

Université de Lille

École doctorale **ED Régionale SMRE 104**

Unité de recherche **Laboratoire de Physique des Lasers, Atomes et Molécules**

Thèse présentée par **Marouane Ayyad**

Soutenue le **21 décembre 2023**

En vue de l'obtention du grade de docteur de l'Université de Lille

Discipline **Physique**

Spécialité **Milieux dilués et optique fondamentale**

Le traitement d'information avec des états chimères optiques

Thèse dirigée par Saliya COULIBALY

Composition du jury

<i>Rapporteurs</i>	Giovanna TISSONI Alejandro M YACOMOTTI	Université Côte d'Azur CNRS-Université de bordeaux	
<i>Examineurs</i>	Valérie VALLET Sylvain BARBAY	CNRS-Université de Lille CNRS-Université Paris-Saclay	président du jury
	Mustapha TLIDI	Université Libre de Bruxelles	
<i>Directeur de thèse</i>	Saliya COULIBALY	MCF HDR à l'Université de Lille	

COLOPHON

Mémoire de thèse intitulé « Le traitement d'information avec des états chimères optiques », écrit par Marouane AYYAD, achevé le 3 février 2024, composé au moyen du système de préparation de document \LaTeX et de la classe `yathesis` dédiée aux thèses préparées en France.

Université de Lille

Doctoral School **ED Régionale SMRE 104**

University Department **Laboratoire de Physique des Lasers, Atomes et Molécules**

Thesis defended by **Marouane Ayyad**

Defended on **December 21, 2023**

In order to become Doctor from Université de Lille

Academic Field **Physics**

Speciality **Dilute media and fundamental optics**

Information processing with optical chimera states

Thesis supervised by Saliya COULIBALY

Committee members

<i>Referees</i>	Giovanna TISSONI	Université Côte d'Azur	
	Alejandro M YACOMOTTI	CNRS-Université de bordeaux	
<i>Examiners</i>	Valérie VALLET	CNRS-Université de Lille	Committee President
	Sylvain BARBAY	CNRS-Université Paris-Saclay	
	Mustapha TLIDI	Université Libre de Bruxelles	
<i>Supervisor</i>	Saliya COULIBALY	HDR Associate Professor at Université de Lille	

Cette thèse a été préparée au

Laboratoire de Physique des Lasers, Atomes et Molécules

2 Avenue Jean Perrin
BP 70478
59658 Villeneuve d'Ascq Cedex
France

☎ (33)(0)3 20 33 60 37

📠 (33)(0)3 21 46 55 75

✉ phlam-contact@univ-lille.fr

Site <https://phlam.univ-lille.fr/>



Remerciements

C'est avec une grande humilité que je tiens à exprimer ma gratitude à l'ensemble de ceux qui ont contribué à la réalisation de cette thèse, une entreprise complexe et exaltante qui m'a poussé au-delà de mes limites personnelles et intellectuelles.

Tout d'abord, je tiens à saluer le défi que représente la réalisation d'une thèse de doctorat. C'est un parcours exigeant qui demande une dévotion inébranlable à la recherche et au savoir. La multidisciplinarité, qui est au cœur de ce travail, a été à la fois une source d'inspiration et de complexité. J'ai été constamment encouragé à explorer de nouveaux horizons, à croiser les disciplines, et à aborder des questions complexes sous des angles différents. C'est précisément cette multidisciplinarité qui, je l'espère, apportera des contributions significatives à notre compréhension du monde.

Je tiens à exprimer ma sincère gratitude envers les membres de jury qui ont accepté de consacrer du temps et de l'expertise à l'évaluation de ce travail. Leurs commentaires et leurs suggestions ont été inestimables, m'aidant à perfectionner cette thèse et à la rendre plus robuste et pertinente. Leur engagement envers l'excellence académique a été une source d'inspiration.

Je tiens à exprimer ma profonde gratitude envers mon encadrant, *Le chef Saliya*, pour son rôle essentiel dans la réalisation de cette thèse. Ses compétences, son soutien et sa compréhension ont été d'une valeur inestimable tout au long de ce parcours. *Saliya* a été bien plus qu'un guide académique. Il a su respecter ma vision scientifique, encourageant la créativité et l'exploration, tout en offrant une orientation précieuse et un cadre solide pour mes recherches. Ses précieux conseils et son expertise ont éclairé le chemin à travers les défis académiques, et j'ai appris énormément à ses côtés. De manière tout aussi importante, je tiens à remercier *Saliya* pour sa compréhension et son soutien envers mes problèmes de santé. Sa flexibilité et son empathie m'ont permis de surmonter les défis personnels tout en poursuivant mon travail de recherche. Cette considération a été un facteur décisif dans la réussite de cette thèse.

Je souhaite également adresser des remerciements chaleureux à mes amis en France et au Bled, dont le soutien indéfectible a été une bouée de sauvetage dans les moments de doute. Leur encouragement constant, leur écoute bienveillante, et leurs sourires m'ont rappelé que, même au milieu des défis quotidiens, la vie est pleine de joie et de sens.

Enfin, ma famille *Angade*, mes parents et mes frères, méritent une mention spéciale. Leurs sacrifices, leur amour et leur soutien inconditionnel ont été le socle sur lequel repose cette réalisation. Leurs encouragements et leurs valeurs m'ont inspiré à poursuivre ce chemin, et je leur suis profondément reconnaissant. À mon neveu *Majd* et mes nièces *Roufayda*,

Jounayda et Nouhayla bien-aimés, voici quelques conseils à emporter avec vous : Cultivez la curiosité, car elle vous mènera vers de nouvelles découvertes passionnantes. Poursuivez vos rêves avec détermination, car ils sont la clé de votre succès. Sachez que l'apprentissage est un voyage sans fin, alors ne cessez jamais d'explorer et de grandir. Enfin, rappelez-vous que la **famille** est un pilier de soutien inestimable, alors chérissez chaque moment passé ensemble. Votre oncle vous aime énormément et croit en votre brillant avenir.

Le traitement d'information avec des états chimères optiques

Résumé

Dans la mythologie grecque, une chimère est une créature fantastique dont certaines parties du corps appartiennent à des animaux différents. Par analogie à cette mythologie, en physique et plus particulièrement dans l'étude des systèmes complexes discrets spatialement étendus, ces états chimères correspondent à la coexistence de deux comportements dynamiques spatio-temporels opposés. La coexistence de deux domaines l'un cohérent et l'autre incohérent dans une chaîne d'oscillateurs non-linéaires couplés en est l'exemple historique, à l'image des différentes parties du corps d'une chimère. Ces auto-organisations spatio-temporelles ont été largement étudiées théoriquement et expérimentalement. Cependant, rares sont les études menées pour explorer les liens entre ce type de dynamique et les automates cellulaires. Ces automates, malgré leur simplicité, possèdent des propriétés dynamiques remarquables et, par conséquent, représentent un des socles de la théorie d'information.

Pour répondre à cette problématique, nous avons considéré des états chimères obtenus dans une chaîne de résonateurs optiques identiques couplés. Ces structures ont alors fait l'objet d'analyses quantitatives et qualitatives par les mêmes outils que ceux utilisés pour caractériser les automates cellulaires. Cela nous a permis de mettre en évidence une dynamique de type automate cellulaire élémentaire cachée dans l'évolution de nos états chimères. Nous avons alors été en mesure de déduire, un ensemble de propriétés en terme de calculabilité, ouvrant des perspectives vers des potentielles applications pour le traitement de l'information.

Par la suite, nous avons utilisé nos états chimères optiques dans le cadre des réseaux de neurones récurrents. Il s'agit d'un nouveau paradigme, qui se distingue par sa grande simplicité, sa rapidité ainsi que son efficacité incontournable dans le traitement de l'information. Cependant, les performances de cette technique d'apprentissage automatique, dépendent notamment du design du réservoir. Nos résultats montrent que l'implémentation de nos états chimères optiques au lieu des réservoirs 'classiques', peut fournir une alternative architecturale prometteuse permettant d'améliorer davantage la vitesse du traitement d'information.

Mots clés : chimera states, spatiotemporal complexity, cellular automata, neural network, reservoir computing

Information processing with optical chimera states

Abstract

In Greek mythology, a chimera is a fantastic creature whose body parts belong to different animals. By analogy with this mythology, in physics and more particularly in the study of spatially extended discrete complex systems, these chimera states correspond to the coexistence of two opposing spatio-temporal dynamic behaviors. The coexistence of two domains, one coherent and the other incoherent in a chain of coupled non-linear oscillators is the historical example, like the different parts of the body of a chimera. These spatio-temporal self-organizations have been widely studied theoretically and experimentally. However, few studies have been carried out to explore the links between this type of dynamics and cellular automata. These automata, despite their simplicity, have remarkable dynamic properties and, consequently, represent one of the foundations of information theory.

To answer this problem, we considered chimera states obtained in a chain of identical coupled optical resonators. These structures were then the subject of quantitative and qualitative analyzes using the same tools as those used to characterize cellular automata. This allowed us to highlight an elementary cellular automaton type dynamic hidden in the evolution of our chimera states. We were then able to deduce a set of properties in terms of computability, opening perspectives towards potential applications for information processing.

Subsequently, we used our optical chimera states in the context of recurrent neural networks. This is a new paradigm, which stands out for its great simplicity, speed and essential efficiency in the processing of information. However, the performance of this machine learning technique depends in particular on the design of the reservoir. Our results show that the implementation of our optical chimeric states instead of 'classic' reservoirs can provide a promising architectural alternative to further improve the speed of information processing.

Keywords: états chimères, complexité spatiotemporelle, automates cellulaires, réseau de neurones, calcul par réservoir

Laboratoire de Physique des Lasers, Atomes et Molécules

2 Avenue Jean Perrin – BP 70478 – 59658 Villeneuve d'Ascq Cedex – France

Introduction générale

Ces dernières décennies, le domaine des systèmes complexes a fait l'objet d'une grande attention par la communauté scientifique. Les exemples souvent cités de systèmes complexes dans la nature et dans la société comprennent la connectivité du cerveau biologique, le système immunitaire, les systèmes biologiques, les cellules, les réseaux métaboliques, les colonies de fourmis, l'internet, les marchés économiques et les réseaux sociaux humains... La science des systèmes complexes explore les mécanismes à l'œuvre dans de vastes ensembles d'entités élémentaires, où le comportement global émerge de manière non-réductionniste en raison des interactions locales non-linéaires entre leurs composants individuels [1-3]. Cette approche vise à une compréhension fondamentale des origines des comportements collectifs qui découlent de ces interactions entre des entités simples. La complexité du comportement global du système est généralement caractérisée en termes des *patterns* qu'il forme, du traitement de l'information qu'il accomplit et du degré d'adaptabilité qu'il manifeste. Pour caractériser ces comportements auto-organisés, les scientifiques des systèmes complexes utilisent des outils issus de diverses disciplines, notamment la dynamique non-linéaire, la théorie de l'information, la théorie de calculabilité et la biologie évolutive. Le domaine des systèmes complexes cherche à expliquer et à découvrir des lois communes pour les comportements émergents et auto-organisés observés dans des systèmes complexes interdisciplinaires [4, 5]. De nombreux scientifiques croient également que la découverte de tels principes généraux sera essentielle pour créer la vie et l'intelligence artificielles [6].

Effectivement, l'hypothèse du *edge of chaos* présente un socle axial dans l'interaction fructueuse entre la dynamique non-linéaire des systèmes complexes, la théorie de calculabilité et la théorie d'information. LANGTON et PACKARD [7, 8] sont les pionniers dans l'édification de l'hypothèse du *edge of chaos*. Les études citées suggèrent que les systèmes complexes atteignent des capacités de calcul optimales si leur dynamique se situe à la frontière d'une transition entre l'ordre et le chaos. Les initiateurs de cette hypothèse ont intégré la théorie de l'information comme un instrument efficace pour mener leurs recherches sur les automates cellulaires. Ces automates, en tant que systèmes dynamiques discrets abstraits,

se révèlent être des outils précieux pour la modélisation de systèmes complexes présents dans la nature [9], et ils ont fourni une plateforme stimulante pour explorer la relation entre la calculabilité et les comportements complexes auto-organisés.

Ce concept s'est répandu pour la première fois au début de l'informatique, où il a été remarqué que des complexités auto-organisées émergeaient spontanément dans les systèmes artificiels. Du point de vue biologique, l'hypothèse du *bord du chaos* est très prometteuse pour les neurosciences [10]. Une bonne compréhension des propriétés des transitions de phase à l'œuvre dans le cerveau établirait un cadre sur lequel on peut construire des modèles de fonctionnement cérébral. En effet, les aspects computationnels intrinsèques à la région du *bord du chaos* offrent une bonne compréhension de la cognition et le traitement d'information dans le cerveau [11-15]. Plus loin, les enseignements tirés de l'étude des différentes signatures du *bord du chaos* dans le cerveau peuvent également trouver des applications dans l'intelligence artificielle [16-18], où l'étude des transitions de phase, de l'auto-organisation et de la complexité demeure un sujet émergent et peut à son tour aider à guider l'utilisation de réseaux de neurones artificiels dans la modélisation du traitement de l'information au sein du cerveau.

L'hypothèse du *bord du chaos* a suscité davantage d'intérêt avec l'avènement du reservoir computing [19-21], un type de réseau neuronal récurrent et dynamique qui est capable d'apprendre des tâches complexes et qui, comme l'ont montrées plusieurs études, fonctionne de manière optimale au bord du chaos [21-24]. En outre, d'autres études ont démontré que cette région était soutenue par une maximisation simultanée des propriétés de stockage et de transfert d'informations à travers le réseau [25].

Dans ce travail de thèse, nous nous sommes intéressés à une forme exotique d'auto-organisation qui peut être accueillie par les systèmes complexes spatialement étendus, connue sous le nom d'*état chimère*. Ce phénomène se distingue par la présence simultanée de deux comportements incongrus au sein du même système. Notre attention est consacrée à des états chimères observés dans un réseau de guides d'ondes optiques 1D couplés localement [26]. La présence des patterns triangulaires de type SIERPINSKI dans ces états chimères optiques, contribue à la richesse des signatures d'auto-organisation et de complexité.

L'objectif de cette thèse est d'explorer les capacités computationnelles de nos états chimères optiques. En d'autres termes, notre but est d'investiguer le rôle de l'auto-organisation manifestée dans nos états chimères optiques (à travers ses différentes empreintes), dans la maximisation du traitement d'information. À cette fin, nous avons adopté la feuille de route suivante :

- La mise en évidence d'une analogie qualitative et quantitative entre nos états chimères optiques et un système calculatoire par excellence qui réside en les automates cellulaires élémentaires.
- L'analyse structurelle/fonctionnelle de l'implémentation de nos états chimères optiques dans l'architecture du reservoir computing.

Cette thèse est structurée de la manière suivante :

- [Le chapitre 1](#) commence par présenter les concepts les plus essentiels, en mettant l'accent sur ceux qui sont pertinents pour comprendre comment l'auto-organisation se manifeste dans les systèmes complexes non-linéaires. Ensuite, nous abordons une forme fascinante d'auto-organisation : les états chimères. Cela nous permettra enfin d'introduire nos propres états chimères optiques, qui seront au cœur des études présentées dans ce manuscrit.
- [Le chapitre 2](#) présente l'hypothèse du 'bord du chaos', qui établit un lien entre la complexité et la calculabilité, tout en examinant en détail les opérations fondamentales du traitement de l'information. Ensuite, nous fournirons une définition des automates cellulaires, qui représentent un excellent moyen concret d'explorer les concepts de calculabilité.
- [Le chapitre 3](#) a pour objectif d'examiner comment nos états chimères optiques peuvent démontrer des propriétés calculatoires similaires à celles des automates cellulaires élémentaires. Dans cette optique, nous cherchons à établir une analogie entre nos états chimères optiques et les automates cellulaires élémentaires, à la fois qualitativement et quantitativement, en utilisant des outils de la théorie de l'information, notamment les concepts d'entropie.
- [Le chapitre 4](#) vise à fournir les fondements théoriques nécessaires à la compréhension du concept de 'Reservoir Computing'. Nous commençons par une recette de base pour la création de réservoirs, ainsi qu'un aperçu des hyper-paramètres essentiels pour régir leur dynamique et optimiser leurs performances. Ensuite, nous explorons les caractéristiques souhaitables d'un réservoir de qualité.
- [Le chapitre 5](#) pour but d'évaluer les performances de nos réservoirs basés sur des états chimères en ce qui concerne l'observabilité, l'inférence et la prédiction de séries chaotiques, ainsi que la reproduction complète d'attracteurs chaotiques. Ensuite, nous approfondirons notre analyse structurelle et dynamique grâce à une approche multidisciplinaire qui connecte les domaines des neurosciences et de la physique des systèmes complexes.

Etat chimère optique

1.1 Introduction

Au cours des quatre dernières décennies, un nouveau paradigme pour la recherche scientifique a émergé : il s'agit de *la complexité*. La science classique, telle qu'illustrée par la mécanique newtonienne, est essentiellement réductionniste : elle réduit tous les phénomènes complexes à leurs composants les plus simples, puis tente de décrire ces composants de manière complète, objective et déterministe [4, 5]. La philosophie de la complexité est que cela est en général impossible. En effet, les *systèmes complexes*, tels que les organismes, les sociétés ou internet, ont des propriétés émergentes qui ne peuvent être réduites aux simples propriétés de leurs parties. De plus, le comportement de ces systèmes présente des aspects intrinsèquement imprévisibles, et ne peut être décrit de manière complète. Au mieux, nous pouvons trouver certaines régularités statistiques dans leurs caractéristiques quantitatives, ou comprendre leur comportement qualitatif à travers des modèles et des simulations informatiques.

Bien que ces qualités 'traditionnelles' de réductionnisme et de prévisibilité qui manquent aux systèmes complexes, ces systèmes présentent également un certain nombre de caractéristiques étonnamment positives, telles que la flexibilité, l'autonomie et la robustesse, qui manquent aux systèmes traditionnels. Ces qualités peuvent toutes être vues comme des aspects du processus d'*auto-organisation* qui caractérise les systèmes complexes : *ces systèmes s'organisent spontanément pour mieux faire face aux diverses perturbations internes et externes*. Cela leur permet d'évoluer et de s'adapter à un environnement en constante évolution.

Les processus d'auto-organisation créent littéralement de l'ordre à partir du désordre [4].

Ils sont responsables de la plupart des modèles, des structures et des arrangements ordonnés ("*patterns*") que nous trouvons dans le monde naturel, et beaucoup de ceux dans les domaines de la physique, de la biologie et de la sociologie [27-31].

Dernièrement, les systèmes complexes ont prouvé leurs capacités d'accueillir une forme exotique d'auto-organisation, connue sous le nom d'*état chimère* [32, 33]. En effet, ce phénomène est caractérisé par la coexistence de deux comportements incongrus au sein du même système.

Le présent chapitre tentera, d'abord, d'introduire les concepts importants de manière simple et cohérente, en mettant l'accent sur ceux qui peuvent nous aider à comprendre l'apparition du phénomène d'auto-organisation dans les systèmes complexes non-linéaires. Ensuite, nous introduisons les états chimères, qualifiés comme étant un type fascinant d'auto-organisation. Ce qui nous permettra, finalement, de présenter nos états chimères optiques, qui feront l'objet des études établies dans le présent manuscrit.

1.2 Système complexe auto-organisé

1.2.1 Systèmes complexes

Il n'existe pas de définition généralement admise de la complexité [1] : différents auteurs ont proposé des dizaines de mesures ou de conceptions, dont aucune ne rend compte de tous les aspects intuitifs du concept. Par exemple, la mesure la plus connue, la complexité de KOLMOGOROV, qui est à la base de la théorie algorithmique de l'information, définissant la complexité d'une chaîne de caractères comme la longueur du programme le plus court qui peut générer cette chaîne [34]. Cependant, cela implique que les chaînes aléatoires sont extrêmement complexes, puisqu'elles ne permettent aucune description plus courte que la chaîne elle-même. Cela contredit notre intuition que les systèmes aléatoires ne sont pas vraiment complexes. Un certain nombre de grandeurs plus variées de cette définition ont été proposées pour résoudre ce problème, mais elles souffrent toujours du fait qu'elles ne s'appliquent qu'aux chaînes, pas aux systèmes du monde réel. De plus, il a été prouvé que la description la plus courte possible est en général non calculable, ce qui implique qu'on ne peut jamais être sûr d'avoir réellement déterminé la véritable complexité d'une chaîne.

En dépit de ces problèmes fondamentaux dans la formalisation de la notion de complexité, il existe un certain nombre de caractéristiques plus intuitives des systèmes complexes qui apparaissent à maintes reprises dans les différentes tentatives de caractérisation [5]. Une idée plus ou moins universellement admise est que la complexité doit se

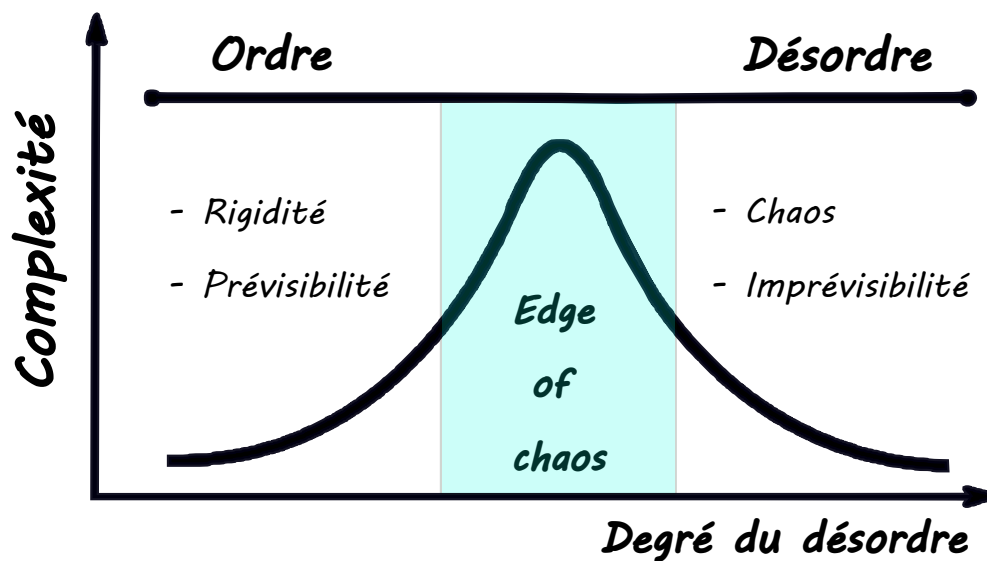


FIGURE 1.1 – Courbe schématique montrant la position du régime "Edge of chaos" dans la transition entre l'ordre et le désordre. Cette hypothèse stipule que la complexité ainsi que la richesse d'information, sont maximisées dans cette région.

situer entre l'ordre et le désordre : les systèmes complexes ne sont ni réguliers et prévisibles (comme l'arrangement rigide et 'figé' des molécules dans un cristal), ni aléatoires et chaotiques (comme la configuration en constante évolution des molécules dans un gaz). Ils présentent un mélange des deux dimensions, étant à peu près prévisibles à certains égards, surprenants et imprévisibles à d'autres. Cette position intermédiaire entre régularité et chaotité, est parfois appelée "Edge of Chaos" (Fig. 1.1). Un certain nombre de théoriciens ont proposé que cet équilibre précaire est précisément ce qui est nécessaire à l'adaptation, à l'auto-organisation et à la vie, et que les systèmes complexes tendent à évoluer spontanément vers ce "bord" [6].

Une autre caractéristique fondamentale est que les systèmes complexes sont constitués de nombreuses (ou au moins plusieurs) parties qui sont connectées via leurs interactions. Leurs composants sont à la fois distincts et connectés, à la fois autonomes et, dans une certaine mesure, mutuellement dépendants. Une dépendance complète impliquerait l'ordre, comme dans un cristal où l'état d'une molécule détermine l'état de toutes les autres. Une indépendance complète impliquerait un désordre, comme dans un gaz où l'état d'une molécule ne vous donne aucune information sur l'état des autres molécules.

Les composants d'un système complexe sont le plus souvent modélisés comme des agents, c'est-à-dire des systèmes individuels qui agissent sur leur environnement en réponse aux événements qu'ils subissent. Les conditions environnementales auxquelles un agent réagit sont

normalement affectées par l'activité d'autres agents. Par conséquent, une action d'un agent déclenchera en général d'autres actions d'un ou plusieurs autres agents, mettant éventuellement en mouvement une chaîne d'activité étendue qui se propage d'agent en agent à travers le système. De telles interactions sont initialement locales : elles n'affectent au départ que les agents du voisinage immédiat de l'acteur initial. Cependant, leurs conséquences sont souvent globales, affectant le système d'agents dans son ensemble, comme une ondulation produite par un objet qui perturbe localement la surface de l'eau, mais s'élargit ensuite pour englober tout l'étang.

1.2.2 Non linéarité

Les processus dans les systèmes complexes sont souvent non linéaires : *leurs effets ne sont pas proportionnels à leurs causes*. Lorsque les effets sont plus grands que les causes, on peut dire qu'il y a amplification ou rétroaction positive : d'abord les petites perturbations se renforcent pour devenir de plus en plus intenses. Un exemple est la propagation d'une maladie, où une seule infection peut éventuellement se transformer en une pandémie mondiale. Lorsque les effets sont inférieurs aux causes, il y a amortissement ou rétroaction négative : les perturbations sont progressivement supprimées, jusqu'à ce que le système revienne à son état d'équilibre.

Les interactions avec rétroaction positive sont très sensibles à leurs conditions initiales : un changement dans cette condition peut être si petit qu'il est intrinsèquement indétectable, mais il peut entraîner un résultat radicalement modifié. C'est ce qu'on appelle l'*effet papillon*, après l'observation qu'en raison de la non-linéarité du système d'équations régissant la météo, le battement d'ailes d'un papillon au Brésil peut provoquer (rétroaction positive) ou empêcher (rétroaction négative) un ouragan à Texas. La non-observabilité des perturbations initiales rend l'issue à priori imprévisible, même si la dynamique du système était parfaitement déterministe : aucun système de surveillance météorologique ne peut être si précis qu'il capte tous les mouvements d'ailes de papillon... C'est pourquoi les prévisions météorologiques ne peuvent pas être vraiment fiables, surtout à plus long terme. Une rétroaction positive amplifiera de petites fluctuations aléatoires en oscillations imprévisibles, rendant le comportement global du système : chaotique. C'est ainsi que le *chaos déterministe* a été découvert pour la première fois par un certain nombre de chercheurs comme le physicien de l'atmosphère EDWARD LORENTZ [35].

Malgré l'omniprésence des fluctuations, la plupart des systèmes qui nous entourent apparaissent relativement stables et prévisibles. Cela est dû à la présence d'une rétroaction

négative, qui supprime les effets de telles fluctuations. Cependant, alors que la rétroaction négative rend un système plus prévisible, elle le rend également moins contrôlable : si nous essayons de changer l'état d'un tel système, nous pouvons constater que nos changements sont contrecarrés, et que quoi que nous fassions, le système revient toujours à son état initial propre (état d'équilibre "préféré").

La dynamique des systèmes complexes présente généralement une combinaison de rétroactions positives et négatives, de sorte que certains changements sont amplifiés et d'autres atténués. Cela rend le comportement global du système imprévisible. De plus, ces systèmes sont normalement ouverts, ce qui signifie qu'ils échangent de la matière, de l'énergie et/ou des informations avec leur environnement plus large. Par exemple, un écosystème dépend du climat et de la quantité de lumière solaire, de pluie et de chaleur qu'il produit. Ces flux entrants et sortants rendent la dynamique encore plus compliquée, puisque nous ne pouvons pas connaître tous les événements externes susceptibles d'affecter le système. Par exemple, un écosystème florissant peut soudainement s'effondrer à cause de l'invasion d'un ravageur étranger (des organismes animaux qui attaquent les plantes cultivées, ou les récoltes stockées). De plus, l'apport d'énergie tend à alimenter les processus d'amplification, de sorte qu'ils n'atteignent jamais l'équilibre qui suivrait l'épuisement des ressources.

Pour les raisons ci-dessus, dans les systèmes complexes non linéaires, simplifier le modèle mathématique en utilisant des approximations est problématique. La manière courante d'approximer l'effet d'interactions complexes en le réduisant au "champ moyen" (c'est-à-dire l'effet moyen de nombreuses actions discrètes effectuées par des agents indépendants) peut en fait conduire à des erreurs fondamentales. Par exemple, une équation différentielle représentant l'effet de «champ moyen» peut prédire qu'une certaine perturbation s'éteindra parce qu'elle est trop petite, tandis qu'une simulation numérique des agents individuels constate que son effet est amplifié par une rétroaction positive jusqu'à ce qu'elle domine le système [30].

En raison de ces difficultés intrinsèques avec les modèles mathématiques, les chercheurs dans le domaine de la complexité préfèrent généralement les simulations numériques, qui, tout en étant bien sûr des approximations, sont plus faciles à manipuler, de sorte que davantage des facteurs et des variations du modèle peuvent être explorés. Ici, l'évolution du système est tracée pas à pas en appliquant de manière itérative les règles qui régissent les interactions des agents, générant ainsi les états ultérieurs du système. Afin d'explorer les comportements possibles du système, de nombreuses manipulations différentes - avec des conditions initiales différentes ou des variations aléatoires au cours du processus/la simulation sont effectuées. Les principales valeurs des variables pour chaque exécution sont

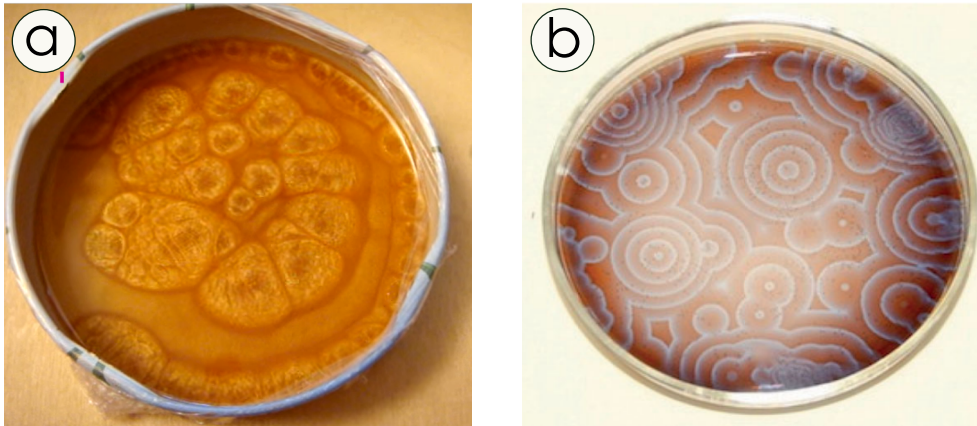


FIGURE 1.2 – (a) : Cellules de convection de RAYLEIGH–BÉNARD. (b) : Des *patterns* de la réaction de BELOUSOV–ZHABOTINSKY [29].

collectées. Ces résultats sont ensuite analysés statistiquement pour découvrir des tendances récurrentes.

Cela produit parfois des résultats très robustes, en ce sens que toutes les exécutions, aussi différentes soient-elles dans leur comportement initial, semblent finalement converger vers le même type de modèle stable. Dans la majorité des cas, les résultats peuvent être classés en un nombre relativement restreint de catégories distinctes. Cela fournit aux chercheurs une image qualitative des résultats les plus probables et un aperçu des facteurs qui favorisent un résultat plutôt qu'un autre (*diagramme de bifurcation*). Ce n'est qu'exceptionnellement qu'aucun schéma clair ne peut être discerné dans les résultats des différentes simulations. La raison pour laquelle les systèmes complexes, malgré leur imprévisibilité intrinsèque, ont tendance à s'installer dans un ensemble relativement restreint de comportements reconnaissables et leur tendance inhérente à *s'auto-organiser*.

1.2.3 Auto-organisation

Le paradoxe thermodynamique

L'émergence spontanée de nouvelles structures est facile à observer, tant en laboratoire que dans notre monde de tous les jours. L'exemple le plus courant est peut-être le phénomène de BÉNARD [27], l'apparition d'un motif de cellules hexagonales ou de rouleaux parallèles dans un liquide chauffé par le bas (Fig. 1.2a). Des exemples plus compliqués sont certaines réactions chimiques, comme la réaction de BELOUSOV-ZHABOTINSKY ou le "Brus-selator" [28], où il suffit de pomper constamment un certain nombre d'ingrédients dans une solution pour voir se développer des spirales éblouissantes de couleur pulsante (Fig. 1.2b).

Ce que ces exemples ont en commun est l'auto-organisation : l'apparition d'une structure ou d'un pattern sans qu'un agent extérieur l'impose. Mais cela ne correspond pas à notre image intuitive du monde. Si un système, comme une fleur ou un bâtiment fait preuve d'organisation, nous avons tendance à supposer que quelqu'un ou quelque chose doit avoir arrangé les composants dans cet ordre particulier. Si nous ne pouvons pas trouver une personne responsable de la conception, nous sommes tentés de l'attribuer à une force intelligente inconnue. Cette intuition est confirmée par *la deuxième loi de la thermodynamique*, qui stipule que dans un système laissé à lui-même l'entropie (désordre) ne peut qu'augmenter, pas diminuer. Ainsi, la première étape pour expliquer l'auto-organisation doit être de la concilier avec la thermodynamique. Dans le cas où le système auto-organisé n'atteint pas l'équilibre, l'explication est moins évidente. Le thermodynamicien belge ILYA PRIGOGINE a reçu un prix Nobel pour ses travaux, à partir des années 1950, sur ce problème, où il a étudié ce qu'il appelle les *structures dissipatives*. Ce sont des motifs observés dans les systèmes tels que les cellules de BÉNARD ou le Brusselator, qui présentent une auto-organisation dynamique. De telles structures sont nécessairement des systèmes ouverts : de l'énergie et/ou de la matière y circulent. Le système génère en permanence de l'entropie, mais cette entropie est activement dissipée, ou exportée, hors du système. Ainsi, il parvient à accroître sa propre organisation au détriment de l'ordre dans l'environnement. Le système contourne la deuxième loi de la thermodynamique simplement en se débarrassant de l'excès d'entropie. Les exemples les plus évidents de tels systèmes dissipatifs sont les organismes vivants. Les plantes et les animaux absorbent de l'énergie et de la matière (la lumière et la nourriture) sous une forme à faible entropie. Ils la réexportent (les déchets) sous une forme à haute entropie. Ceci leur permet de réduire leur entropie interne, contrecarrant ainsi la dégradation impliquée par la deuxième loi de la thermodynamique.

La première conférence sur les systèmes auto-organisés, tenue en 1959 à Chicago, a été organisée par le même groupe multidisciplinaire de scientifiques qui avait fondé la discipline de la cybernétique. Cette discipline représente la science qui étudie les mécanismes de communication et de régulation dans les machines et chez les êtres vivants. Le cybernéticien britannique W. ROSS ASHBY a noté qu'un système dynamique, indépendamment de son type ou de sa composition, tend toujours à évoluer vers un état d'équilibre, ou ce qu'on appelle dans le jargon de la dynamique non linéaire "*un attracteur*". Cela équivaut à une autre vision du phénomène de l'auto-organisation. L'équilibre qui en résulte peut être interprété comme un état où les différentes parties du système s'adaptent mutuellement. L'idée est très simple : plus un système est amené à se déplacer largement dans son espace d'états, plus il finira rapidement dans un attracteur. PRIGOGINE a proposé le principe connexe de

"l'ordre par les fluctuations" [4]. Les systèmes non linéaires ont en général plusieurs attracteurs. Lorsqu'un système réside entre des attracteurs, ce sera en général une variation fortuite, appelée "fluctuation" en thermodynamique, qui le poussera dans un autre de ses attracteurs.

Caractéristiques d'un système auto-organisé

Différentes études ont été menées afin de décortiquer un certain nombre de propriétés fondamentales ou "signatures", qui distinguent les systèmes auto-organisés des systèmes traditionnels étudiés en science. Nous allons essayer d'énumérer cette variété de caractéristiques :

- **Systèmes complexes** : L'auto-organisation a lieu dans un système complexe, où les interactions entre ses parties ne sont pas organisées de manière centralisée mais de manière distribuée et il est difficile de prédire l'ensemble du comportement même si l'on connaît dans une large mesure le nombre de ses éléments et les liens entre eux.
- **Autonomie, Adaptabilité et Robustesse** : La localité des interactions permet qu'une perturbation passe d'une région à une autre, tout en passant d'abord par toutes les régions intermédiaires. Le processus est véritablement collectif, c'est-à-dire parallèle et réparti sur tous les agents. Cela rend l'auto-organisation qui en résulte intrinsèquement robuste et résistante aux perturbations.
- **Non-linéarité, rétroaction et brisure de symétrie** : Dans les systèmes auto-organisés, la relation de cause à effet n'est pas linéaire : de petites causes peuvent avoir de grands effets, et de grandes causes peuvent avoir de petits effets. Cette non-linéarité peut être comprise à partir de la relation de rétroaction qui existe entre les composants de ce type de systèmes. Il y aura plusieurs boucles de rétroaction positives et négatives imbriquées, de sorte que les changements dans certaines directions sont amplifiés tandis que les changements dans d'autres directions sont supprimés. Cela peut conduire à un comportement très compliqué et difficile à prévoir. Cependant, après l'auto-organisation, une direction ou une configuration domine toutes les autres, et donc la symétrie est brisée. Ce phénomène est étroitement lié à la notion de "bifurcation". Ce choix que prend le système, est influencé par la valeur critique des paramètres de contrôle.
- **Système ouvert et loin de l'équilibre thermodynamique** : est plus fragile et sensible aux changements de l'environnement, mais aussi plus dynamique et capable de réagir. Les changements influencent la structure organisée. Afin de maintenir cette

structure, il doit y avoir une dynamique constante qui gère ces changements. En d'autres termes, le système doit s'éloigner de l'équilibre pour maintenir la structure.

1.3 Etat chimère

1.3.1 Bref historique

Comme constaté dans la section précédente, dans les systèmes complexes où un grand nombre d'éléments sont couplés ensemble, les comportements collectifs des unités en interaction peuvent être non triviaux. La dynamique collective des systèmes complexes a suscité une grande attention au cours des dernières décennies [36, 37]. Les systèmes d'oscillateurs couplés sont des modèles paradigmatiques pour révéler et comprendre le mécanisme de divers phénomènes collectifs différents. Par exemple des oscillateurs de KURAMOTO [38], ont été utilisés pour modéliser une grande variété de formation de motifs dans des systèmes spatio-temporels, tels que des réseaux de neurones, des fluides convectifs, des réseaux laser et des oscillateurs biochimiques couplés. Ces systèmes sont connus pour présenter de riches dynamiques collectives, allant de la synchronisation, l'intermittence au chaos spatio-temporel et à la turbulence [36, 39, 40].

Le pouvoir de synchronisation des réseaux dynamiques d'oscillateurs couplés peut se décomposer en deux ou plusieurs domaines synchronisés et désynchronisés et un tel comportement spatio-temporel complexe fascinant est appelé *état chimère* [32, 33] (Fig. 1.3a).

Dans la mythologie grecque, une chimère est une créature fantastique dont certaines parties du corps appartiennent à des animaux différents. Ces organismes hétérogènes réussissant ainsi à coexister au sein d'un même animal (Fig. 1.3b). Les états chimères n'ont cessé de susciter un intérêt considérable depuis deux décennies. Leur ressemblance avec plusieurs phénomènes omniprésents dans la nature, tels que le sommeil uni-hémisphérique chez certains mammifères, les crises d'épilepsie et la schizophrénie, a contribué à la compréhension de certains de leurs mécanismes de base [41-43].

Par analogie à cette mythologie, en physique et plus particulièrement dans les systèmes complexes discrets spatialement étendus, ces états chimères correspondent à la coexistence de deux comportements dynamiques spatio-temporels opposés. La coexistence de deux domaines l'un cohérent et l'autre incohérent dans une chaîne d'oscillateurs non-linéaires couplés en est l'exemple historique, à l'image des différentes parties du corps d'une chimère (Fig. 1.3b).

En 2002, le physicien japonais YOSHIKI KURAMOTO et son collaborateur DORJSU-

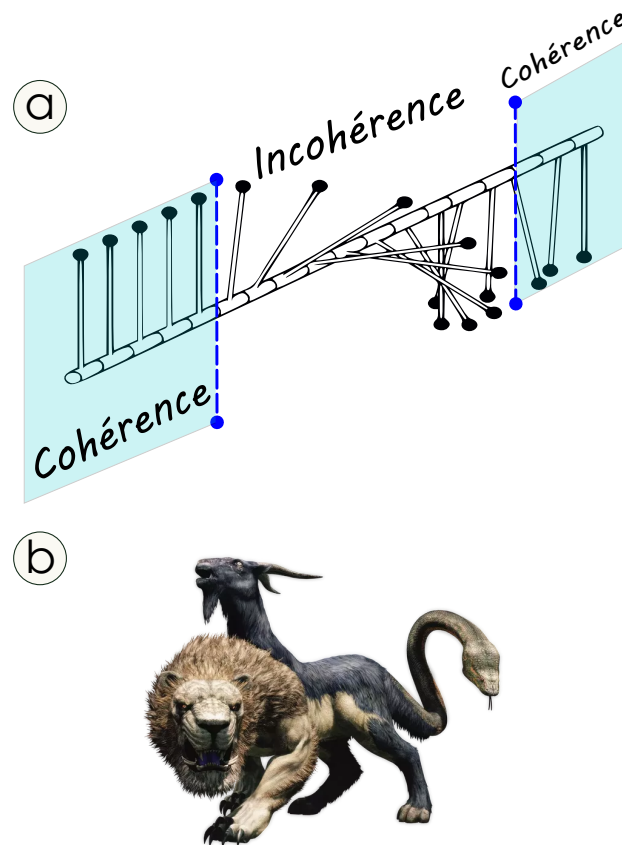


FIGURE 1.3 – (a) : Une chaîne d'oscillateurs couplés schématisant la coexistence de domaine : un cohérent et un incohérent. (b) : Illustration de la créature mythique "chimère".

REN BATTOGTOKH ont inauguré les investigations théoriques pionnières sur ces auto-organisations spatio-temporelles [32]. Ces investigations ont été menées sur un anneau d'oscillateurs identiques ayant la phase couplée non localement. Ce système est décrit par l'équation de phase suivante :

$$\frac{\partial \phi(x, t)}{\partial t} = \omega - \int G(x - x') \sin \left(\phi(x, t) - \phi(x', t) + \alpha \right) dx' . \quad (1.1)$$

Ici $\phi(x, t)$ est la phase instantanée de l'oscillateur à la position x à l'instant t et satisfait une condition aux limites périodiques, donc cette équation peut décrire un anneau d'oscillateurs couplés. Comme les oscillateurs sont réglés pour être identiques, la valeur de la fréquence ω et du déphasage α , sont choisis constants et indépendants de l'emplacement spatial de l'oscillateur. $G(x, x')$ décrit la fonction du couplage non local entre les oscillateurs.

Remarquablement, sous certaines conditions initiales, des oscillateurs identiques couplés peuvent se comporter différemment les uns des autres (Fig. 1.4). Autrement dit, certains des

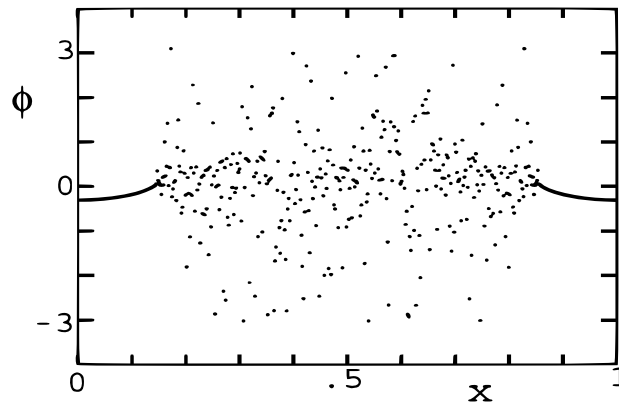


FIGURE 1.4 – Répartition spatiale instantanée des phases [32], obtenue à partir de l'Eq. 1.1. On constate une cohérence et une incohérence : la région de cohérence s'étend de 0 à 0.1 et de 0.9 à 1.

oscillateurs pouvaient se synchroniser tandis que d'autres restaient incohérents. Ce n'était pas un état transitoire, mais apparemment un phénomène persistant stable combinant certains aspects de l'état synchrone avec d'autres aspects de l'état incohérent.

En 2004, ABRAMS et STROGATZ ont nommé ce fascinant paradigme spatio-temporel "état chimère", inspiré du monstre grecque mythologique, et ont introduit quelques explications théoriques à l'existence d'un tel comportement [33].

Des efforts considérables ont été conduits afin d'explorer l'émergence des états chimères et pour une meilleure compréhension des mécanismes de rupture de symétrie spontanée dans les modèles spatio-temporels concernés [44, 45]. Ces études théoriques ont été focalisées sur :

- **la géométrie du réseau** : Initialement, les états chimères étaient étudiés dans un anneau d'oscillateurs de phase couplés non localement [32, 33, 46, 47], puis dans le cas d'un plan bidimensionnel [48, 49], d'un tore [50], et d'une sphère [45].
- **la nature du réseau** : D'autres travaux ont montré que les états chimères ne se limitent pas aux oscillateurs de phase de Kuramoto, mais peuvent également être trouvés dans une grande variété de systèmes différents, par exemple, cartes couplées chaotiques discrètes dans le temps [51], modèles chaotiques continus en temps [52], systèmes neuronaux et réseaux [53], systèmes d'oscillateurs de ROESSLER et LORENZ [54] et les systèmes à retard [55].
- **l'effet de la topologie du couplage** : Au début, ces états ont été observés dans des réseaux couplés non localement et après cela, ils ont également été trouvés dans le cadre du couplage global [56-58], dans des réseaux couplés localement [59-61] (Fig. 1.5).
- **l'effet des conditions initiales** : La multistabilité est un phénomène très im-

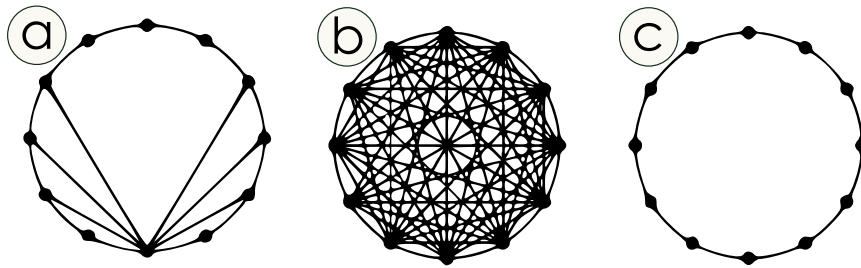


FIGURE 1.5 – Topologies du couplage : (a) : non local. (b) : global. (c) : local.

portant dans les systèmes dynamiques non linéaires [62, 63]. Ce phénomène est caractérisé par la coexistence de deux ou plusieurs attracteurs. Cette coexistence a motivé plusieurs études sur les effets de différentes conditions - de différents bassins d'attraction - sur l'émergence des états chimères [64, 65].

Outre les travaux mentionnés ci-dessus principalement basés sur des simulations et analyses théoriques, récemment, des états chimères ont été démontrés expérimentalement dans un certain nombre de systèmes de laboratoire. Il s'agit notamment de populations d'oscillateurs chimiques couplés [66], méta-matériaux supraconducteurs [67], un réseau d'oscillateurs mécaniques [68] et dans des oscillateurs électroniques à retard non linéaires [69]. Ces expériences indiquent les potentielles applications des états chimères dans la vie réelle.

Dans les systèmes optiques, les observations expérimentales de ce type de patterns auto-organisés sont multiples. En utilisant un dispositif optoélectronique à rétroaction retardée, l'observation des états chimères résultait d'un couplage non linéaire entre la longueur d'onde du laser et le courant électrique de polarisation [70]. Dans [71], les états chimères résultent du verrouillage-déverrouillage de la phase entre plusieurs modes longitudinaux dans une diode laser couplée à un absorbeur saturable non linéaire. Il a été démontré que ces auto-organisations fascinantes peuvent être observées et contrôlées dans la dynamique de polarisation de la lumière dans un système de laser diodes grâce à la rétro-action [72].

1.3.2 Nos états chimères optique (OCS)

Ces dernières années, les systèmes optiques discrets, par exemple un réseau de guides d'ondes/microcavités couplés (Fig. 1.6), ont suscité une attention croissante, et leurs propriétés non linéaires ont été étudiées [73]. Ces systèmes dissipatifs peuvent confiner la lumière sous forme de structures localisées comme les solitons discrets [74, 75]. En propagation libre, le confinement spatial est attribué à l'équilibre entre la diffraction et la non linéarité d'une part, et d'autre part, entre la dissipation d'énergie due aux miroirs qui doit

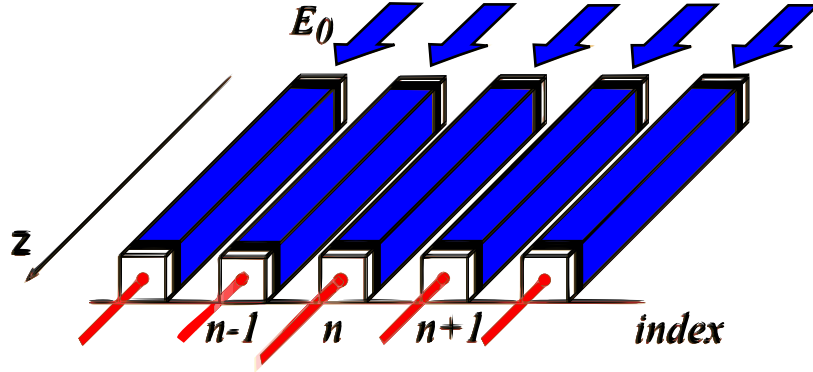


FIGURE 1.6 – Représentation schématique d'un réseau de guides d'ondes couplés pilotés par un champ électrique externe d'intensité E_0^2 avec une fréquence de fonctionnement ω proche de la fréquence de résonance de la cavité ω_0 .

être compensée par l'injection optique. Nous considérons un réseau de guides d'ondes non linéaires faiblement couplés, qui est sujet à une injection d'un champ de commande externe E_0 . Les résonateurs sont alors constitués d'un milieu Kerr auto-focalisant. La dynamique spatio-temporelle de ce simple dispositif optique est décrite théoriquement par le modèle discret de Lugiato-Lefever [74, 75]. Le champ intracavité obéit :

$$\partial_T \psi_n = E_0 - (1 + i\Delta)\psi_n + i|\psi_n|^2\psi_n + iC(\psi_{n+1} - 2\psi_n + \psi_{n-1}) \quad (1.2)$$

Où ψ_n est l'enveloppe à variation lente du champ électrique circulant dans le n ème micro-résonateur. $\Delta \equiv \omega - \omega_0$ représente le désaccord entre la fréquence de résonance ω_0 de la cavité et la fréquence d'entrée ω . C symbolise le couplage avec les plus proches voisins (cavités). Le temps d'évolution $t = T\tau_{ph}$ est mesuré dans l'unité de durée de vie du photon τ_{ph} (est normalisé au temps de décroissance de la cavité). Toutes les simulations numériques ont été réalisées à l'aide d'un code aux différences finies avec l'algorithme du quatrième ordre de *Runge - Kutta* et les conditions aux limites de *Neumann*, compatibles avec une chaîne finie de guides d'ondes couplés.

Dans la limite continue où $C \rightarrow \infty$, des investigations théoriques [76-78] et expérimentales [79] ont révélé une coexistence entre une solution constante et une solution de type chaos spatio-temporel dans un certain régime de paramètres incluant E_0 et Δ . Cependant, ces 2 solutions localisées ne sont pas observées. C'est-à-dire qu'en augmentant la constante de couplage du plus proche voisin C , et en fonction des autres paramètres, la structure localisée chaotique se rétrécit ou s'étend de sorte que le système ne présente qu'un seul type d'état étendu en tant qu'un équilibre, étant soit un état homogène, soit un état

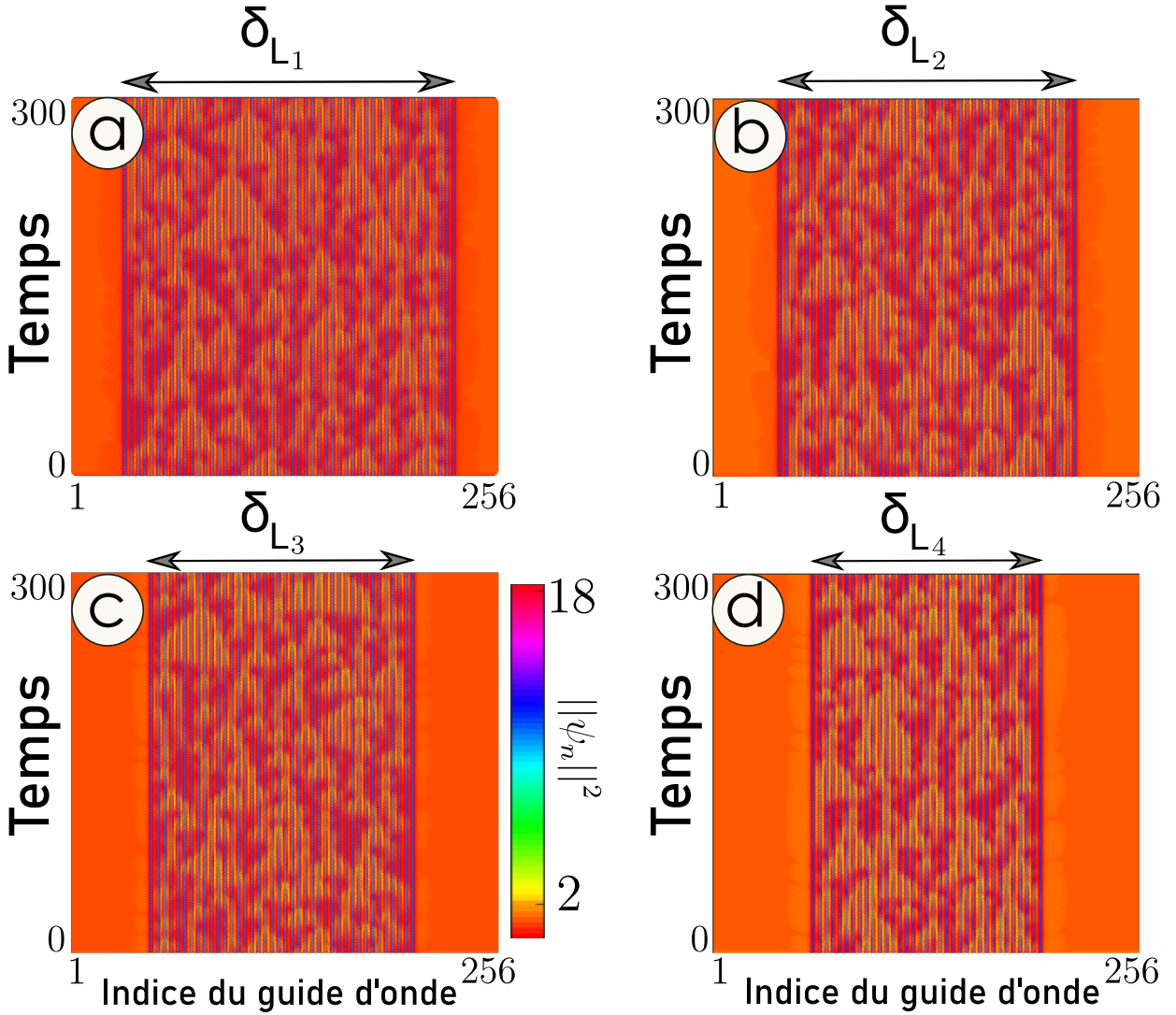


FIGURE 1.7 – Évolutions spatio-temporelles de l'intensité $\|\psi_n(t)\|^2$ pour des tailles différentes d'OCS ($\delta_{L1}, \delta_{L2}, \delta_{L3}$ et δ_{L4}), obtenues à partir de simulations numériques du modèle Eq. 1.2 : en prenant $C = 1.9$, $\Delta = 7$ et $E_0 = 6$.

spatio-temporel chaotique.

En effet, l'interaction des fronts est attrayante et alors les états localisés spatio-temporels sont instables. Cependant, en considérant une constante de couplage faible du plus proche voisin [$C \sim O(1)$], il est possible de stabiliser des états localisés spatio-temporels sous la forme d'un ÉTAT DE CHIMÈRE (Fig. 1.7). Il a été montré que le couplage local dans les systèmes dissipatifs provoque une défaillance de propagation ou verouillage des fronts, dans une large région de paramètres connue sous le nom de *plage de verouillage* [80, 81]. À savoir, les fronts reliant différents états deviennent immobiles dans un grand espace de

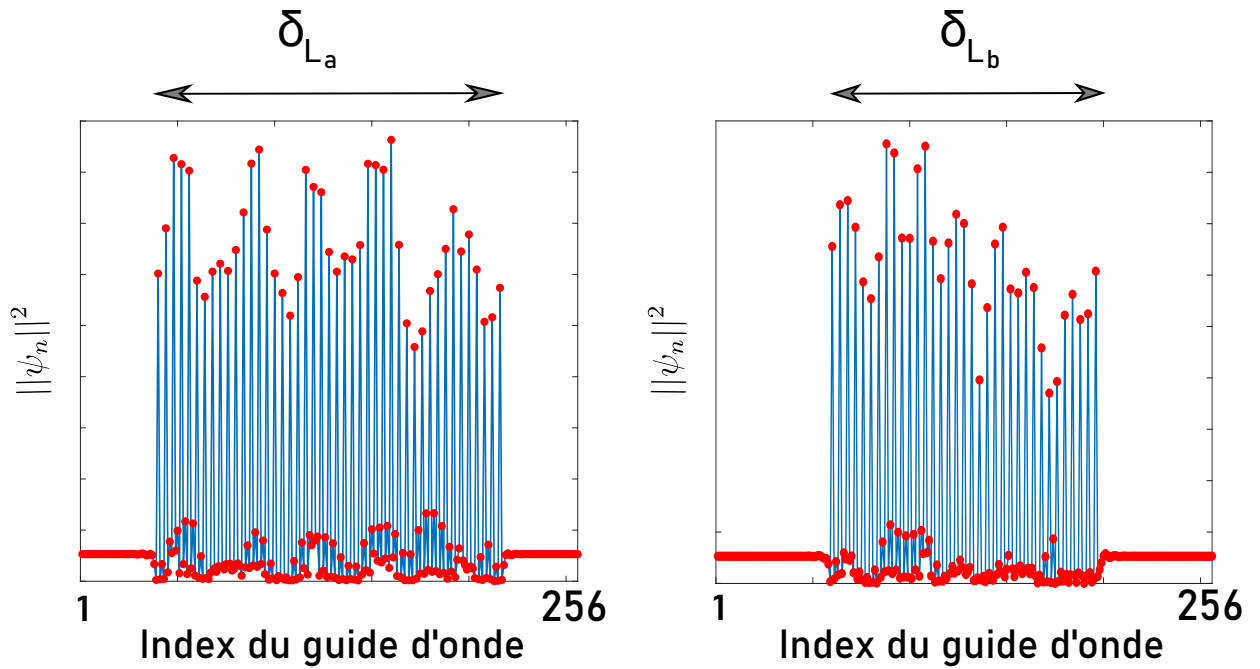


FIGURE 1.8 – Profil d'intensité de l'amplitude du champ électrique $||\psi_n(t)||^2$ à un instant fixe pour des tailles différentes d'OCS ($\delta_{L_a}, \delta_{L_b}$), obtenues à partir de simulations numériques du modèle Eq. 1.2 : en prenant $C = 1.9$, $\Delta = 7$ et $E_0 = 6$.

paramètres. Ce phénomène, résultat de la discrétisation, induit un potentiel effectif sur la dynamique des fronts, qui se caractérise par le recouvrement de l'interaction attractive des fronts et du potentiel de PEIERLS-NABARRO [81]. Ainsi, le couplage local empêche l'état incohérent d'envahir l'état cohérent. En effet, le potentiel de *Peierls-Nabarro* induit par la discrétisation, génère une famille de différentes tailles d'états spatio-temporels chaotiques localisés stables (Fig. 1.7). Par conséquent, ces états chimères stabilisés dans ce système de micro-résonateurs couplés, sont le résultat du couplage local combiné à la nature discrétisée du système.

Une autre caractéristique qu'on peut remarquer à travers les évolutions spatio-temporelles (Fig. 1.7), est la présence des motifs de TYPE SIERPINSKI. Ces patterns triangulaires restent une solution non linéaire universelle et typique de l'intermittence spatio-temporelle [82, 83]. La nature intermittente de la dynamique se caractérise par des alternances désordonnées entre un comportement régulier et irrégulier (Fig. 1.7). En augmentant la taille des états localisés (en jouant sur les conditions initiales injectées à l'équation de LUGIATO-LEFEVER (Eq. 1.2)), nous favorisons l'émergence de motifs conduisant à un comportement dynamique spatio-temporel plus complexe de TYPE SIERPINSKI.

Dans la suite de ce manuscrit, nous allons utiliser la taille de la partie incohérente δ_L de

nos états chimères comme un **paramètre déterminant pour quantifier la complexité spatio-temporelle** au sein de ce système optique.

1.4 Conclusion

Dans un premier temps, nous avons défini les notions de complexité, système complexe et non-linéarité. Nous avons expliqué qu'il s'agit des caractéristiques fondamentales d'un système complexe auto-organisé. Dans un second temps, nous avons introduit les états chimères, comme une des formes du phénomène d'auto-organisation. Par la suite, nous avons mis en évidence les travaux pionniers, théoriques et expérimentaux, établis afin de comprendre les mécanismes de formation de ce nouveau paradigme. Finalement, nous avons utilisé l'ensemble de ces notions afin de décrire nos états chimères optiques stabilisés dans un réseau de guides d'ondes couplés non linéairement. D'abord, nous avons décrit le système optique discret ainsi que l'équation de LUGIATO-LEFEVER (Eq. 1.2), qui régit sa dynamique spatio-temporelle. Nous avons expliqué que le couplage local ainsi que la nature discrétisée du système sont derrière la stabilisation des états chimères optiques dans ce dispositif optique. Nous avons remarqué la présence des patterns triangulaires de type SIERPINSKI dans les domaines incohérents de nos états chimères optiques. Nous allons voir, dans le chapitre suivant, l'importance de ces patterns spatio-temporels notamment dans le traitement de l'information.

Capacités de calcul dans les automates cellulaires

2.1 Introduction

La science des systèmes complexes est l'étude des grandes collections d'entités simples, où le comportement global est un résultat non trivial de la nature locale des interactions entre ses éléments individuels. Cette approche cherche à comprendre de manière fondamentale comment de tels comportements collectifs résultent de ces interactions entre des individus simples. Bien qu'aucun cadre commun n'ait été établi pour l'analyse des séries temporelles dans la science des systèmes complexes [84], de plus en plus la notion de "calculabilité" est importée pour expliquer la dynamique collective au sein de ces systèmes complexes [85]. La calculabilité dans les systèmes complexes est décrite via les concepts de mémoire, communication et traitement de l'information [86]. Par conséquent, on dit qu'un système possède ces propriétés de calculabilité s'il est capable d'accomplir trois opérations fondamentales en terme du stockage, du transfert et de la modification de l'information [87]. Le concept de calculabilité peut sembler assez distinct de ce domaine, mais l'application est bien raisonnée puisque les interactions entre les individus d'un système complexe, peuvent être vues comme une communication, et le système dans son ensemble comme un outil de traitement de l'information via une calculabilité distribuée.

Une panoplie d'études ont démontré que les conditions optimales du traitement d'information sont liées avec le comportement complexe situé dans la transition de phase entre l'ordre et le désordre. Le pionnier de cette hypothèse connue sous le nom '*edge of chaos*', a adopté la théorie d'information comme étant un outil efficace, pour mener ses investiga-

tions sur les automates cellulaires [7]. Ces derniers sont des systèmes dynamiques discrets conceptuels utiles pour modéliser les systèmes complexes dans la nature [88] et ont offert une plateforme enrichissante pour explorer le lien entre la calculabilité et le comportement complexe [7].

Dans ce chapitre, nous allons introduire l'hypothèse du 'edge of chaos' qui connecte la complexité avec la calculabilité, tout en détaillant les opérations basiques du traitement de l'information. Puis, nous définirons les automates cellulaires, qui présentent une voie royale pour comprendre d'une façon concrète les notions de calculabilité.

2.2 Calculabilité dans les systèmes complexes

2.2.1 Calculabilité et dynamiques de l'information

La nature du calcul distribué présente depuis longtemps un sujet d'intérêt pour la science des systèmes complexes, la physique, la vie artificielle, la bio- et la neuro-informatique. Dans tous ces domaines pertinents, le calcul distribué est généralement abordé en termes de mémoire, de communication et de traitement :

- **Mémoire** : fait référence au stockage d'informations par un agent ou un processus, pour les utiliser dans son avenir. Il a été étudié dans le mouvement coordonné des robots modulaires [89], dans la dynamique des temps de distribution inter-événements [90], et dans la synchronisation entre systèmes couplés [91].
- **Communication** : fait référence au transfert d'informations entre un agent ou un processus et un autre. Il s'est avéré pertinent pour les systèmes biologiques (par exemple, l'interaction dipôle-dipôle dans les microtubules [92] et dans la transduction du signal par les ions calcium [93]), les animaux sociaux (par exemple, le comportement collectif des poissons [94]) et l'influence des agents sur leurs environnements dans un système [95] et l'induction d'une structure neuronale émergente [96]).
- **Traitement** : Se réfère à la combinaison ou à la modification d'informations stockées et/ou transmises dans une nouvelle forme. Il a été discuté en particulier pour les réseaux de neurones biologiques et leurs modèles [97-100] (où il a été suggéré comme un moteur biologique potentiel), et également en ce qui concerne le calcul basé sur les collisions (par exemple [101, 102], et y compris la dynamique des solitons et les collisions [103]).

Le calcul distribué est défini comme étant, tout processus mené par plusieurs agents ou entités qui implique ces opérations sur des informations. Des exemples notables incluent :

l'évolution dans le temps de systèmes dynamiques discrets tels que les automates cellulaires [104], le traitement de l'information dans le cerveau [105], les réseaux de régulation des gènes calculant les comportements cellulaires [106], les troupeaux ou les bancs calculant leur cap collectif [94], les colonies de fourmis calculant les voies les plus efficaces vers les sources de nourriture [107], ou le comportement collectif dans les systèmes artificiels auto-organisés [84].

En effet, ces opérations existent dans tous les systèmes, que le système ait été ou non explicitement conçu pour calculer ou semble remplir une fonction utile : on parle alors de **calcul intrinsèque** [86, 87]. C'est pourquoi l'univers peut être considéré comme calculant son propre avenir, selon Lloyd [108] : «Qu'est-ce que l'univers calcule ? Il se calcule lui-même». Cette compréhension du calcul intrinsèque soutient également pourquoi "l'information est physique et la physique est l'information" [109]. De manière significative, ces trois opérations sur l'information sont formellement des fonctions primitives du calcul universel de Turing [7] :

- **stockage d'information.**
- **transfert (ou transmission) d'information.**
- **modification d'information.**

Ces opérations sont particulièrement importantes du point de vue théorique de la science des systèmes complexes, où elles font l'objet d'un certain nombre de conjectures importantes concernant la nature fondamentale du calcul distribué et sa relation avec le comportement complexe émergent.

L'information est le *langage* du calcul, donc toute tentative de caractériser la dynamique du calcul distribué, ou de décrire des systèmes complexes en terme de ces trois opérations fondamentales, doit nécessairement impliquer la théorie de l'information. Alors que cette théorie a traditionnellement été associée au traitement du signal, elle a été de plus en plus utilisée pour étudier des systèmes complexes [84, 110-112]. GELL-MANN [113] a dit : " *Bien que (les systèmes adaptatifs complexes) diffèrent largement dans leurs attributs physiques, ils se ressemblent dans la manière dont ils traitent l'information. Cette caractéristique commune est peut-être le meilleur point de départ pour explorer leur fonctionnement*". Cette déclaration fait allusion à trois caractéristiques importantes de la perspective de la théorie de l'information : sa généralité, sa concentration sur le calcul et les informations utiles connues qu'elle produit. Il est important de noter que l'utilisation de la théorie de l'information signifie que les mesures à base d'entropies, capturent les effets *non linéaires* et sont *génériques* et *portables* entre les domaines d'application.

Afin de quantifier la dynamique de l'information dans un système complexe, on fait

appel à deux grandes écoles de pensées. La première propose des mesures **moyennes** ou **globales** (e.g. [7, 114]), pour caractériser la dynamique du calcul d'un système complexe. Alors que la deuxième école [115] suggère que la quantification et la compréhension du calcul distribué nécessitera d'étudier ces dynamiques d'information et leurs interactions à l'échelle **locale** dans l'espace et dans le temps. Cela montrera comment le calcul se déroule dans le temps, et la dynamique de la façon dont des agents distincts interagissent pour accomplir une tâche collective. Nous reviendrons sur ces deux perspectives : moyenne et locale, afin de comprendre les travaux faits sur les automates cellulaires (section. 2.3.4).

2.2.2 Edge of chaos

La théorie de la calculabilité a été développée dans le cadre de la tentative de comprendre les aspects du traitement de l'information dans les systèmes. Une définition familière du traitement de l'information pourrait être "la transformation d'une entrée donnée en une sortie souhaitée". Cependant, afin d'appliquer la notion de traitement de l'information à des systèmes complexes et de la relier à la théorie des systèmes dynamiques, la notion doit être enrichie pour inclure la production d'informations ainsi que son stockage, sa transmission et sa manipulation logique.

L'adaptation de KOLMOGOROV et SINAI de la théorie d'information de Shannon (1950) [116, 117], témoigne les efforts investis afin de relier les capacités de traitement de l'information dans les systèmes non linéaires et leurs comportements temporels. Un résultat important est qu'un système chaotique déterministe peut être considéré comme un générateur d'information [118]. Une autre remarque précieuse est que la complexité de la prédiction d'un comportement de systèmes chaotiques croît de façon exponentielle avec le temps [119]. La relation entre la difficulté de la prédiction et le caractère aléatoire dynamique est simplement résumée par l'affirmation que le taux de croissance de la complexité descriptive est égal au taux de production d'information [120]. Ces résultats donnent une vision du chaos déterministe qui met l'accent sur l'imprévisibilité qui résulte de la dynamique produite. Ce sont probablement les premiers liens entre la dynamique et la calculabilité.

Les systèmes dynamiques naturels qui évoluent et apprennent de l'interaction avec leur environnement immédiat présentent à la fois un ordre structurel et un chaos dynamique [121]. L'ordre est le fondement de la communication entre les éléments à tout niveau d'organisation, qu'il s'agisse d'une population de neurones, d'abeilles ou d'humains. Pour un organisme, l'ordre est le processus qui nécessite de la *mémoire* proportionnellement à la longueur du motif qu'il répète. La diversité comportementale, est fondamentale pour

la survie d'un organisme. Aucun organisme ne peut modéliser l'environnement dans son intégralité. Le Chaos, tel que nous le comprenons maintenant, est le mécanisme *dynamique* par lequel la nature développe un hasard contraint et utile.

Dans cette double vision mémoire/dynamique, ces deux types de processus sont simples à décrire statistiquement. Entre ces extrêmes se trouvent cependant les processus complexes hautement structurés qui utilisent à la fois le caractère aléatoire dynamique et le stockage d'information pour produire leur comportement. Ces processus sont plus complexes à décrire statistiquement que les processus chaotiques ou périodiques simples.

En un sens, le titre *complexe* rend compte de la nature difficile de l'analyse de ces systèmes. Dans un autre sens, il saisit également l'essence de leur nature intéressante d'un point de vue de la théorie de la calculabilité. En effet, l'interaction des éléments au niveau microscopique peut donner lieu à une organisation sophistiquée du système au niveau macroscopique. En d'autres termes, ils peuvent présenter un comportement global qui est intéressant mais qui n'est pas une conséquence évidente des interactions locales. On parle alors de comportement *émergent* [85]. Certains auteurs vont plus loin dans leurs définitions des systèmes complexes. Par exemple, MITCHELL [85] propose qu'un système complexe soit : *"un système dans lequel un grand nombre de composants sans contrôle central et des règles de fonctionnement simples donnent lieu à un comportement collectif complexe, à un traitement sophistiqué de l'information et à une adaptation via l'apprentissage ou l'évolution"*. Il ne fait aucun doute que ces termes sont une partie essentielle de notre compréhension des systèmes complexes et leur calculabilité.

Il est important de noter que les systèmes complexes sont généralement décrits comme combinant des éléments d'ordre et d'aléatoire pour créer un comportement vraiment complexe [122]. Par exemple, les économies impliquent une régulation et une perception d'un comportement rationnel en même temps que de grandes variations dans les comportements individuels et des fluctuations imprévues dans la dynamique du marché. En effet, un comportement vraiment complexe n'est ni complètement ordonné, ni complètement aléatoire. Les systèmes ordonnés sont parfaitement structurés et donc simples à prédire [122]. Les systèmes complètement désordonnés ne peuvent pas du tout être prédits au niveau individuel, mais la prédiction du comportement moyen est non seulement possible mais triviale. Les systèmes complexes, quant à eux, incarnent une dualité entre dépendance et indépendance de leurs composants, rendant leur prédiction possible mais non triviale. De telles observations sont généralisées par certains auteurs en suggérant que le comportement complexe se produit lors d'une transition de phase entre le comportement ordonné et le comportement chaotique (généralisé dans l'hypothèse **"edge of chaos"** [7, 123]). Il existe certainement de

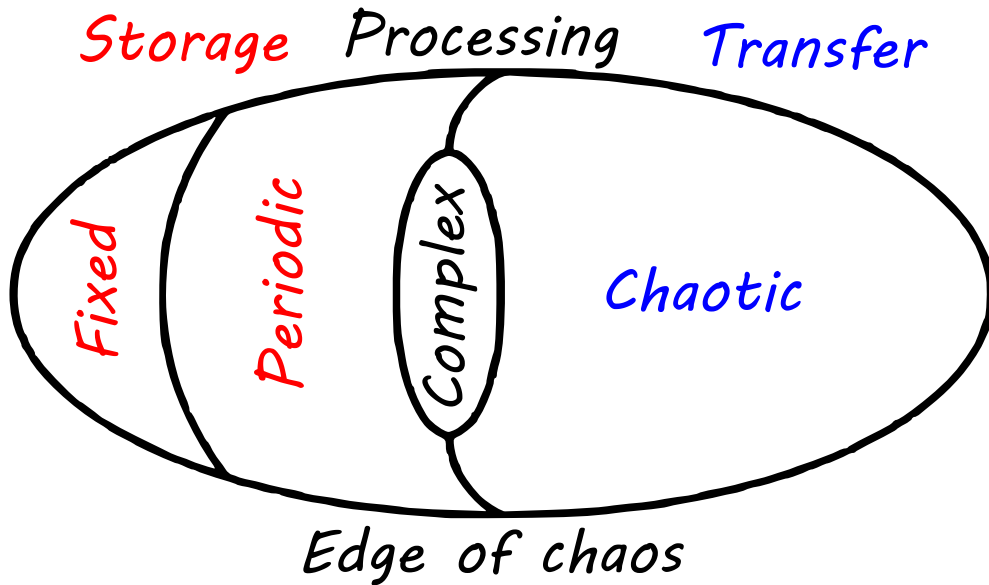


FIGURE 2.1 – Le célèbre diagramme d'œuf de LANGTON, connectant le comportement complexe au traitement optimal de l'information [7].

nombreuses transitions de phase ordre-chaos (en particulier lorsqu'elles sont dictées par un seul paramètre ordre-chaos) où la complexité a été mesurée d'une manière ou d'une autre pour être maximisée dans la transition.

Une application importante de cette hypothèse d'**edge of chaos** (EOC) est le comportement adaptatif : il a été suggéré par plusieurs auteurs, biologistes et physiciens, qu'un comportement au EOC pourrait être optimal du point de vue de l'adaptation. Ce point de vue peut également être lié aux processus évolutifs : par exemple, KAUFFMAN [123] a soutenu que *les "systèmes de traitement parallèle" se trouvant au niveau de l'EOC sont les plus capables non seulement de s'adapter, mais aussi d'évoluer*. Ici, il y a deux manières possibles d'explorer cet énoncé qui ne sont pas du tout incompatibles mais plutôt complémentaires :

Le stockage, le transfert et le traitement des informations sont optimisés au EOC (Fig. 2.1). Une critique peut être formulée : il ne suffit pas de stocker et de transmettre des informations, elles doivent également être transformées et traitées d'une manière ou d'une autre, mais cela signifie simplement qu'être au EOC est une condition nécessaire mais pas suffisante. Cependant, ce point est intéressant et a été développé théoriquement par CRUTCHFIELD et YOUNG [124] et par LANGTON [7]. En particulier, ce dernier a montré que l'information mutuelle (une condition de base pour que le calcul coopératif ait lieu est la présence d'une grande quantité d'information mutuelle entre les éléments d'un système) est maximisée dans les automates cellulaires élémentaires à la frontière entre

ordre/périodique et aléatoire/comportement chaotique. D'autres auteurs (PACKARD [8], KANEKO [125], SOLÉ [126] 1995) ont proposé des outils pour la modélisation des systèmes biologiques et mis en évidence un traitement optimal de l'information au EOC. Bien qu'il y ait eu des résultats contradictoires [127], plusieurs études ont étudié la propagation et le traitement de l'information dans les réseaux, en particulier les transitions de phase de ces propriétés entre les régimes ordonnés et chaotiques. SOLÉ et Valverde [128] ont étudié l'effet de la variation du taux de génération de messages dans un modèle de réseaux informatiques, trouvant des transitions de phase maximisant le nombre de paquets réellement livrés et les informations mutuelles dans l'état des nœuds couplés. Ils en déduisent que le transfert d'information est maximisé à l'état critique. KINOUCI et COPELLI [97] ont étudié la variation du "rapport de ramification" (en fait un niveau d'activité) dans un réseau d'éléments excitables, trouvant des transitions de phase maximisant la plage dynamique de la sortie de l'élément et en déduisant une maximisation du traitement de l'information à la criticité.

Une autre vision menée par HAKEN [129] et GOODWIN [130], qui ont défini l'EOC comme un point de rupture de symétrie, où l'on trouve un compromis optimal entre exploration et exploitation. Évidemment, si un système doit être capable de faire face aux changements environnementaux, c'est-à-dire de briser les symétries de la bonne manière lorsque cela est nécessaire, il doit se maintenir au voisinage d'un point de rupture de symétrie (le EOC).

Il est intéressant de noter que ces deux points de vue ne sont pas opposés : En effet, la capacité d'un essaim ou de tout système collectif à déplacer avec souplesse ou pertinence son activité en réponse à une perturbation est optimisée au EOC en raison de propriétés dynamiques qui permettent la collecte et le 'traitement' de l'information au niveau collectif (HOHENBERG [131]).

2.3 Automates cellulaires

2.3.1 Bref historique

Les automates cellulaires sont des systèmes spatialement étendus constitués d'un grand nombre de composants identiques simples avec une connectivité locale. De tels systèmes ont le potentiel d'effectuer des calculs complexes avec un degré élevé d'efficacité et de robustesse ainsi que de modéliser le comportement de systèmes complexes dans la nature. Pour ces raisons, les automates cellulaires et les architectures associées ont été largement étudiées dans les mathématiques des sciences naturelles et en informatique.

Le concept original d'automates cellulaires est plus fortement associé au scientifique et mathématicien JOHN VON NEUMANN. Selon l'histoire racontée par BURKS [132], VON NEUMANN était profondément intéressé par les liens entre la biologie et la nouvelle science des dispositifs informatiques, la théorie des automates. Le phénomène biologique d'auto-reproduction prédominait dans son esprit et il avait posé une question fondamentale : 'Quel type d'organisation logique est suffisant pour qu'un automate puisse se reproduire. L'idée d'utiliser des automates cellulaires comme cadre pour répondre à cette question a été suggérée à VON NEUMANN par STANISLAW ULAM [133]. Ainsi, le concept original d'automates cellulaires peut être attribué à Ulam, tandis que le développement précoce du concept peut être attribué à VON NEUMANN. Il croyait fermement qu'une théorie générale du calcul dans des réseaux complexes d'automates tels que les automates cellulaires serait essentielle à la fois pour comprendre les systèmes complexes dans la nature et pour concevoir des systèmes complexes artificiels. La solution détaillée de Von Neumann à sa question : "Quel type d'organisation logique est suffisante pour qu'un automate puisse se reproduire ?" a été présentée dans son livre *Theory of Self-Reproduction Automata* [132]). Le manuscrit était incomplet au moment de la mort de VON NEUMANN en 1957. Le manuscrit a été édité et complété par Burks, qui donne également une excellente vue d'ensemble détaillée du système de VON NEUMANN dans l'essai de son livre " *Essays on Cellular Automata*" [133].

Sur la base du concept théorique d'universalité exploré par WOLFRAM dans son livre "A New Kind of Science" [104], les chercheurs ont tenté de développer des architectures d'automates cellulaires plus simples et plus pratiques qui peuvent être utilisées pour modéliser des domaines d'application très divergents. À cet égard, deux développements notables peuvent être attribués à CONWAY et WOLFRAM. Dans les années 1970, le mathématicien JOHN CONWAY a proposé son désormais célèbre jeu de la vie [134] qui a suscité un large intérêt parmi les chercheurs. Au début des années 80, STEPHEN WOLFRAM a étudié en détail une famille de règles simples d'automates cellulaires unidimensionnels (maintenant célèbres règles de WOLFRAM [9]) et a montré que même ces règles les plus simples sont capables d'émuler un comportement complexe.

Ensuite, des chercheurs de divers domaines, sans nécessairement être conscients du cadre mentionné ci-dessus, ont intuitivement identifié la dynamique des automates cellulaires avec des problèmes dans leurs propres domaines de cellules et population d'organismes [135, 136]. L'application réussie des automates cellulaires dans la modélisation du système immunitaire a été explorée par CELADA, SEIDEN et DEBOER [137, 138], et dans les modèles de développement tumoral [139], et dans la détection des troubles génétiques des cellules cancéreuses [140, 141]. En écologie, ils ont été utilisés pour modéliser l'écosystème prédateur-

proie [142], pour détecter la nature de la migration des poissons dans rivières [143], et pour modéliser la croissance de la population végétale [144, 145]. La chimiotaxie modélisée par RESNIK [142] a été utilisée pour déterminer la marche aléatoire des animaux en réponse à un gradient chimique. L'effet de la marche aléatoire simple par un seul individu et les marches aléatoires multiples par un certain nombre d'individus dans un système ont été modélisés et largement étudiés [146]. Un certain nombre d'applications d'automates cellulaires ont été rapportées dans les domaines des séquences d'ADN [147-149]. Les automates cellulaires ont été utilisés pour modéliser la cinétique des systèmes moléculaires et la croissance cristalline en chimie [150] et le phénomène d'absorption-désorption important pour analyser l'empoisonnement d'une surface lors d'une catalyse hétérogène [151], l'inter-diffusion d'atomes de deux matériaux [152], le système de diffusion entraîné où le champ externe biaise le mouvement de chaque espèce dans la direction opposée [153], le processus de solidification avec une analyse spéciale de la transformation de phase de la substance de liquide à solide [154], la formation d'alliages [150]. En physique, les applications couvrent l'étude des systèmes dynamiques à partir de l'interaction des particules jusqu'à l'agrégation des galaxies [155] des systèmes de spin [156, 157], des modèles pour divers types de formes régulières, croissance dendritique et aléatoire basée sur des automates cellulaires bidimensionnels [150], la formation de motifs dans les systèmes de réaction-diffusion [158-160], modélisation des systèmes hydrodynamiques [161]. JAMES M.SAKODA a été le premier à développer un modèle automate cellulaire basé sur les sciences sociales. SAKODA a publié l'article "The Checker board Model of Social Interaction" en 1971 [162]. OOMES a utilisé l'automate cellulaire pour modéliser l'effet des inégalités économiques sur les marchés émergents [163]. Les automates cellulaires ont également été utilisés pour modéliser le flux de trafic [153] ainsi qu'un outil de conception pour le développement urbain [164], et pour développer un modèle pour les électeurs [165]. Une bonne analyse théorique de la modélisation des automates cellulaires dans une perspective sociale est rapportée dans [166].

Au vu d'une telle diversité, nous avons présenté les principales applications qui ont non seulement porté la recherche sur les automates cellulaires vers de nouveaux sommets, mais ont également permis à des chercheurs de différents domaines de se joindre et d'exploiter collectivement le monde passionnant des automates cellulaires.

2.3.2 Définition

Les automates cellulaires font partie des modèles standards des systèmes dynamiques spatialement étendus, qui se démarquent des autres modèles par leur discrétion spatio-

temporelle. Un automate cellulaire comprend un réseau de N cellules ou sites identiques. L'état de chaque cellule (i) au pas de temps (t) est noté \mathbf{X}_i^t .

Pour bien définir un automate cellulaire, il faut préciser : la dimension (d) de l'arrangement des cellules, le nombre d'états possibles par cellule (k) à partir de l'alphabet S , le nombre de voisins reliés à la cellule centrale (i) dans chaque direction, appelé rayon (r) et une règle de transition locale (φ) jouant le rôle d'une équation du mouvement. Selon cette dernière (φ), chaque cellule (i) est mise à jour vers l'état suivant \mathbf{X}_i^{t+1} en fonction de son état actuel \mathbf{X}_i^t et ceux de ses voisins pris en compte (r).

Dans ce manuscrit, nous traiterons des automates cellulaires élémentaires (ECA), proposés par WOLFRAM [9, 88] qui sont définis comme un réseau unidimensionnel ($d = 1$) de cellules, soumis à des conditions aux limites périodiques ($\mathbf{X}_i^t = \mathbf{X}_{i+N}^t$), avec deux états binaires possibles par site ($k = 2$, $S = \{0, 1\}$), et où la taille (m) du voisinage spatial d'une cellule (i) est liée au rayon ($r = 1$, une cellule de chaque côté) avec la relation $m = 2r + 1 = 3$ (Fig. 2.2).

Dans ce cas des automates cellulaires d'état binaires, la règle de mise à jour locale ($\varphi(m)$) est connue sous le nom de table de règles, car elle énumère les 8 ($k^m = 2^3$) configurations possibles ($\mathbf{X}_{i-r}^t \ \mathbf{X}_i^t \ \mathbf{X}_{i+r}^t$) et spécifie l'état de la cellule (i) à l'itération suivante X_i^{t+1} , comme :

$$\mathbf{X}_i^{t+1} = \varphi(\mathbf{X}_{i-1}^t, \mathbf{X}_i^t, \mathbf{X}_{i+1}^t) \quad (2.1)$$

A chaque pas de temps (t), toutes les cellules du tableau évoluent de manière synchrone en utilisant une règle de mise à jour locale homogène $\varphi(m)$.

Au total, il y a ($k^{k^m} = 2^8$) 256 règles de mise à jour possibles φ . WOLFRAM a décrit un schéma de numérotation pour étiqueter chaque règle φ parmi les 256 ECA, par un nombre entier compris entre 0 et 255 [104]. Il convient de souligner que l'espace de règles ECA peut être réduit à 88 ensembles de règles équivalents en raison de la réflexion miroir et de la 0-1 symétrie [104] qui peuvent être affichés dans les tables de règles (Fig. 2.2). Par exemple, la règle 255 est équivalente à la règle 0 (Fig. 2.2).

Contrairement aux automates cellulaires avec une dimension supérieure (d) ou des états (k), le comportement des ECA peut être visualisé par un diagramme espace-temps (Fig. 2.3) avec des cellules noires et blanches correspondant aux états 0 et 1.

2.3.3 Classification

L'un des problèmes posés par les automates cellulaires est leur classification. En effet, il a été démontré qu'à partir d'une condition initiale arbitraire, on ne peut pas décider à

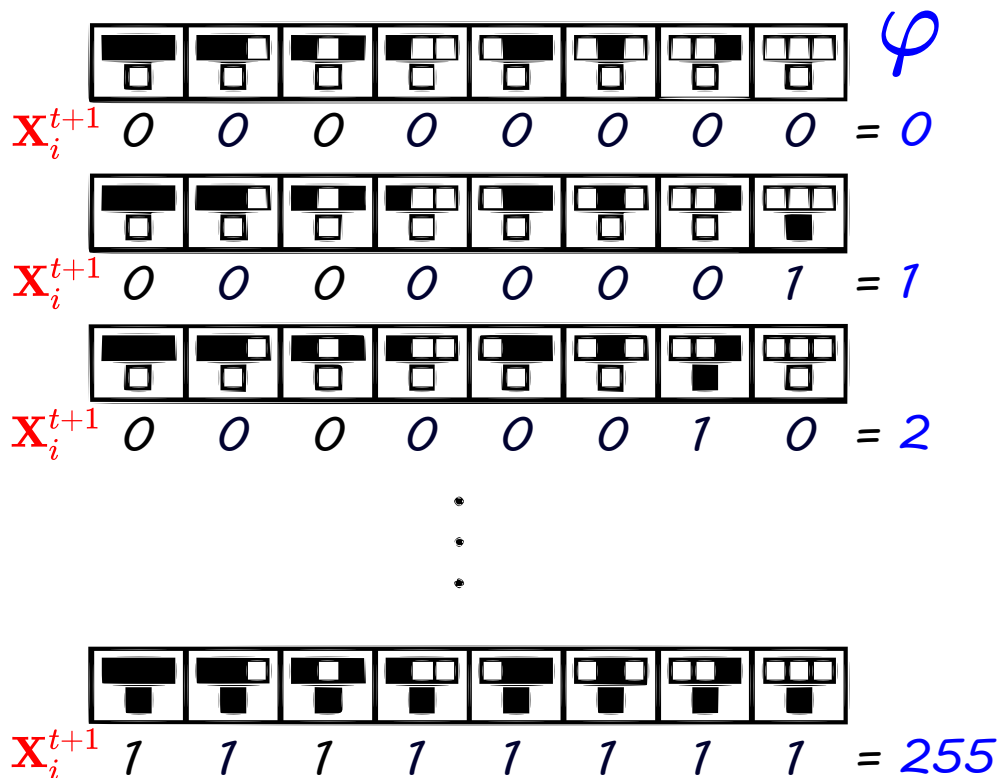


FIGURE 2.2 – Tables des règles des automates cellulaires élémentaires (0,1,2 et 255), énumérant l'évolution des 8 situations possibles.

quelle classe un automate cellulaire donné, appartient [167]. Généralement, il existe deux grandes écoles de classification. Une classification *phénotypique* basée sur la quantification de la dynamique spatio-temporelle observée [7, 168, 169] et une classification *génotypique* fondée sur la paramétrisation des tables de règles [7, 170].

La tentative pionnière a été proposée par Wolfram pour classer les ECA [88], où il a basé sa classification heuristique sur l'étude du comportement spatio-temporel "moyen" observé sur un échantillon de configurations initiales aléatoires (Fig. 2.3). En effet, considérant les automates cellulaires comme des systèmes dynamiques discrets spatialement étendus, il a proposé une classification qualitative du comportement des automates cellulaires à peu près analogue aux classifications de la théorie des systèmes dynamiques [88]. L'analogie est grossière puisque des concepts tels que 'chaos', 'attracteur', 'bifurcation', 'dépendance sensible aux conditions initiales', sont rigoureusement définis pour les systèmes dynamiques à temps continu à état continu alors que les automates cellulaires sont des systèmes à temps discret à état discret. Ainsi, WOLFRAM a regroupé les ECA en quatre classes (Fig. 2.3) :

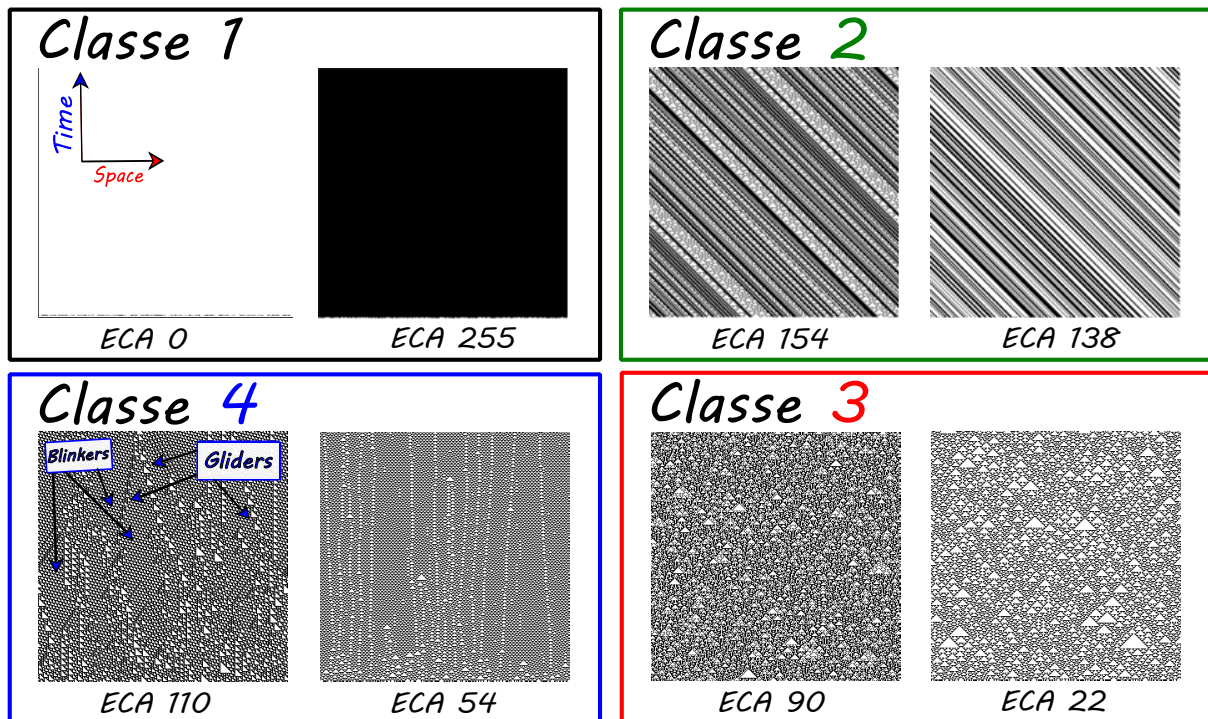


FIGURE 2.3 – Des diagrammes espace-temps illustrant les comportements typiques des quatre classes de Wolfram d’automates cellulaires. Chaque classe est présentée par deux exemples d’ECA qui y appartiennent. Ces évolutions sont le résultat d’une génération à partir de conditions initiales aléatoires sur un réseau de 300 cellules.

Classe I

Presque toutes les configurations initiales évoluent vers un état homogène.

Classe II

Presque toutes les configurations initiales évoluent après une période transitoire vers un point fixe ou un cycle temporellement périodique.

Classe III

Presque toutes les configurations initiales se évoluent après une période transitoire vers un comportement chaotique. Le terme chaotique fait ici référence à un comportement spatio-temporel apparemment imprévisible.

Class IV

Presque toutes les configurations initiales évoluent vers des patterns complexes, des structures localisées corrélées parfois de longue durée.

Ici, les classes I et II représentent un comportement ordonné, et la classe III représente un comportement chaotique. La classe IV représente un comportement complexe et est

considérée comme se situant entre les classes ordonnées et chaotiques.

WOLFRAM [88] a spéculé l'hypothèse que la classe IV complexe, est capable du calcul universel en se basant sur l'utilité des structures localisées telles que les '*blinkers*' et propageantes telles que les '*gliders*' (Fig. 2.3).

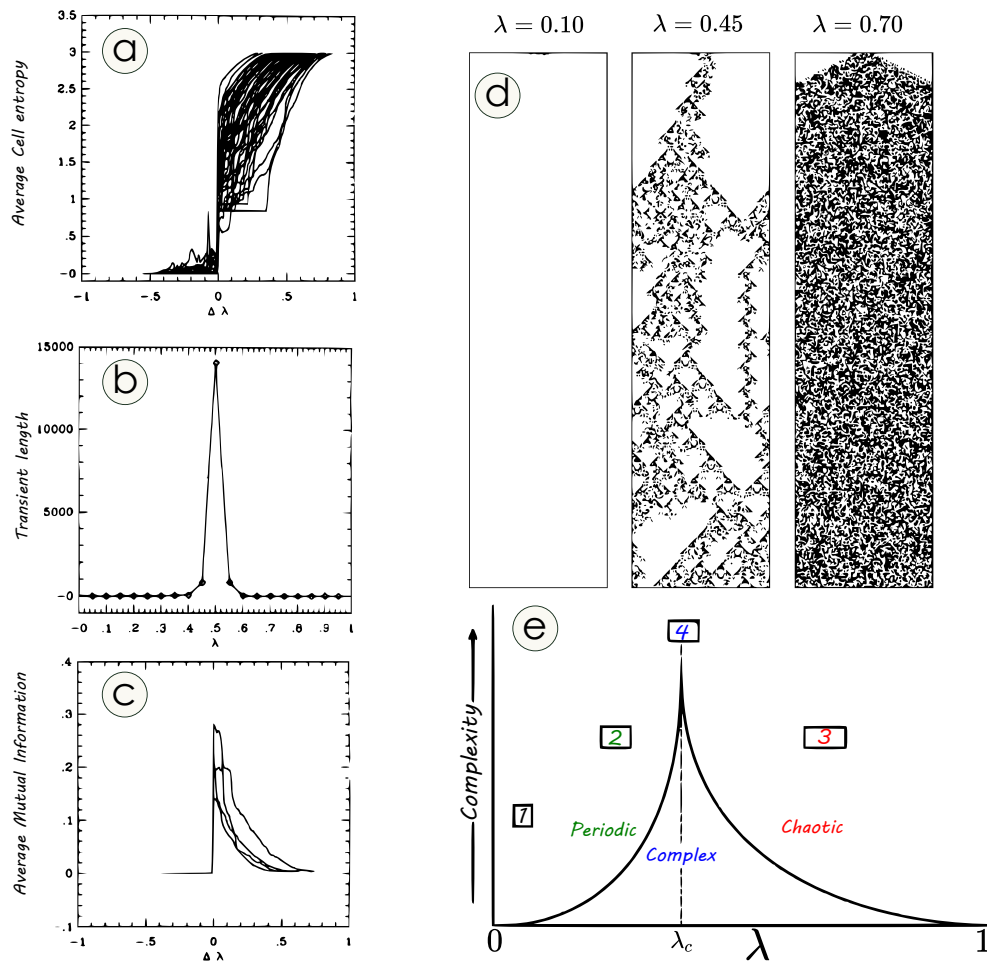


FIGURE 2.4 – Paramétrisation de l'espace des règles proposée par Langton [7]. (a) : Superposition de 50 événements de transition superposés de la variation moyenne de l'entropie d'une seule cellule lorsque nous varions λ ($\Delta\lambda = \lambda - \lambda_c$). (b) : Durée transitoire moyenne en fonction de λ dans un automate de 128 cellules. (c) : Information mutuelle moyenne en fonction de λ . (d) : Évolutions spatiotemporelles d'automates cellulaires unidimensionnelles particulières en changeant λ , générées à partir de configurations initiales partiellement aléatoires. (e) : Emplacement des classes Wolfram dans l'espace de λ .

À la suite des travaux de WOLFRAM, plusieurs chercheurs se sont questionnés comment les propriétés statiques des règles des automates cellulaires étaient liées au comportement dynamique des automates cellulaires. CHRISTOPHER LANGTON a par exemple étudié la relation entre le comportement dynamique moyen des automates cellulaires et une statis-

tique particulière (λ) d'une table de règles d'automates cellulaires [7]. Pour les automates cellulaires élémentaires (à état binaire), λ est simplement la fraction des '1' dans les bits de sortie de la table de règles.

LANGTON a réalisé un certain nombre d'échantillons Monte Carlo d'automates cellulaires bidimensionnels commençant par $\lambda = 0$ et augmentant progressivement jusqu'à $\mathbf{1}$ (les tables de règles les plus homogènes aux plus hétérogènes). Langton a utilisé diverses statistiques telles que l'entropie d'un seul site, les informations mutuelles sur deux sites et la durée transitoire moyenne pour classer le comportement moyen des automates cellulaires à chaque valeur λ (Fig. 2.4). La notion de '*comportement moyen*' visait à capturer le comportement le plus probable observé avec une configuration initiale choisie au hasard pour des automates cellulaires sélectionnés au hasard dans un sous-espace λ fixe. Ces études ont révélé une certaine corrélation entre les différentes statistiques et λ . La corrélation est assez bonne pour les valeurs λ très basses et très élevées (Fig. 2.4). Cependant, pour les valeurs λ intermédiaires dans les automates cellulaires à états finis, il existe un degré élevé de variation de comportement (Fig. 2.4).

LANGTON a affirmé sur la base de ces statistiques que lorsque λ est incrémenté, le comportement moyen des automates cellulaires subit une transition de phase d'un comportement ordonné (point fixe ou cycle limite après une courte période transitoire) à un comportement chaotique (apparemment imprévisible après une courte période transitoire). Lorsque λ atteint une 'valeur critique' λ_c , on prétend que les règles ont tendance à avoir des phases transitoires de plus en plus longues (Fig. 2.4). De plus, LANGTON a affirmé que les automates cellulaires proches de λ_c ont tendance à présenter des patterns complexes non périodiques mais non aléatoires à longue durée de vie. Il a aussi proposé que le régime λ_c corresponde approximativement à la classe complexe IV de WOLFRAM.

En outre, LANGTON a conjecturé l'hypothèse que les automates cellulaires capables d'effectuer des calculs complexes se trouveront très probablement dans ce régime 'edge of chaos'. *Car le calcul complexe dans les automates cellulaires nécessite des transitoires et des longueurs de corrélation espace-temps suffisamment longs.* Une revue de ce travail qui est sceptique quant aux relations entre λ et les propriétés dynamiques et computationnelles des automates cellulaires est donnée dans (MITCHELL [127]) et (PACKARD [8]).

2.3.4 Calculabilité émergente dans les automates cellulaires

De nombreux systèmes spatialement étendus, à la fois physiques et biologiques, ont tendance à générer des patterns complexes [171-174]. Les patterns de ces systèmes découlent

des interactions locales non-linéaires des composants individuels. En d'autres termes, le comportement du système dans son ensemble est plus qu'une simple superposition linéaire des comportements des composants individuels lorsqu'ils sont considérés isolément. Cette propriété d'*émergence* est capturée par l'expression populaire "le tout est plus que la somme des parties". De cette manière, la dynamique globale incarnée par la formation de motifs dans les systèmes spatialement étendus peut être considérée comme une propriété *émergente* de ces systèmes [175, 176].

Souvent, les patterns émergents dans les systèmes spatialement étendus donnent lieu à une certaine forme de comportement globalement coordonné, ou de traitement global de l'information. Ces patterns sont utilisés pour stocker, transmettre et traiter des informations afin de prendre des décisions ou des classifications à un niveau global, qui peuvent être naturellement interprétées comme des calculs. Ainsi, de nombreux systèmes spatialement étendus peuvent être considérés pour effectuer des calculs médiés par la formation de patterns émergents.

On prétend que les systèmes qui présentent un calcul émergent ont plusieurs avantages potentiels par rapport aux systèmes informatiques traditionnels en termes d'efficacité, de flexibilité et de robustesse [177, 178].

Il y a peu de compréhension théorique de la façon dont la dynamique (c'est-à-dire le comportement dans le temps) d'un système spatialement étendu donne lieu à un calcul émergent. En d'autres termes, on ne sait pas comment les patterns émergents dans la dynamique d'un système encodent, transmettent et traitent les informations nécessaires pour effectuer des calculs, ou du moins comment nous pouvons interpréter ces patterns de cette manière. Ainsi, une question importante est : *quelle est la relation entre la dynamique (en particulier, la formation de patterns émergents) et la capacité de calculabilité dans les systèmes spatialement étendus ?*

Pour permettre une étude formelle de la relation entre la dynamique, le calcul émergent et l'évolution dans les systèmes spatialement étendus, des modèles mathématiques et informatiques sont utilisés. En particulier, les automates cellulaires sont utilisés comme une classe des modèles de systèmes spatialement étendus.

Puisque les automates cellulaires ont la capacité d'une large gamme de dynamiques ainsi que de nombreux types de calcul, ils forment une classe idéale de modèles pour étudier la relation entre la dynamique et la capacité de calcul dans les systèmes étendus spatialement. La classification de WOLFRAM du comportement des automates cellulaires, ainsi que sa spéculation sur les automates cellulaires capables de calcul, ont fourni un premier pas dans cette direction. Cependant, cette étude ne traitait que le comportement des automates

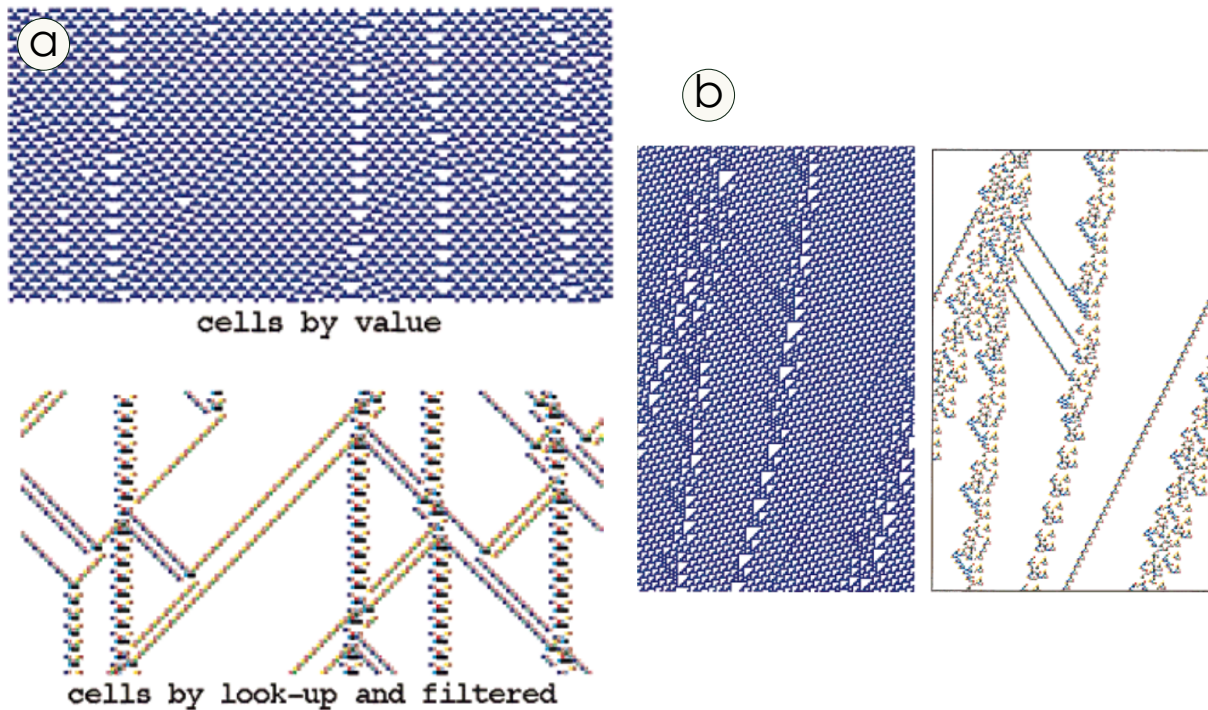


FIGURE 2.5 – Deux exemples d'évolutions d'automates cellulaires complexes, contenant des patterns spatio-temporels et supportant des *gliders* en interaction, filtrés par la méthode de Wuensche [169]. (a) : la règle 54. (b) : la règle 110.

cellulaires "générique" ou "moyen" et, de plus, n'utilisait pas de schéma de classification bien défini ni de mesure de la capacité de calculabilité.

LANGTON [7] a établi l'hypothèse que le comportement complexe des automates cellulaires au "edge of chaos" présentait les trois opérations de composants nécessaires au calcul universel. Il a suggéré que plus un système devient chaotique, plus la *transmission* d'informations augmente, et plus un système devient ordonné, plus il *stocke* d'informations (Fig. 2.1). Comme dit précédemment, un comportement complexe se produisait à une transition de phase entre ces extrêmes nécessitant un niveau intermédiaire à la fois de stockage et de transmission de l'information. Il a suggéré par exemple que les niveaux intermédiaires de transmission d'informations peuvent être supportés par *grandes longueurs de corrélation*. Il poursuit en suggérant que les *particules* forment la base de la transmission de l'information, car elles semblent faciliter la communication de la dynamique d'une zone de l'automate cellulaire à une autre zone. Pour compléter l'identification qualitative des éléments de calcul dans les automates cellulaires, il a également suggéré que les *blinkers* constituaient la base du stockage de l'information, et que les collisions entre ces structures propagatives (particules) et statiques (blinkers) *peuvent modifier les informations stockées*

ou transmises dans le cadre d'un calcul global, ou processus de décision sur la dynamique (Fig. 2.3). Il a également fait des tentatives rudimentaires pour quantifier le transfert moyen d'informations (et dans une certaine mesure le stockage d'informations), via l'information mutuelle (Fig. 2.4).

Reconnaissant l'importance des structures émergentes pour le calcul, il existe plusieurs exemples de tentatives d'identification automatique des règles d'automates cellulaires qui donnent lieu à des particules et *gliders* (voir [169] par exemple), suggérant qu'il s'agit des règles d'automates cellulaires les plus intéressantes et les plus complexes (Fig.2.5).

Plusieurs auteurs critiquent cependant les approches susmentionnées consistant à tenter de classer les automates cellulaires en fonction de leur comportement générique ou de "propriétés statistiques globales", suggérant que le large éventail de dynamiques différentes qui se déroulent dans les automates cellulaires rend cela problématique [127]. Malgré l'hypothèse de Langton et l'introduction de mesures utiles de la complexité des règles des automates cellulaires (par exemple [87, 114, 169], il n'y a pas de transition de phase établie dictée par un seul paramètre d'ordre-chaos dans les automates cellulaires [87]. En outre, GRAY suggère qu'il peut effectivement y avoir des classes de automates cellulaires capables de calculs plus complexes que le calcul universel seul [179]. Plus important encore, HANSON et CRUTCHFIELD [180] critiquent l'accent mis sur la capacité de calcul universelle comme s'éloignant de la capacité d'identifier les "propriétés de calcul génériques", c'est-à-dire qu'un manque de capacité de calcul universel ne signifie pas qu'un automate cellulaire n'entreprend aucun calcul du tout.

Alternativement, ces études suggèrent que l'analyse de la riche dynamique spatio-temporelle au sein des automates cellulaires est un objectif plus approprié. Ainsi, les références [127, 180] et d'autres ont analysé la dynamique locale du calcul intrinsèque ou autre calcul spécifique, en se concentrant sur les rôles computationnels des structures émergentes. Ils s'alignent sur les observations de Langton sur les *blinkers stockant l'information, les particules facilitant le transfert d'information et les collisions facilitant la modification ou le traitement de l'information*, et l'analyse utilisant de tels filtres pour analyser les règles d'automates cellulaires sélectionnées via un calcul évolutif pour effectuer des tâches de classification par Mitchell et al. [182]. Sont également pertinentes les études qui explorent en profondeur la nature des particules et leurs interactions (Fig. 2.6), par exemple : les types de particules et leurs produits d'interaction identifiés pour des automates cellulaires particuliers dans [182-185], les règles établies pour leurs produits d'interaction dans [186], et les études de "calcul basé sur les collisions" dans [102, 187] .

La mécanique computationnelle est un cadre général d'analyse des systèmes dynamiques

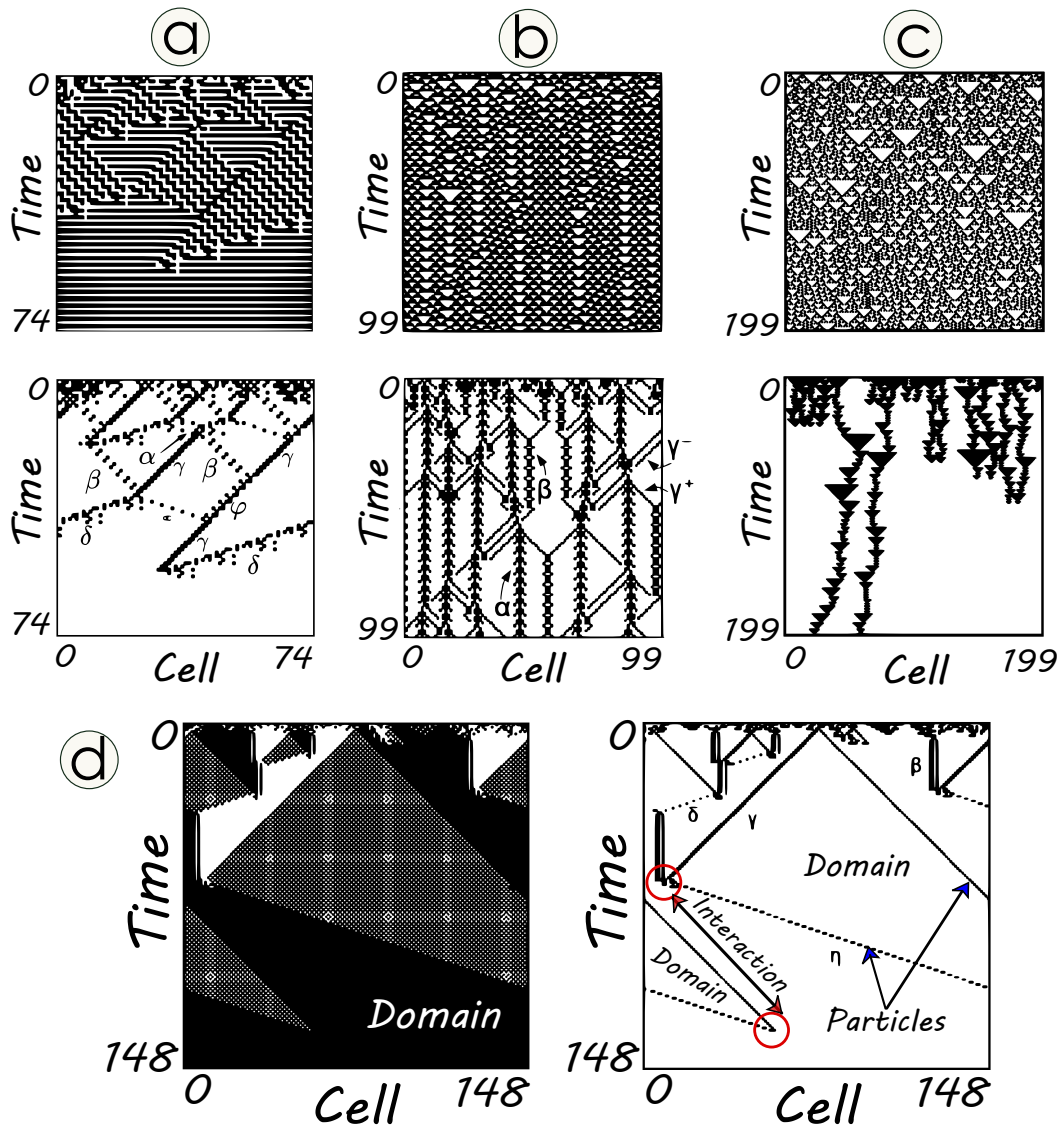


FIGURE 2.6 – La mécanique computationnelle [181]. (a) : Diagramme espace-temps de ϕ_{sync5} et sa version filtrée. (b) : Diagramme espace-temps de ECA 54 et la version filtrée. (c) : Diagramme espace-temps de ECA 18 et la version filtrée. (d) : schéma explicatif des éléments essentiels permettant le traitement de l'information, dans le cadre de la mécanique computationnelle.

qui permet de déduire un modèle du processus caché qui a généré le comportement observé d'un système. Ce modèle capture les patterns ou régularités observés dans ce comportement. Dans le cadre mécanique computationnelle, les patterns dominants dans la dynamique d'un automate cellulaire sont identifiés puis classés en *domaines réguliers*, *particules* et *interactions de particules* (Fig. 2.6).

Cette perspective des rôles computationnels des structures émergentes est importante

non seulement pour notre compréhension théorique de la nature du calcul distribué, mais est également importante sur le plan pratique .

Malgré cet intérêt, il serait plus intéressant de simplement quantifier le rôle des structures émergentes dans le calcul des automates cellulaires. Apporter une telle preuve nécessite de pouvoir quantifier la dynamique computationnelle du calcul au regard des trois opérations fondamentales, à une *échelle locale* dans l'espace et dans le temps au sein de des automates cellulaires. Seule l'échelle locale montrera explicitement les rôles computationnels des structures émergentes.

LIZIER a présenté un cadre pour la dynamique de l'information du calcul distribué, qui est nouveau dans la *quantification séparée* du stockage, du transfert et de la modification de l'information [115, 188, 189]. Sa plateforme basée sur les entropies (stockage, transfert et modification) a fourni la première preuve quantitative directe de plusieurs conjectures importantes de longue date concernant la facilitation du calcul dans les automates cellulaires via des structures émergentes (Fig. 2.6). Autrement dit, Lizier a montré que les *blinkers* implémentent le stockage d'informations, que les *particules* en mouvement (*gliders* et *domain walls*) sont des agents de transfert d'informations dominants et que les *collisions de particules* sont des événements de modification d'informations. Cela a démontré que son cadre quantitatif s'alignait sur les concepts populaires de mémoire, de communication et de traitement.

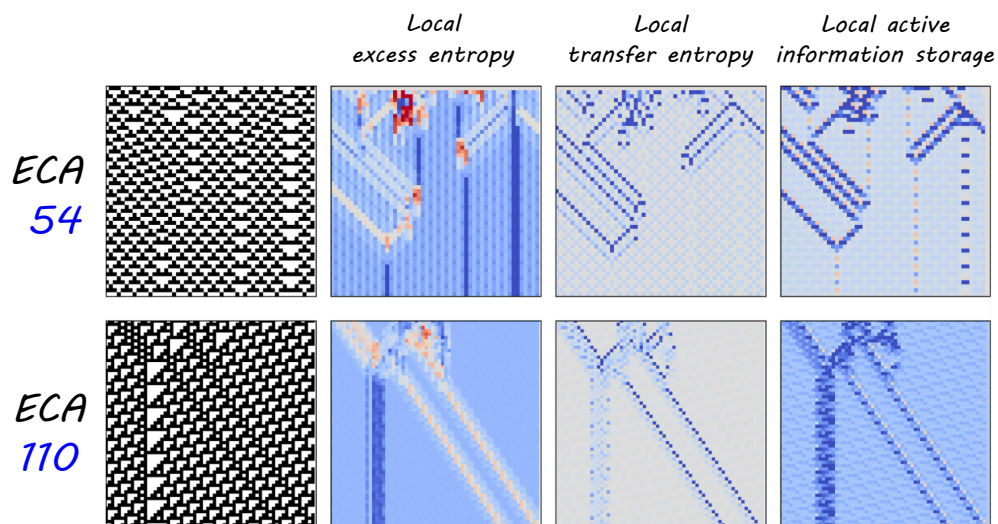


FIGURE 2.7 – Détection de structures cohérentes dans les automates cellulaires par la méthode quantitative de Lizier [188]. Mesures d'informations locales dans les automates cellulaires, appliquées aux simulations des règles 54 (en haut) et 110 (en bas). Pour toutes ces mesures, le bleu et le rouge indiquent respectivement les valeurs positives et négatives.

2.4 Conclusion

D'abord, nous avons présenté le lien entre la théorie de calculabilité et les systèmes complexes. Nous avons défini les opérations fondamentales permettant le traitement de l'information. Ensuite, nous avons mis en évidence l'hypothèse du 'edge of chaos' qui connecte le comportement complexe et les ingrédients d'une calculabilité optimale et efficace.

Dans la deuxième partie de ce chapitre, nous avons projeté concrètement ces notions sur les automates cellulaires. Dans un premier temps, nous avons défini ce type de systèmes spatialement étendus, et présenté leurs classifications. Puis, nous avons détaillé les différentes approches qui s'alignent sur le rôle computationnel des patterns émergent supportés par la dynamique des automates cellulaires.

Les automates cellulaires dans les états chimères optiques

3.1 Introduction

Comme nous l'avons vu dans les deux chapitres précédents, l'auto-organisation spontanée et les comportements collectifs complexes au sein des systèmes spatialement étendus ont fait l'objet d'un intérêt central pour les recherches en physique, en chimie et en biologie. L'intérêt pour les systèmes discrets couplés a été renouvelé après qu'il ait été démontré qu'ils supportent la coexistence de domaines cohérents et incohérents [32]. Les objets dynamiques résultants sont connus sous le nom d'états chimères [46] et leur étude est actuellement un sujet d'actualité en dynamique non linéaire. Une caractéristique remarquable des états chimères est la diversité des motifs produits par leur dynamique. Un cas précis a retenu notre attention. En effet, des chimères observées dans un réseau de guides d'ondes optiques 1D couplés localement, des patterns triangulaires ont été observés [26]. Nous avons remarqué dans le chapitre 1, qu'une telle évolution connue sous le nom de motifs SIERPIŃSKI est une solution non linéaire universelle [82, 83] et typique de l'intermittence spatio-temporelle. Remarquablement, nous avons constaté dans le chapitre 2, que les motifs de SIERPIŃSKI sont également un comportement typique dans une autre classe de systèmes discrets : les automates cellulaires élémentaires.

Les automates cellulaires font partie des modèles standards des systèmes dynamiques spatialement étendus, qui se démarquent des autres modèles par leur discrétion spatio-temporelle. Ils ont été introduits par WOLFRAM qui s'intéressait à la relation entre le comportement des systèmes dynamiques et leurs capacités de calcul [190]. En particulier, la

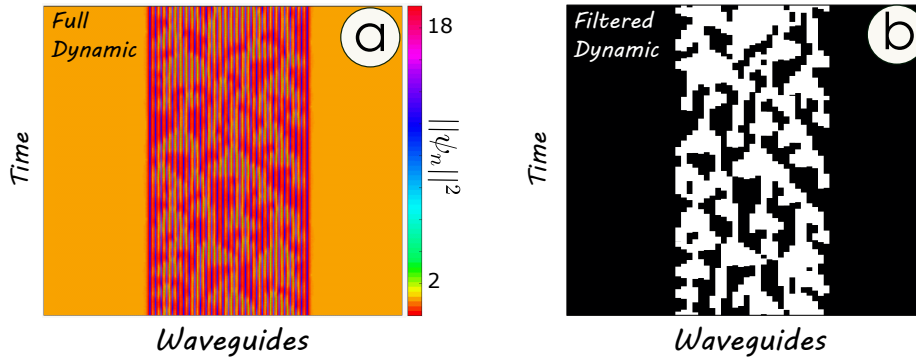


FIGURE 3.1 – États chimères optiques **OCS**. (a) : Évolution spatio-temporelle de l’intensité $\|\psi_n(t)\|^2$, montrant la dynamique complète. (b) : la version filtrée de (a), affichant une dynamique submergée de type automates cellulaires.

classe complexe IV des automates cellulaires qui supportent les structures localisées émergentes appelées *blinkers* et les structures propageantes nommées *gliders* (voir Fig. 3.2a). WOLFRAM a supposé que toutes les règles de classe IV étaient capables de calcul universel [190, 191]. Il convient de mentionner qu’une calculabilité optimale est conditionnée par l’équilibre entre trois fonctionnalités essentielles : *stockage d’informations*, *transfert d’informations* et *modification d’informations* [7].

Malgré toutes ces propriétés de complexité spatio-temporelle et ces capacités de traitement de l’information, les automates cellulaires restent principalement des systèmes de calcul discrets et abstraits. En effet, pour l’instant très peu d’implémentations d’automate cellulaire ont été rapportées dans des systèmes physiques.

Le but de ce chapitre est d’étudier la capacité de nos états chimères optiques (OCS) à fournir des propriétés calculatoires d’automates cellulaires élémentaires (ECA). A cet effet, nous nous sommes intéressés à la mise en évidence qualitative et quantitative d’une analogie entre nos OCS et les ECA, par le biais des outils de la théorie d’information notamment les notions d’entropies.

3.2 Aperçu qualitatif de la dynamique des ECA et nos OCS

La formation des patterns émergents dans des systèmes décentralisés spatialement étendus donne lieu à une certaine forme de comportement globalement coordonné, ou de traitement global de l’information, qui est utilisé par le système pour se soutenir et prendre certaines décisions. Ce traitement global de l’information dans des systèmes spatialement

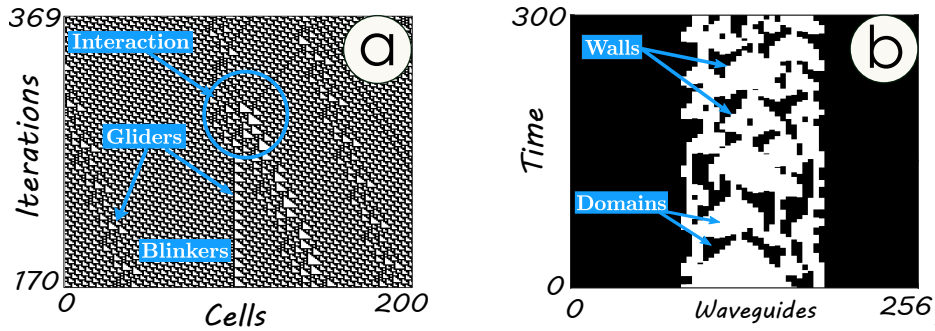


FIGURE 3.2 – Illustration de l’analogie phénotypique entre les patterns émergents montrés par nos OCS et les ECA complexes. (a) diagramme espace-temps illustrant le comportement typique de l’ECA complexe 110, à partir de conditions initiales aléatoires sur un réseau de cellules 300. (b) Évolution spatio-temporelle d’une OCS filtrée de l’intensité binarisée $||\psi_n(t)||^2$.

étendus, par le biais de la formation de patterns émergents, est connu sous le nom de calcul émergent [178].

Nous avons vu, dans le chapitre précédent, la façon dont la dynamique spatio-temporelle des automates cellulaires donne lieu à un calcul émergent. En effet, nous avons rapporté un certain nombre d’études effectuées dans l’optique d’identifier les mécanismes du calcul émergent. Ces études convergent vers l’attribution du stockage de l’information aux *blinkers*, la transmission de l’information aux *gliders* et la modification de l’information aux *collisions* entre ces derniers (Fig. 3.2a).

Maintenant, prenons la dynamique spatio-temporelle montrée par nos OCS. Après un processus de binarisation décrit dans (3.3.1), nous obtenons l’évolution de la Fig. 3.1b). Il convient de noter que ces OCS démontrent la robustesse, la flexibilité et l’évolutivité de leurs patterns spatio-temporels. En effet, par la simple modification des paramètres d’entrée, on peut favoriser la génération d’une large gamme de structures géométriques dans l’espace-temps, conduisant à plus de complexité. La présence de ces patterns triangulaires de type SIERPINSKI dans les domaines incohérents de nos OCS a motivé l’analogie établit tout au long de ce chapitre.

Guidés par l’importance des structures émergentes dans le calcul, on peut spéculer que la dynamique spatio-temporelle montrée par nos OCS présente de nombreux ingrédients de base nécessaires pour effectuer le traitement de l’information, du point de vue de la théorie de la calculabilité. Par exemple, les domaines, les particules embarquées (Walls entre les domaines) et les collisions entre ces particules illustrées sur la Fig. 3.2b).

Afin de confirmer davantage notre hypothèse, nous avons adopté le cadre quantitatif

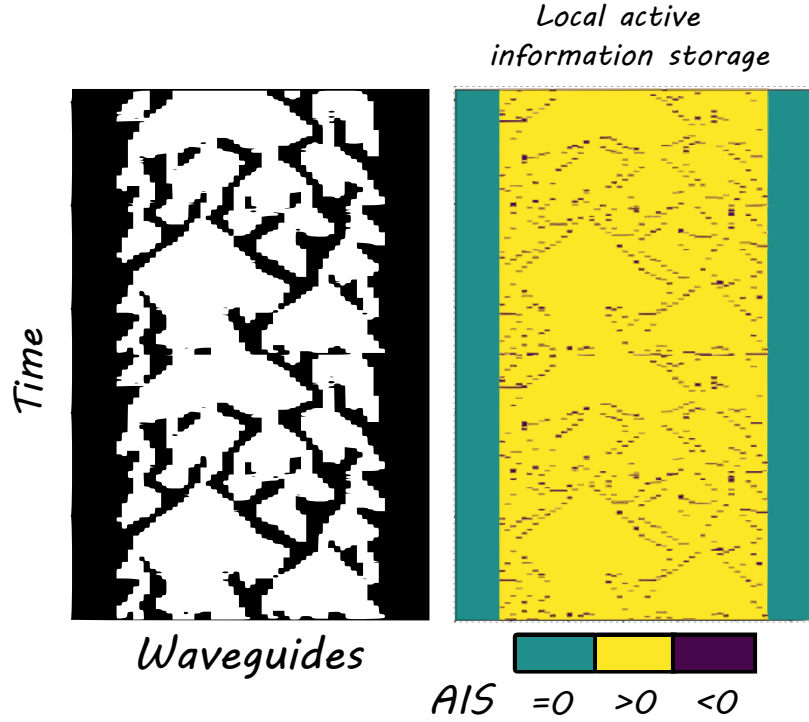


FIGURE 3.3 – Mise en évidence des rôles computationnels des patterns émergents dans nos OCS par la méthode quantitative de LIZIER [192]. (a) : Diagramme spatio-temporel filtré de l'intensité binarisée $||\psi_n(t)||^2$. (b) : Profil local de **AIS** sur (a).

de LIZIER (chapitre 2) qui concerne la dynamique de l'information au sein des calculs distribués (Fig. 3.3).

Nous avons calculé le *Stockage actif local d'informations* (**AIS**) dans la dynamique spatio-temporelle de nos OCS (Fig. 3.3). En effet, il s'agit d'une grandeur proposée par LIZIER dans [192] afin de quantifier *localement* deux opérations fondamentales dans le cadre du traitement de l'information qui se réside dans l'*information storage* et l'*information transfer*. Les détails de la méthode de calcul sont fournis dans l'annexe A.

Concernant l'*information storage*, LIZIER l'a définie comme '*l'information dans le passé d'un agent ou d'une variable qui peut être utilisée pour prédire son avenir*'. En appliquant ce concept sur nos OCS, nous avons remarqué que les profils de **AIS** (Fig. 3.3b) montrent que les domaines réguliers triangulaires représentent les principales entités de stockage d'informations (**AIS** > 0). En effet, les états des cellules à l'intérieur des domaines sont fortement prévisibles à partir de leur histoire passée, car ils sont temporellement corrélés. D'autre part, les composants locaux négatifs du **AIS** sont concentrés dans les zones de déplacement des particules ou ce que nous appelons les Walls (**AIS** < 0) (Fig. 3.3b). Ces profils locaux de **AIS** fournissent un bon filtre spatio-temporel des Walls. En effet, lorsqu'un

Wall est rencontré dans une cellule donnée, l'historique de cette cellule (faisant partie du domaine d'arrière-plan) est peu informatif sur l'état suivant, car la séquence de domaine était plus susceptible de continuer que d'être interrompu. Par contre, ces mesures au niveau local ont révélé que les Walls (séparant les domaines réguliers) sont les entités dominantes en terme du transfert d'informations.

Finalement, on peut conclure que nos résultats s'alignent avec ceux établis sur les automates cellulaires concernant les rôles computationnels des patterns émergents. En effet, cette analogie qualitative qui a motivé la première partie du travail de ce chapitre, était confirmée par la ressemblance phénotypique et fonctionnelle des domaines réguliers et des Walls dans nos OCS, aux *blinkers* et aux *gliders* dans les ECA.

3.3 Aperçu quantitatif de la dynamique des ECA et nos OCS

3.3.1 Méthode d'analyse de données

Les comportements ordonnés et aléatoires sont simples à discerner visuellement ou à décrire statistiquement [110]. Néanmoins, les comportements et les patterns complexes sont faciles à identifier qualitativement, mais leur quantification reste un casse-tête lancinant [121]. Comme mentionné précédemment, la reconnaissance du rôle fondamental des structures émergentes et interactives a conduit une panoplie d'études statistiques à résoudre ce problème de quantification, en essayant de définir un emplacement relatif où se trouvent des systèmes physiques au comportement complexe [7, 121, 169, 192-194]. C'est à la lumière de ces outils statistiques que nous mènerons une analyse quantitative des dynamiques spatio-temporelles présentées par les ECA et nos OCS. Cependant, nous concentrerons nos efforts sur la recherche de grandeurs permettant de classer les ECA selon leur évolution dynamique.

Tout d'abord, nous classons les différents types de dynamique dans les ECA, en particulier, pour filtrer la classe complexe. Ensuite, nous étendons l'étude à la dynamique de nos OCS, en cherchant à quelle classe d'ECA, elle pourrait être comparée, et nous permettant enfin de construire le socle de l'analogie prometteuse proposée tout au long de ce chapitre. Le concept d'entropie pour quantifier la quantité d'informations perdue et pour mesurer le degré d'aléatoire et d'imprévisibilité dans un processus dynamique sera notre outil clé. En effet, nous adoptons la complexité de Lempel-Ziv (LZ) [195, 196], l'entropie approximative

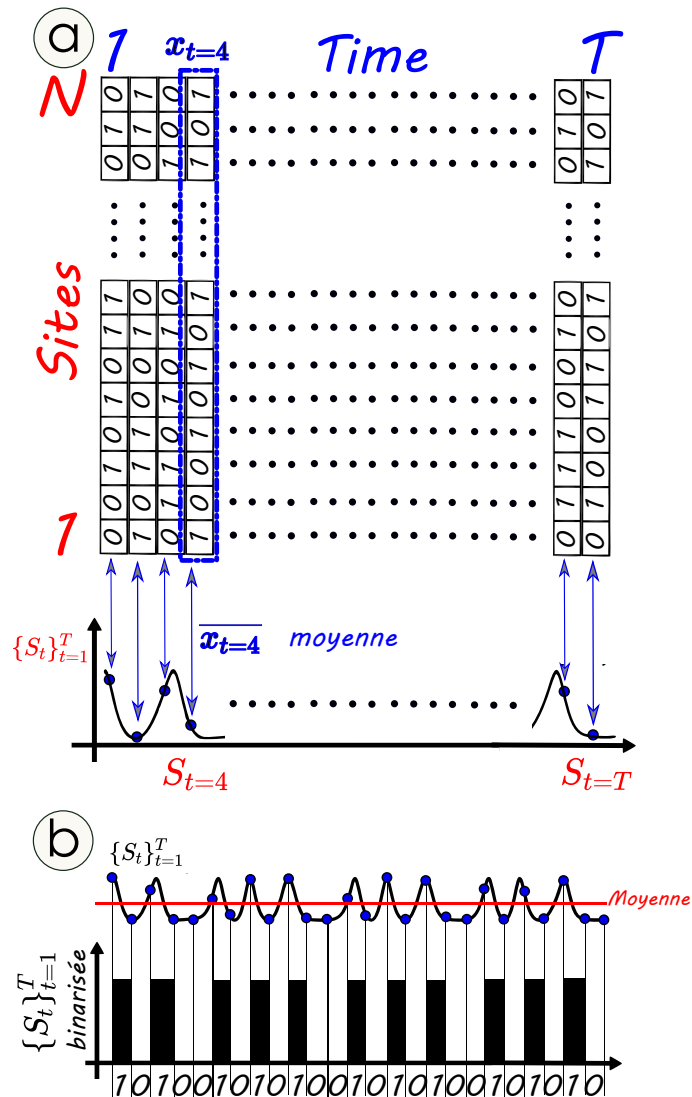


FIGURE 3.4 – Méthode d’analyse des données : (a) : Construction d’une série temporelle à partir d’un diagramme spatio-temporel binarisé. (b) : schéma de binarisation obligatoire des séries temporelles pour l’estimation de la complexité LZ.

(ApEn) [197], et l’entropie de l’échantillon (SampEn) [198]. Ces trois grandeurs ont des aspects théoriques différents, mais toutes ont leurs racines dans la théorie de l’information et la théorie du chaos. Aussi, elles présentent un lieu commun dans l’analyse des séries temporelles discrètes (voire l’annexe A pour le formalisme mathématique).

Il faut souligner que ces trois quantités déterministes sont maximisées pour les processus aléatoires et s’annulent pour les processus parfaitement ordonnés [195-199]. Inspirée de l’ensemble de la littérature traitant la caractérisation de la complexité au sein de signaux expérimentaux, à l’aide d’une seule observable [199-201], notre étude sera basée sur une

analyse de séries temporelles.

Décrivons les grandes lignes de notre méthode. En fait, les données spatio-temporelles d'ECA ou d'OCS sont binarisées si nécessaire et disposées sous la forme $\{x_t\}_{t=1}^T = \{x_1, x_2, \dots, x_T\}$ où $x_t = (x_t(1), x_t(2), \dots, x_t(N))$ est la configuration à l'instant t d'un système de N sites (Fig. 3.4a). Chaque composante $x_t(i)$ donne l'état du i -ème site à l'instant t . Ensuite, nous ne considérerons que la valeur moyenne \bar{x}_t à chaque pas de temps. En suivant ce schéma, nous pouvons construire une série temporelle $\{S_t\}_{t=1}^T = \{S_1, S_2, \dots, S_T\}$ où $S_t = \bar{x}_t \in [0,1]$, pour chaque diagramme spatio-temporel (Fig. 3.4a). C'est sur ces données moyennées que nous appliquons les trois algorithmes statistiques choisis, à savoir LZ-ApEn-SampEn, pour quantifier le degré d'aléatoire et classer les différents types de dynamiques montrées par les ECA et nos OCS. Il faut souligner que les données doivent être de type binaire pour l'estimation de la complexité LZ. Ainsi, pour cette dernière quantité le $\{S_t\}_{t=1}^T$ sera binarisé suivant une valeur seuil donnée par sa valeur moyenne (Fig. 3.4b).

Concernant ECA, des simulations ont été effectuées pour chaque règle, sur un réseau de 300 cellules sur 3000 itérations où les 500 premiers états transitoires sont ignorés. Ensuite, le processus est répété pour 100 conditions initiales aléatoires.

Concernant nos OCS, le système optique évolue pendant T unités de temps et son intensité binarisée $|\psi_n(t)|^2$ est stockée dans un tableau $N \times \frac{T}{dt}$ ($T = 250$, $dt = 0.0833$), pour différentes tailles de région désynchronisée (δ_L). Le processus de binarisation a été effectué grâce à une valeur de coupure d'intensité, nous permettant de filtrer l'évolution spatio-temporelle de l'OCS pour mettre en évidence les patterns similaires à ceux qui surviennent dans l'ECA (Fig. 3.1b).

3.3.2 Résultats et discussion

Classification par clustering

Pour la classification des ECA, nous utilisons la stratégie de clustering. Par conséquent, pour chaque ECA, nous avons calculé l'ApEn, la complexité LZ et la SampEn pour des centaines de conditions initiales aléatoires différentes. Dans les Fig. 3.5a) et b), nous avons représenté ces entropies dans le plan (LZ, ApEn) et (LZ, SampEn), respectivement. Comme on peut le voir sur ces figures, nous avons différents clusters dans les plans d'entropie.

Remarquez que nous retrouvons les classifications précédemment proposées dans [170, 190, 202] (Tab. 3.1). La classe des règles périodiques représente la région ordonnée (nuage de points orange), ayant de faibles valeurs d'entropie, tandis que la classe des règles chaotiques avec des valeurs d'entropie élevées, représente la région désordonnée (nuage de points

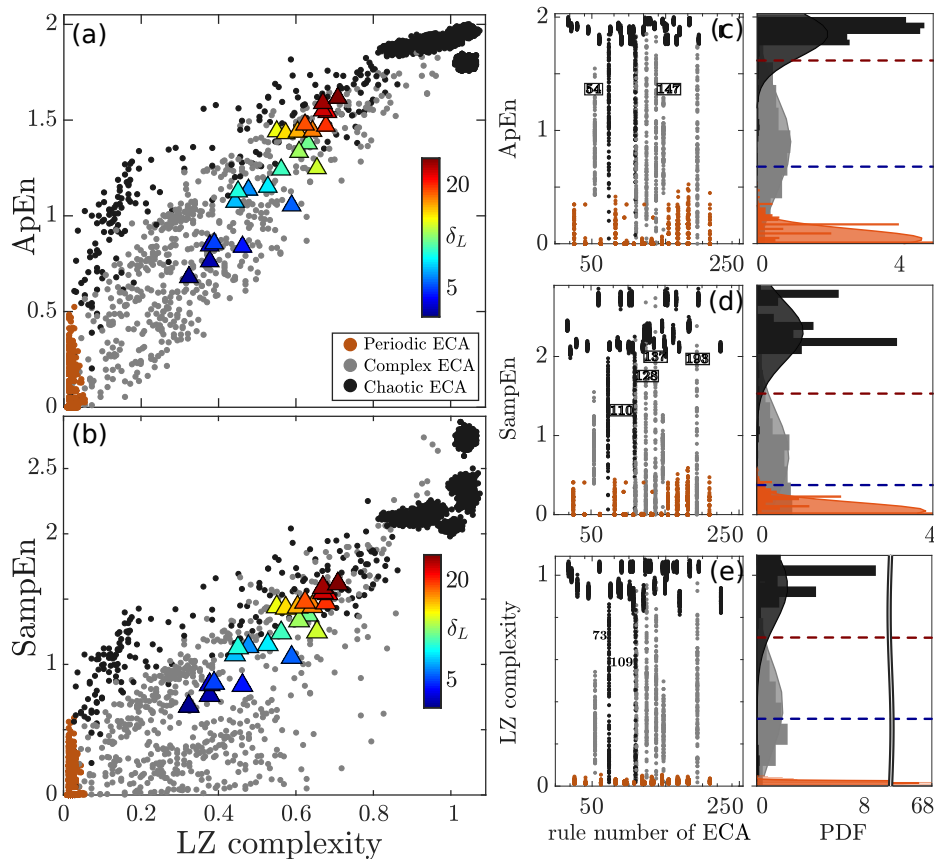


FIGURE 3.5 – "Edge of chaos" : Aperçu statistique de la dynamique spatio-temporelle de nos OCS et des ECA. (a) et (b) correspondent respectivement aux cartes ApEn vs LZ et SampEn vs LZ. Les points correspondent aux ECA et les triangles à nos OCS. Les triangles sont obtenus pour différentes tailles δ_L d'OCS, avec $C = 1.9$, $\Delta = 7$ et $E_0 = 6.05$. Les panneaux de gauche montrent les trois grandeurs statistiques avec leurs histogrammes de fréquence (PDF) : ApEn (c), SampEn (d) et complexité LZ (e), pour toutes les règles ECA. Les valeurs minimales (bleues) et maximales (rouges) de nos OCS sont illustrées par des lignes en pointillés.

noirs). Comme prévu, les classes de règles à comportement complexe (nuage de points gris), caractérisées par une entropie intermédiaire, sont susceptibles de se trouver dans la région critique, située entre les régions ordonnée et désordonnée. Nous avons également représenté les trois entropies, pour chaque règle ECA et toutes les conditions initiales aléatoires dans les Figs. 3.5c), d) et e). Remarquablement, nous observons à partir de ces figures une corrélation intéressante concernant les valeurs à caractère aléatoire calculées à partir du comportement d'une règle ECA donnée et de ses équivalents, sur l'échantillon de conditions initiales aléatoires, telles que les règles complexes : 54 et 147 ou la règle 110 et 124, 137, 193 (Tab. 3.2).

TABLEAU 3.1 – Classification des ECA selon LI et PACKARD (1990) [170].

Classification		
classe	nombre	règles
Nulle	8	0, 8, 32, 40, 128, 136, 160, 168
Point fixe	32	2, 4, 10, 12, 13, 24, 34, 36, 42, 44, 46 56, 57, 58, 72, 76, 77,78, 104, 130, 132 138, 140, 152, 162, 164, 170, 172, 184 200, 204, 232
Périodique	32	1, 3, 5, 6, 7, 9, 11, 14, 15, 19, 23, 25 27, 28, 29, 33, 35, 37,38, 41, 43, 50 51, 62, 74, 94, 108 134, 142,154, 156, 178
Complexe	2	54, 110
Localement chaotique	1	73
Chaotique	13	18, 22, 30, 45, 60, 90, 105, 106, 132 126,146, 150, 122

Ensuite, nous avons étendu notre analyse à la dynamique de nos OCS. Les résultats sont donnés par les triangles des Figs. 3.5a) et b). On peut voir qu'ils appartiennent au nuage de classe complexe (nuage de points gris). Cela montre que notre méthodologie capture et quantifie parfaitement le comportement complexe et confirme également que ce dernier réside dans la transition de la phase régulière vers la phase chaotique.

Pour tester la robustesse de nos résultats, des calculs ont été effectués pour différentes tailles de réseau (N) d'ECA sur la Fig. 3.6 et différents pas de temps (longueur de la série temporelle L) sur la Fig. 3.7, ne montrant aucun changement dans nos résultats. De plus, notre technique montre le potentiel de classer les configurations spatio-temporelles, pour un spectre de dynamiques ordonnées, complexes et chaotiques, sans scinder l'analyse spatiale et temporelle, grâce à la compression de l'information spatiale dans une série temporelle observable telle que la valeur moyenne (Fig. 3.4).

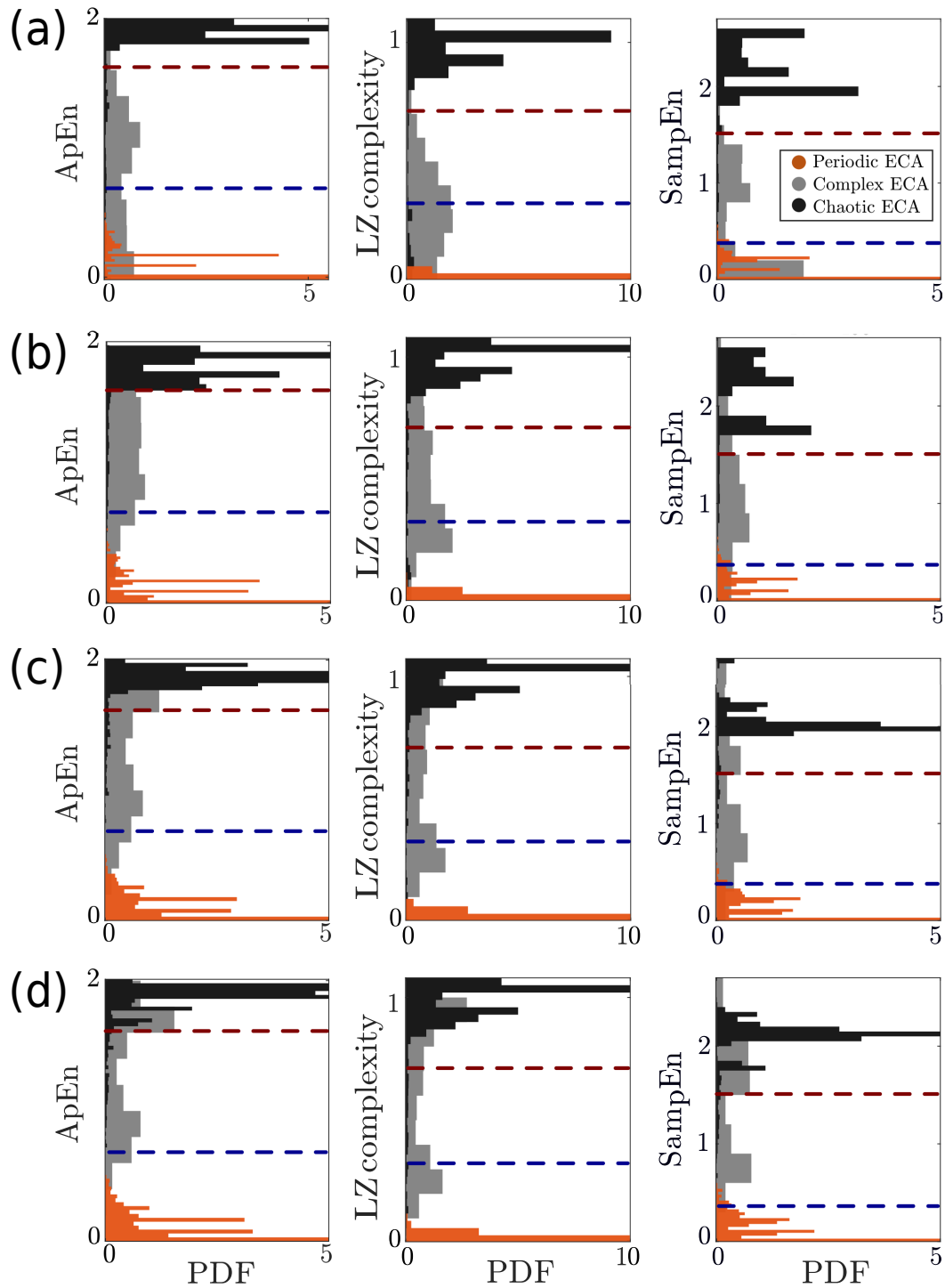


FIGURE 3.6 – Robustesse des résultats : différentes tailles de réseau des ECA. (a) : $N = 200$, (b) : $N = 400$, (c) : $N = 600$, (d) : $N = 800$. Les valeurs minimales (bleues) et maximales (rouges) de nos OCS sont illustrées par des lignes pointillées (données de la Fig. 3.5).

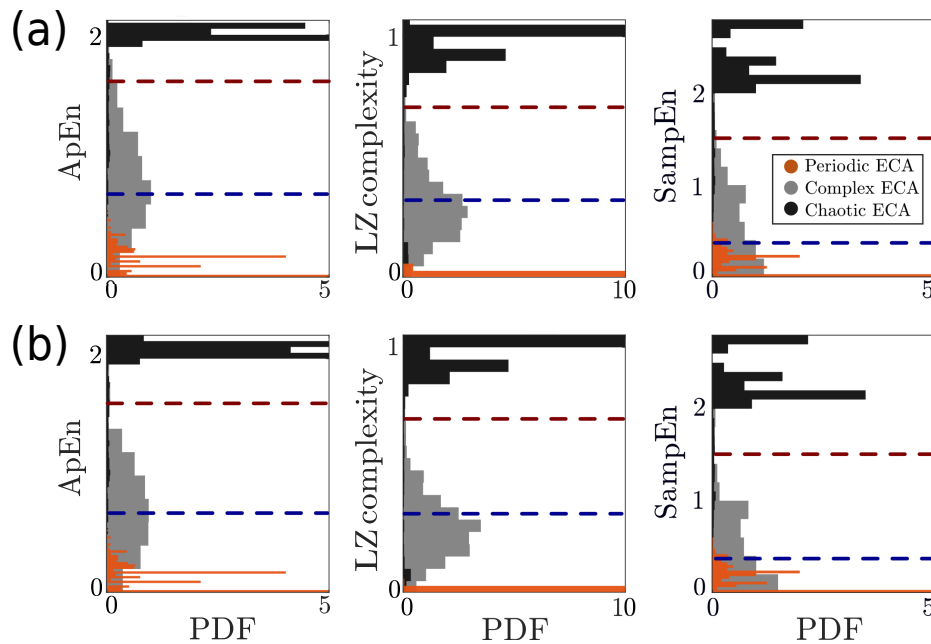


FIGURE 3.7 – Robustesse des résultats : différentes longueurs de séries temporelles (L) pour $N = 300$ taille de réseau ECA. (a) : $L = 3500$, (b) : $L = 4500$. Les valeurs minimales (bleues) et maximales (rouges) de nos OCS sont illustrées par des lignes pointillées (données de la Fig. 3.5).

Calculabilité de nos OCS

Hypothèse 'edge of chaos'

En rassemblant les résultats ci-dessus, nous pouvons conclure que nos OCS peuvent être classés dans la classe de comportement complexe des ECA. Cette classe joue un rôle central dans l'hypothèse de "edge of chaos" (EOC), qui à son tour est un concept clé de la théorie de l'information (voire Chapitre 2).

Comme nous l'avons exposé dans le chapitre précédent, les défenseurs de cette fameuse hypothèse stipulent que les systèmes positionnés dans un tel régime critique améliorent remarquablement leurs capacités de calcul. Plus loin, ils ont connecté le comportement complexe au traitement sophistiqué d'information via l'adaptivité et l'évolution.

Le raisonnement statistique derrière cette conjecture est que le **stockage d'informations** nécessite une faible entropie (ordre), tandis que la *transmission d'informations* nécessite une entropie élevée (désordre), ainsi le régime EOC peut permettre le compromis optimal [7]. D'autre part, comme déjà mentionné(chapitre2), le fait d'exister dans la région EOC n'est pas suffisant pour une meilleure calculabilité. En effet, Il faudra accomplir la

troisième opération qui se réside dans la **modification d'informations**. Concernant les automates cellulaires, nous avons vu qu'une panoplie d'études (chapitre 2) ont attribué ces fonctionnalités très importantes aux *gliders* et *blinkers* (Fig. 3.2). Ce dernier point est vérifié dans le cas de nos OCS dans la section précédente 3.2.

Dépendance aux paramètres de contrôle

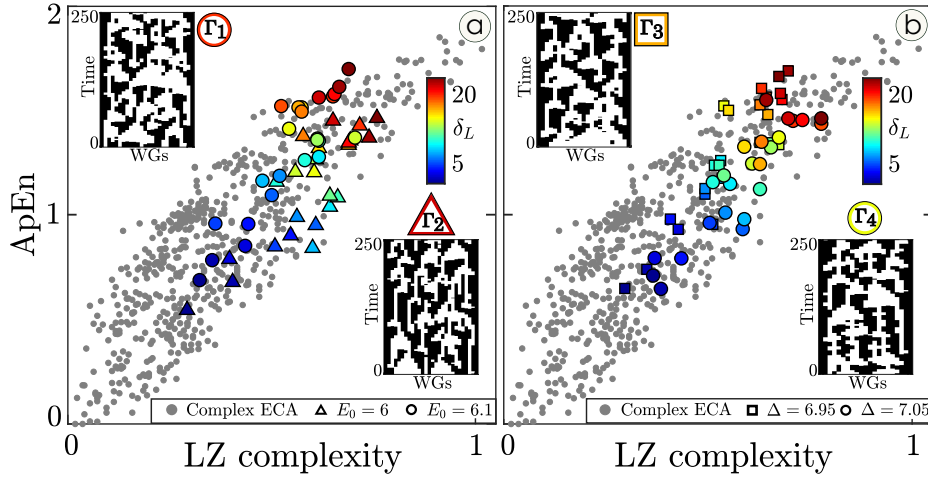


FIGURE 3.8 – Sensibilité des ECA et de nos OCS aux paramètres de contrôle. Deux cartes : ApEn contre la complexité LZ (données de la Fig. 3.5) affichant uniquement le nuage de la classe complexe ECA (gris) et les différentes tailles d'OCS (δ_L) générées avec différents paramètres d'entrée (triangle, carré et marqueurs de cercle). (a) : $C = 1.9$, $\Delta = 7$ et $E_0 = 6, 6.1$ différent ; (b) : $C = 1.9$, $E_0 = 6, 05$ et différents $\Delta = 6.95, 7.05$. Γ_1 ($\delta_L = 20$), Γ_2 ($\delta_L = 22$), Γ_3 ($\delta_L = 17$) et Γ_4 ($\delta_L = 15$) sont des évolutions dynamiques sélectionnées pour illustrer la configurabilité de nos OCS.

Considérons maintenant la capacité du système à effectuer une tâche de calcul utile [127]. Dans ce contexte, l'équation du mouvement qui régit le système est considérée comme le programme/algorithm, la condition initiale comme l'entrée et la configuration souhaitable comme la sortie.

ZENIL [203, 204] a exploré les liens possibles entre le comportement dynamique d'un système et l'universalité de TURING en termes de capacité du système à (efficacement) transmettre et manipuler l'information. Dans son travail, ZENIL a fournis certains arguments à l'aide d'un coefficient de transition défini basé sur la compression, qui quantifie la sensibilité d'un système à la programmation. A l'issue de ces études, ZENIL a établi la conjecture suivante [203] : "nous conjecturons qu'un système sera capable de calcul universel s'il a un grand coefficient de transition, au moins supérieur à zéro, par exemple."

Dans la même ligne directrice, Wolfram a émis l'hypothèse que la classe ECA complexe est capable de calcul universel, en raison de l'indécidabilité concernant son comportement [190, 191]. Dans notre étude, cette indécidabilité est reflétée par les grandes valeurs d'intervalle des mesures statistiques prises par la classe complexe des ECA située dans la transition de phase entre les classes ordonnées et chaotiques (voir Fig. 3.5). Notez que l'imprévisibilité de la dynamique de nos OCS par rapport aux conditions initiales a été démontrée en calculant les exposants de LYAPUNOV [26]. Par conséquent, nos OCS peuvent être considérés comme un système dynamique complexe avec les propriétés des ECA.

De plus, ZENIL a conjecturé [204] : *"la capacité d'un système à effectuer des calculs universels implique qu'il soit capable de supporter des informations en reflétant les perturbations apportées à l'entrée et de les transmettre à la sortie"*. En d'autres termes, un système universel doit montrer sa sensibilité aux stimuli externes.

Une caractéristique intéressante de nos OCS est que leurs propriétés peuvent être contrôlées par l'ensemble des paramètres pertinents du système. Une illustration de cette fonctionnalité peut être observée sur la Fig. 3.8. Sur cette figure, on montre dans le plan (LZ, ApEn) les entropies de la classe complexe ECA et de nos OCS pour différentes valeurs des paramètres de pompe E_0 (marqueurs de couleur sur la Fig. 3.8a) et différentes valeurs du désaccord Δ (marqueurs de couleur sur la Fig. 3.8b). La carte des couleurs dans les deux figures représente le changement de la taille du domaine incohérent δ_L de nos OCS. Il apparaît que les propriétés des ECA héritées par nos OCS peuvent être ajustées en continu par presque tous les paramètres du système ainsi que par le profil de l'état chimère. Par conséquent, avec notre méthode de clustering, nous sommes en mesure de quantifier le changement qualitatif entre les différents types d'évolution spatio-temporelle au fur et à mesure que les paramètres changent (voir les diagrammes spatio-temporels en encadré de la Fig. 3.8).

Comportement de mise à l'échelle

Pour compléter notre analogie, nous avons approfondi l'étude pour comparer les propriétés de nos OCS et celles des règles complexes ECA. Nous considérons la dynamique à long terme. Divers auteurs [7, 205] ont rapporté qu'au voisinage de la transition de phase, les ECA peuvent présenter de longs transitoires qui sont fonction de la taille du réseau (Fig. 3.9).

Par conséquent, on peut s'attendre à ce que les entropies soient aussi fonction de la taille. C'est ce que nous avons observé sur la Fig. 3.10a), où nous avons tracé la moyenne de la valeur d'entropie estimée par ApEn sur 100 configurations initiales aléatoires en

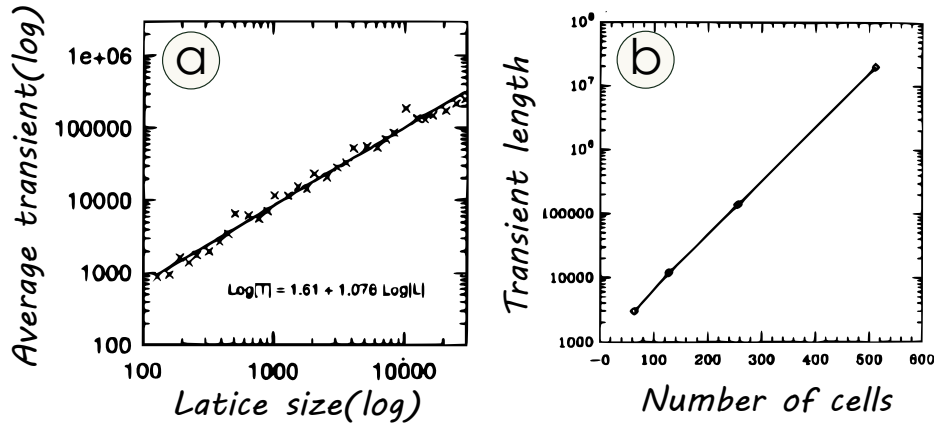


FIGURE 3.9 – Comportement d'échelle du temps transitoire des automates cellulaires dans la région du *Edge of chaos*. (a) : Temps transitoire moyen de la règle 110 en fonction de la taille du système N (dans l'échelle logarithmique)[205]. (b) : Croissance des transitoires moyens en fonction de la taille du réseau pour $\lambda = 0, 50$ [7].

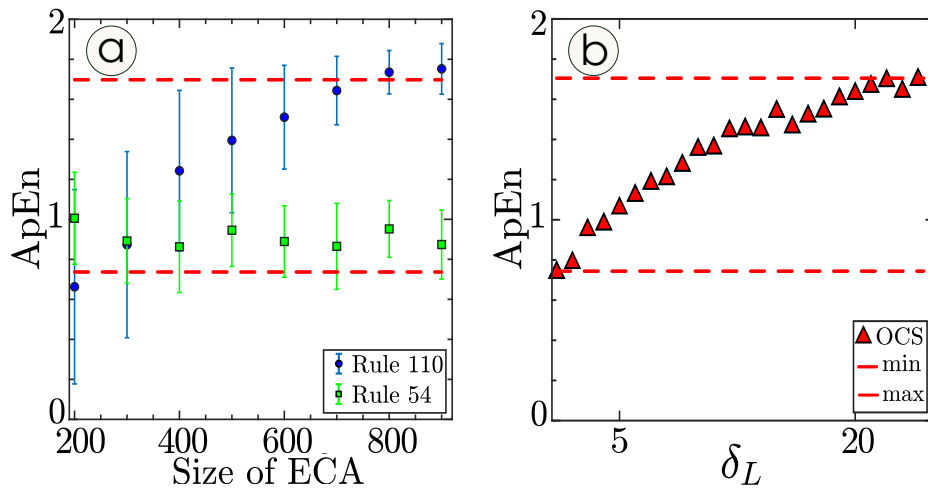


FIGURE 3.10 – Comportement d'échelle des règles complexes des ECA et de nos OCS : (a) et (b) représentent l'effet de taille sur ApEn des ECA complexes et nos OCS, respectivement.

fonction de la taille du réseau N pour les règles complexes ECA 110 (bleu) et 54 (vert). Contrairement à la règle 54, qui reste moins complexe, la dynamique de la règle 110 montre une sensibilité à la taille, à travers la tendance à la croissance d'ApEn avec la taille du système. Ce comportement de mise à l'échelle de la règle 110 peut s'expliquer comme une conséquence directe de son temps transitoire, qui augmente avec la taille du système [205].

Ici aussi, nous observons le même comportement avec notre OCS comme on peut le voir sur la Fig. 3.10b) où nous montrons l'ApEn en fonction de la taille de la chimère δ_L .

Notez que, COOK [206] a prouvé que la règle 110 était universelle (ou Turing) en mon-

trant qu'il était possible d'utiliser la règle pour émuler n'importe quel autre modèle de calcul. Il a montré l'importance des longs transitoires afin de garantir le maximum d'opérations via les structures (*gliders* et *blinkers*) et donc atteindre une calculabilité universelle.

Le bruit $(1/f)^\alpha$

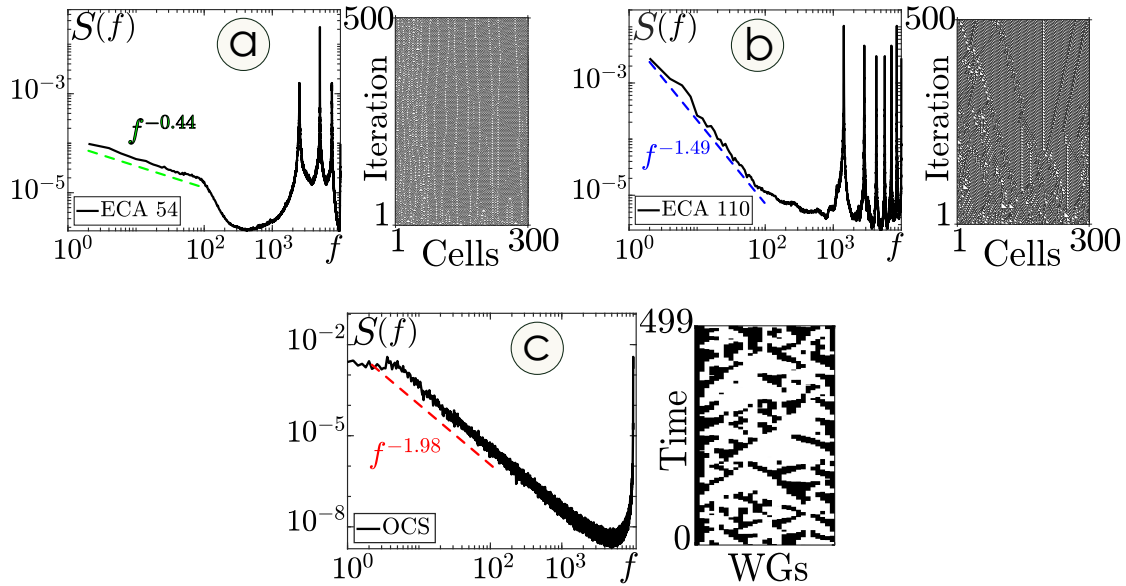


FIGURE 3.11 – Analyse spectrale des règles ECA complexes et de nos OCS. (a), (b) et (c) : les panneaux de gauche affichent le spectre de puissance $S(f)$ du comportement temporel présenté par les règles complexes ECA (54, 110) et nos OCS, respectivement. Les lignes pointillées représentent l'ajustement log-log du spectre de puissance dans la gamme des basses fréquences $f = 2 - 10^2$ montrant une loi de puissance avec différentes pentes. Les panneaux de droite affichent les comportements spatio-temporels typiques des règles ECA complexes et du domaine incohérent de nos OCS.

Une autre propriété suggérée pour avoir une relation avec l'universalité de calcul est la propriété du bruit $1/f^\alpha$ du spectre. Par conséquent, pour explorer plus d'horizons concernant l'analogie entre nos OCS et les règles complexes des ECA, nous avons effectué une analyse spectrale pour étudier le comportement temporel de leurs dynamiques. En effet, une classification des ECA basée sur la forme de leurs spectres de puissance $S(f)$ a été proposée dans [207].

Observation intéressante, le spectre de puissance $S(f)$ des règles ECA complexes (par exemple 110 et 54), présente une loi de puissance aux basses fréquences [207] comme illustré sur la Fig. 3.11a) et b). Les calculs ont été effectués à partir d'évolutions partant d'une

configuration initiale aléatoire de 3000 cellules pour 5000 pas de temps. Cette décroissance spectrale, connue sous le nom de bruit $1/f^\alpha$ [208], reflète la forte influence des événements passés sur le futur. NINAGAWA a établi que seule la règle 110 présente le bruit $1/f^\alpha$ pendant les pas de temps les plus longs, grâce à son comportement transitoire étendu [205].

Le problème le plus controversé dans le bruit $1/f^\alpha$ est de savoir si il est permanent ou non. D'une manière générale, s'il existe un temps de corrélation fini τ dans une fluctuation, le spectre de puissance avec une longueur d'observation de $T > \tau$ a une densité de puissance presque égale à des fréquences inférieures à $1/\tau$. De même, s'il y a un temps de corrélation fini T dans l'évolution d'un automate cellulaire, le spectre de puissance avec une longueur d'observation de $T > \tau$ a une densité de puissance presque égale à des fréquences inférieures à T/τ . Ainsi, le spectre de puissance devient proche d'un plateau aux basses fréquences lorsque la longueur d'observation T devient plus longue que le temps de corrélation τ . De plus on peut supposer que le temps de corrélation τ dépend de la taille du réseau N car le pas de temps transitoire moyen T_{ave} de la règle 110, supposée pertinente pour τ , augmente algébriquement avec la taille du réseau N , $T_{ave} = N^\alpha$ [205].

NINAGAWA a affirmé que le comportement transitoire génère de l'intermittence et cause le bruit $1/f^\alpha$ (les phases chaotiques et périodiques, produites par la collision ou non des **gliders**). Ainsi, partant du constat que l'intermittence représente l'un des principaux mécanismes conduisant au bruit $1/f^\alpha$, dans un système dynamique chaotique [209], nos OCS intermittents [26] tirent leur légitimité de montrer un spectre de puissance dont la forme ressemble plus à celle présentée par l'ECA 110 comme on peut le voir sur les Fig. 3.11b) et c), respectivement. Plus particulièrement, la présence des fluctuations $1/f^\alpha$ dans deux automates cellulaires computationnellement universelles, telles que ECA 110 [206] et le jeu de la vie [210], a conduit NINAGAWA à conjecturer une relation entre universalité computationnelle et le bruit $1/f^\alpha$ dans les automates cellulaires [211].

3.4 Conclusion

En conclusion, les états chimères sont des structures dynamiques avec une caractéristique intrigante pour soutenir la coexistence de domaines cohérents et incohérents dans un réseau d'oscillateurs identiques. Ce sont des objets universels qui ont été observés dans une grande variété de systèmes dynamiques. Dans ce travail, nous avons considéré les états chimères présentés par un réseau de guides d'ondes optiques couplés localement. Nous avons étudié la capacité des états chimères optiques à imiter la dynamique des automates cellulaires, qui ont de nombreuses applications en mécanique computationnelle. Pour cela,

nous avons dans un premier temps, classifié les automates cellulaires élémentaires à l'aide d'une méthode de clustering basée sur trois métriques : complexité LZ, ApEn et SampEn. Par conséquent, les états chimères optiques peuvent être projetés dans cette carte pour identifier de quelle classe d'automates cellulaires ils héritent leurs propriétés et le type de calcul qu'ils peuvent effectuer. À notre connaissance, il s'agit de l'une des premières études suggérant que l'état chimère peut offrir des propriétés d'objet computationnel. Cela ouvre la possibilité d'une implémentation physique d'automates cellulaires, avec des applications prometteuses dans de nombreux domaines.

En effet, dans le chapitre suivant, nous allons explorer ces performances de calcul dans le cadre de réseaux de neurones récurrents.

TABLEAU 3.2 – 88 règles ECA et leurs équivalentes.

régle	régle équivalente	régle	régle équivalente
0	255	56	98, 185, 227
1	127	57	99
2	16, 191, 247	58	114, 163, 177
3	17, 63, 119	60	102, 153, 195
4	223	62	118, 131, 145
5	95	72	237
6	20, 159, 215	73	109
7	21, 31, 87	74	88, 173, 229
8	64, 239, 253	76	205
9	65, 111, 125	77	-
10	80, 175, 245	78	92, 141, 197
11	47, 81, 117	90	165
12	68, 207, 221	94	133
13	69, 79, 93	104	233
14	84, 143, 213	105	-
15	85	106	120, 169, 225
18	183	108	201
19	55	110	124, 137, 193
22	151	122	161
23	-	126	129
24	66, 189, 231	128	254
25	61, 67, 103	130	144, 190, 246
26	82, 167, 181	132	222
27	39, 53, 83	134	148, 158, 214
28	70, 157, 199	136	192, 238, 252
29	71	138	174, 208, 224
30	86, 135, 149	140	196, 206, 220
32	251	142	212
33	123	146	182
34	48, 187, 243	150	-
35	49, 59, 115	152	188, 194, 230
36	219	154	166, 180, 210
37	911	56	98
38	52, 155, 211	160	250
40	96, 235, 249	162	176, 186, 242
41	97, 107, 121	164	218
42	112, 171, 241	168	224, 234, 248
43	113	170	240
44	100, 203, 217	172	202, 216, 228
45	75, 89, 101	178	-
46	116, 139, 209	184	226
50	179	200	236
51	-	204	-
54	147	232	-

Reservoir Computing : Echo States Networks

4.1 Introduction

Dans le premier travail de cette thèse, nous avons étudié les possibilités calculatoires offertes par la complexité spatio-temporelle des systèmes spatialement étendus. L'auto-organisation à travers ses différentes signatures, a motivé la mise en évidence d'une analogie qualitative et quantitative entre les dynamiques supportées par nos états chimères optiques et celles d'un système computationnel par excellence qui réside dans les automates cellulaires élémentaires. Nous avons conclu que nos états chimères optiques, grâce à leur richesse dynamique, possèdent les ingrédients nécessaires permettant d'effectuer le traitement d'information.

Personne ne peut nier les capacités computationnelles incontournables du cerveau humain dans le traitement d'information. Une excellence qui a attisé la convoitise des chercheurs dans le domaine des neurosciences [15], souhaitant comprendre et déterminer les propriétés et processus biologiques qui garantissent au cerveau l'accomplissement de diverses tâches. Plusieurs études ont démontré que cette intelligence computationnelle est mise en évidence par l'activité neuronale du cerveau humain [15, 212, 213]. C'est ainsi que nous assistons à une révolution dans le domaine de l'apprentissage automatique (*machine learning*), inspiré essentiellement par le fonctionnement des neurones biologiques du cerveau.

Le *Reservoir Computing* représente un carrefour interdisciplinaire qui réunit le *machine learning*, les systèmes complexes et la dynamique non-linéaire. Cette interdisciplinarité fait de ce nouveau paradigme d'apprentissage automatique, un chantier fertile pour explorer les capacités computationnelles de nos états chimères optiques.

Ce chapitre a pour but de fournir les éléments théoriques nécessaires dans la compréhension du *Reservoir Computing*. Nous commençons par une recette de base sur la façon de créer des réservoirs de type *Echo States Networks*, et un aperçu des hyper-paramètres essentiels pour contrôler ses dynamiques et ainsi maximiser ses performances. Ensuite, nous abordons les propriétés souhaitables dans un bon réservoir. Enfin, nous présentons un état de l'art sur les implémentations physiques de cette technique du *machine learning*.

4.2 Machine learning et systèmes complexes

4.2.1 Etat de l'art

L'apprentissage automatique (un sous-ensemble de l'intelligence artificielle) fait référence à des méthodes qui ont la capacité "d'apprendre" de l'expérience, permettant l'exécution des tâches désignées. Des exemples d'apprentissage automatique sont les tâches comme la détection des motifs, la reconnaissance, le diagnostic, l'optimisation et la prédiction. L'apprentissage automatique peut aussi souvent être utilisé dans différents domaines de recherche sur les systèmes complexes impliquant l'identification de la structure du système (par exemple, les nœuds et les liens d'un réseau) et l'étude du comportement dynamique des systèmes non-linéaires (Par exemple, déterminer les exposants de LYAPUNOV, la prédiction de l'évolution future et l'inférence de la causalité d'interactions [214, 215]). Dans le domaine de l'application aux problèmes physiques complexes, l'apprentissage automatique a été utilisé avec succès pour construire et récupérer les structures et les dynamiques complexes du climat, des systèmes de régulation génétique, les systèmes chaotiques spatio-temporels et les réseaux de neurones. D'autre part, les systèmes complexes se produisent dans une grande variété de contextes pratiques, y compris l'ingénierie, les neurosciences, les réseaux sociaux, l'économie, etc...

Depuis, les systèmes complexes et l'apprentissage automatique ont eu une relation étroite entre eux. Par conséquent, explorer le fonctionnement de l'apprentissage automatique pour les problèmes impliquant des systèmes complexes a fait l'objet d'importantes recherches. Avec l'avènement de l'apprentissage automatique, il est devenu possible de développer de nouveaux algorithmes et stratégies d'identification, de contrôle et d'analyse des données des systèmes complexes, favorisant ainsi l'application de l'apprentissage automatique dans de nombreux domaines.

Concrètement, l'apprentissage automatique représente un atout considérable pour :

- **La modélisation des systèmes complexes** : La reconstruction de la structure

du système à partir des données est un problème basique en science des réseaux complexes et ses diverses applications, parce que certaines variables ainsi que les liens possibles entre les nœuds sont souvent indisponibles ou même inconnus [216-223].

- **La prédiction et manipulation des systèmes complexes** : Les algorithmes d'apprentissage automatique ont fourni une grande commodité en termes d'analyse et de prédiction de la dynamique des processus chaotiques et des systèmes complexes [224-228]. Notamment, la prédiction d'un système complexe est d'une grande importance dans la compréhension et la prise de décision.
- **La recherche en Data-Driven** : Récemment, le machine learning et le deep learning ont pris part à la compétition avec les méthodes caractéristiques traditionnelles dans les réseaux complexes, et de nombreuses méthodes basées sur le deep learning ont obtenu de meilleurs résultats en raison de l'abondance des données. Par ailleurs, la combinaison des méthodes traditionnelles et les techniques du deep learning, a reçu progressivement une grande attention et certains travaux issus de ce mariage en ont profité [229-232].
- **La classification et la détection** : La classification est définie comme le processus de reconnaissance, de compréhension et de regroupement d'objets et d'idées dans des catégories prédéfinies appelées "sous-populations". La classification et la détection constituent un problème préoccupant pour la communauté de la dynamique non-linéaire et des systèmes complexes. De nombreux travaux basés sur le machine et le deep learning, se sont intéressés à cette problématique, et ils ont réussi à garantir de belles performances [233-235].

4.2.2 Intelligence computationnelle et réseaux de neurones

Parmi les nombreuses approches utilisées dans l'apprentissage automatique, l'exploration et le traitement de données, les approches basées sur l'*intelligence computationnelle* sont peut-être les plus passionnantes et les plus inspirantes. Motivées par des processus observés dans le *cerveau* biologique, de telles approches ont été largement et avec succès appliquées à des domaines allant des prédictions boursières [236] aux interfaces cerveau-machine [237], de la reconnaissance vocale [238] à la robotique [239], à la détection d'anomalies [240] et à la recherche améliorée de documents [241].

En effet, Le cerveau est, de loin, l'organe le plus complexe du corps. Avec environ 85 milliards de neurones et 10000 milliards synapses par cm^3 (connexions entre neurones)

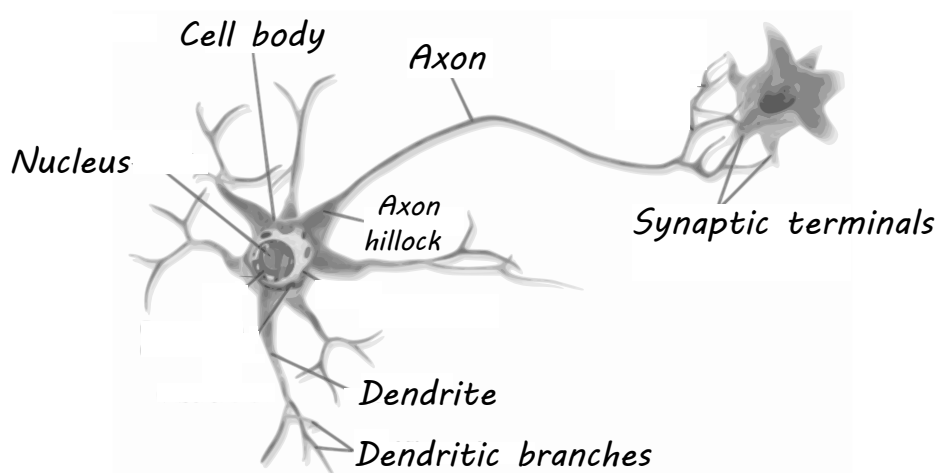


FIGURE 4.1 – Le schéma d'un neurone.

dans le cortex humain [242], il s'agit d'un traitement des données et d'un parallélisme à grande échelle qui permet de comprendre facilement pourquoi les fonctions du cerveau sont encore loin d'être entièrement comprises. Ce problème a peut-être été mieux décrit par LYALL WATSON [243] : "*Si le cerveau était si simple que nous pouvions le comprendre, nous serions si simples que nous ne le pourrions pas.*". Un neurone reçoit une entrée d'un autre neurone ou cellule (un signal provenant d'un nerf optique connecté à l'œil par exemple) le long d'une connexion pré-synaptique connue sous le nom de dendrites et envoie un signal (ou sortie) à d'autres neurones le long d'une projection post-synaptique connue sous le nom d'axone. La différence entre les deux connexions est leur forme et leur distribution ; les dendrites ont de nombreuses branches (le nom vient de leur ressemblance avec les arbres) et sont de forme irrégulière, alors qu'un axone est souvent beaucoup plus long, a moins de branches et est de forme plus lisse [244]. En effet, les axones jouent un rôle majeur dans le transfert d'informations sur de plus grandes distances dans le système nerveux [245]. Les autres parties principales du neurone comprennent le soma, qui est le corps cellulaire de forme grossièrement sphérique du neurone. La figure 4.1 montre la structure d'un neurone.

Une fois qu'un neurone a reçu une entrée via ses connexions et que son potentiel de corps cellulaire atteint un certain seuil, il libère une sortie vers d'autres neurones le long de son axone. L'axone est connecté à d'autres arbres dendritiques d'autres neurones via des connexions synaptiques. Une connexion synaptique a deux côtés : présynaptique et post-synaptique, et une fente synaptique entre eux. Lorsqu'un signal est envoyé d'un neurone à un autre, le signal électrique envoyé par le neurone déclencheur est converti en un signal chimique (appelé neurotransmetteurs) qui traverse la fente synaptique. Une fois du côté post-synaptique de la connexion, le signal chimique est reconverti en un signal électrique.

L'effet des neurotransmetteurs reçus peut être soit d'exciter (augmenter le potentiel de membrane cellulaire) soit d'inhiber (diminuer le potentiel de membrane cellulaire) le neurone récepteur. Un aperçu détaillé des neurones et des différentes façons de les modéliser ainsi que le cerveau sort du cadre de cette recherche. Un aperçu complet peut être trouvé dans BEARET [245] et GERNSTER et KISTLER [246].

Les réseaux de neurones artificiels (Artificial Neural Networks ANNs) (ici "artificiel" est largement utilisé pour distinguer les modèles de réseaux de neurones computationnels de leurs équivalents biologiques qui ont été brièvement présentés juste au-dessus) sont devenus un domaine largement étudié et utilisé dans le domaine de l'**intelligence computationnelle**. S'inspirant directement de la fonctionnalité des neurones biologiques du cerveau, leur attrait réside non seulement dans leurs fondements biologiques, mais également dans leur capacité à apprendre et à reconnaître les caractéristiques contenues dans les données auxquelles ils sont exposés. Grâce à cette expérience d'apprentissage, les ANN sont capables de reconnaître et de classer, tout comme le cerveau, de nouveaux exemples de données qu'ils n'ont jamais vus auparavant. Cela les rend très attrayants pour les domaines où de nouvelles données sont très susceptibles d'apparaître et qui nécessitent une réponse rapide. Alors que le cerveau humain est capable d'apprendre en ajustant ses connexions synaptiques entre ses neurones, les ANN adoptent une approche similaire en modifiant les connexions pondérées entre chacun de leurs neurones pendant l'entraînement jusqu'à ce qu'un niveau approprié d'erreur de sortie soit atteint.

Les ANN présentent de nombreuses caractéristiques qui les rendent attrayants pour la résolution de problèmes dans une grande variété d'applications. L'une de leurs caractéristiques les plus utiles est leur capacité à effectuer des transformations non-linéaires sur leurs entrées : ce qui leur permet de traiter des signaux non-linéaires que l'on trouve dans de nombreux domaines du monde réel tels que le traitement de la parole et du son. Les ANN ont également la capacité d'apprendre la cartographie entre les données d'entrée et la sortie cible sans avoir besoin d'y programmer des règles explicites. Cela leur permet de classer correctement les données d'entrée à l'aide d'un algorithme d'apprentissage qui ajuste les poids des connexions entre chacun des neurones.

Comme aucune règle explicite n'est intégrée dans les ANN considérés ici, une fois qu'un réseau est formé, il est difficile d'analyser quelles parties du réseau sont responsables d'un comportement particulier. Chaque connexion pondérée entraînée n'a souvent aucun sens en soi : c'est seulement collectivement que ces connexions effectuent une tâche. Les ANN sont souvent appelés approches de *boîte noire*, en raison de la difficulté d'évaluer la contribution d'un neurone particulier au comportement du réseau.

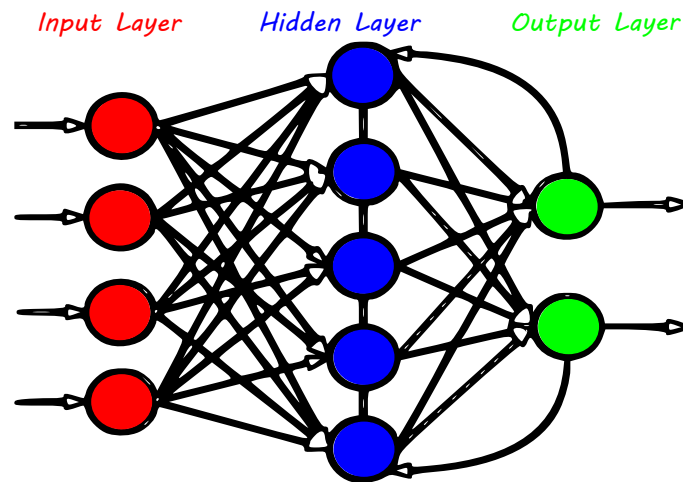


FIGURE 4.2 – Un RNN à trois couches avec des connexions récurrentes entre les neurones de la couche cachée et les neurones de sortie qui rétroagissent dans la couche cachée.

Les architectures des ANN sont généralement classées en réseaux *feedforward* [247] et en réseaux **récurrents** [248], dont le choix dépend du type de tâche de calcul. Les réseaux de neurones **feedforward** (FNN) sont principalement utilisés pour le traitement de données statiques (non temporelles), car les données d'entrée individuelles sont traitées indépendamment même si elles sont données de manière séquentielle. D'autre part, les réseaux de neurones *récurrents* (RNN) sont adaptés au traitement de données dynamiques (temporelles), car ils peuvent intégrer la dépendance temporelle des entrées dans leur comportement dynamique. En d'autres termes, les RNN sont capables de représenter des systèmes dynamiques pilotés par des entrées séquentielles en raison de leurs connexions de rétroaction. La figure 4.2 montre une vue d'ensemble schématique d'une architecture RNN.

Les RNN, comme leur nom l'indique, contiennent des connexions récurrentes qui sont soit intra-couche (les neurones d'une couche rétroagissent vers d'autres neurones de la même couche, ou eux-mêmes), soit inter-couche (la sortie rétroaction des neurones dans la couche cachée, par exemple). Ces connexions rétroagissent sur les activations de leurs neurones qui ont un effet sur la sortie du réseau à des pas de temps futurs [249]. Cela les rend bien adaptés aux domaines contenant des données de séries chronologiques où l'entrée actuelle a été influencée par les entrées précédentes, ce qui rend les RNN bien adaptés aux domaines tels que la robotique [250], la reconnaissance de l'écriture manuscrite [251], les prévisions météorologiques [252], la prévision de charge d'alimentation [253] et la reconnaissance vocale [254]. Les RNN se sont également révélés être des approximateurs universels des systèmes dynamiques, où un réseau avec N sorties peut approximativement réaliser un

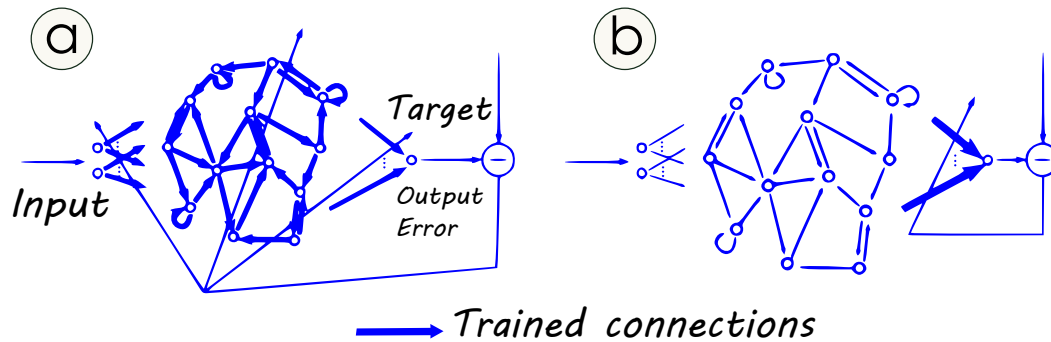


FIGURE 4.3 – (a). Les méthodes d’entraînement RNN traditionnelles basées sur la descente de gradient adaptent tous les poids de connexion (flèches en gras), y compris les poids d’entrée à RNN, RNN interne et RNN à sortie. (b). Dans le Reservoir Computing, seules les pondérations RNN-sortie sont adaptées.

système dynamique de N dimensions [255].

Malgré ces caractéristiques potentiellement puissantes, comme le soutiennent LUKOŠEVIČIUS et JAEGER, l’utilisation, et donc l’impact des RNN a été restreint depuis leur création initiale [256]. La raison en est que la plupart des RNN sont entraînés à l’aide d’algorithmes de descente de gradient qui souffrent de plusieurs problèmes tels que la non-convergence, les longs temps d’entraînement et l’annulation des gradients. Le problème d’annulation des gradients se produit lorsque les valeurs des gradients deviennent souvent très petites après seulement quelques itérations d’entraînement. Cela rend l’entraînement difficile et aboutit à un réseau qui ne peut pas bien gérer les entrées bruyantes ou qui est incapable d’apprendre une corrélation quelconque entre la sortie cible et les entrées récurrentes. Dans les cas extrêmes, cela rend l’entraînement presque impossible [244]. Pour surmonter les problèmes associés aux algorithmes complexes d’apprentissage lent des réseaux de neurones récurrents traditionnels, une nouvelle approche, où seuls les poids de sortie sont entraînés à l’aide d’un algorithme de régression linéaire simple, a récemment été proposée, nommée *Reservoir Computing* (RC) (Fig. 4.3b). En utilisant cette approche, il a été démontré que des performances similaires, et souvent améliorées, peuvent être obtenues tout en évitant la complexité des approches des RNN traditionnelles. Comme le *Reservoir Computing* est au centre de ce travail, nous allons le décrire en détails dans la section suivante.

4.3 Reservoir computing

Le Reservoir Computing (RC) est apparu il y a vingt ans comme un nouveau domaine prometteur dans le domaine des réseaux de neurones artificiels récurrents (RNN). L'idée derrière le RC a été (re)découverte plusieurs fois. Cela peut être dû à la simplicité conceptuelle des fonctions de calcul via une expansion temporelle de grande dimension de l'entrée. Le terme *Reservoir Computing* remonte à [257], dans lequel les auteurs ont expérimentalement unifié les **ESN** (Echo State Network) et les **LSM** (Liquid State Machine) dans un cadre général de calcul utilisant un **réservoir**. En effet, il y a vingt ans, une tendance alternative de compréhension, d'entraînement et d'utilisation des RNN a été proposée avec les réseaux d'état d'écho (ESN) [258] en Machine Learning et les machines à état liquide (LSM) [19] en neuroscience computationnelle.

La recherche concernant le RC s'est développée grâce aux collaborations entre les premiers pionniers. Cela a abouti à de puissantes démonstrations pratiques et au développement d'une théorie mathématique pour ancrer le domaine. L'une des premières motivations pour utiliser le RC comme mécanisme d'entraînement efficace pour les RNN a été largement remplacée par les méthodes du deep learning, qui ont maîtrisé les difficultés de l'entraînement par descente de gradient que le RC contourne. Néanmoins, la recherche sur le RC est florissante [259] dans une variété de domaines, ce qui reflète les nombreuses perspectives qui peuvent être prises dans l'étude du RC :

- En Machine learning, le RC présente un algorithme d'apprentissage simple à mettre en œuvre et efficace pour les cartographies entrées-sorties temporelles par une expansion temporelle non linéaire (voir section 2.4).
- Dans la recherche sur les ANN, le RC présente un mécanisme d'entraînement efficace. Les ESN [258] ont été introduits à l'origine dans le contexte de l'apprentissage automatique avec les ANN.
- Dans la théorie des systèmes dynamiques, le RC aide à approximer des systèmes dynamiques potentiellement complexes. Dans le domaine des systèmes dynamiques non linéaires pilotés par les entrées, les systèmes réservoirs eux-mêmes constituent un sujet d'investigation intéressant [260, 261].
- En neurosciences computationnelles, le RC propose un modèle simple de calcul cortical. Les LSM [19] ont été introduits dans ce contexte.
- Pour les scientifiques des matériaux et des systèmes non linéaires ou les ingénieurs des dispositifs computationnels, le RC fournit une méthode simple pour exploiter une dynamique complexe pour le calcul ou, simplement, un moyen de programmer

un système physique. Ceci est souvent appelé le RC physique [262, 263] qui a des racines dans le RC ainsi que dans l'informatique non conventionnelle [264].

4.3.1 Le principe général

Compte tenu de cette vision large du calcul, le principe général du RC est intuitif. Il permet la construction d'un système dynamique qui implémente une cartographie souhaitée des signaux d'entrée aux signaux de sortie. Cela inclut la prédiction de séries chronologiques, la classification et d'autres tâches. À cette fin, un ensemble d'exemples (ensemble de données d'apprentissage) qui démontre la relation souhaitée entre les signaux d'entrée et les signaux de sortie est généralement requis. Dans la terminologie du Machine learning, il s'agit d'une tâche d'apprentissage supervisé temporelle.

Dans cette configuration, un signal d'entrée d'entraînement $u_{\text{train}}(t) \in \mathbb{R}^{N_u}$ avec un signal de sortie cible souhaité $y_{\text{train}}(t) \in \mathbb{R}^{N_y}$ est donné, où N_u, N_y est la dimensionnalité de l'entrée et de la sortie, respectivement. Dans la prédiction de séries temporelles, le signal d'entrée peut montrer une série temporelle et le signal de sortie cible sa continuation. Dans la classification, le signal d'entrée peut présenter un signal de parole auditif, puis cibler le signal de sortie des phonèmes correspondants qui sont prononcés. Notez que si plusieurs de ces exemples d'entrées-sorties existent, ils peuvent simplement être concaténés dans le temps. Ce que l'on souhaite, c'est un système qui, lorsqu'il est alimenté en entrée $u_{\text{train}}(t)$, génère un signal de sortie $\hat{y}_{\text{train}}(t) \approx y_{\text{train}}(t)$.

On peut supposer que les données sont générées par une fonction $\psi(u_{\text{train}})(t) = y_{\text{train}}(t) + e$, où ' e ' est un terme d'erreur additif, par exemple, le bruit, et u_{train} est la série chronologique d'entrée entière. En traitement du signal, un tel système est appelé filtre. Le système du réservoir, noté $\hat{\psi}_\theta$ (où θ est le vecteur des paramètres du modèle déterminant le système réservoir), vise alors à se rapprocher le plus possible du véritable système générateur ψ .

La solution du RC à cette tâche d'apprentissage se fait en trois étapes :

1. Le réservoir est préparé comme un système dynamique non-linéaire de grande dimension qui peut être piloté par le signal d'entrée $u_{\text{train}}(t)$. Pour un système performant, la dynamique collective du réservoir doit satisfaire les conditions résumées dans les sections 4.3.2 et 4.3.2. Cependant, les composants individuels du réservoir peuvent être aléatoires. Les états de grande dimension du réservoir sont donnés par $\mathbf{X}(t) \in \mathbb{R}^N$, et de nombreuses variables d'état doivent être accessibles pour être lues comme $\mathbf{x}(t) \in \mathbb{R}^{N_x}$, qui consiste en $x_i(t)$ pour $i = 1, \dots, N_x$. Tous les états du réservoir n'ont pas besoin d'être observables, il est possible que $N_x < N$.

2. Le réservoir est piloté avec l'entrée $u_{\text{train}}(t)$ et les variables d'état correspondantes du réservoir en réponse à cette entrée sont lues et stockées en tant que $x_i^{\text{train}}(t)$.
3. Enfin, une fonction de lecture (readout) F , qui mappe tous les vecteurs d'état enregistrés $x_i^{\text{train}}(t)$ vers une sortie $\hat{y}(t) \approx y(t)$ est apprise (estimée) :

$$F(x_1^{\text{train}}(t), \dots, x_N^{\text{train}}(t)) = \hat{y}(t) \approx y(t). \quad (4.1)$$

Ceci est fait en minimisant une fonction d'erreur choisie $E(y, \hat{y})$, qui évalue la différence entre la sortie du réservoir $\hat{y} = \hat{\psi}_\theta(u)$ et la sortie désirée $y = \psi(u)$. Plus précisément, nous minimisons l'erreur moyenne sur les données d'entraînement données :

$$\bar{E}_{\text{train}}(\hat{\psi}_\theta) = 1/N_T \sum_t E(y_{\text{train}}(t), \hat{\psi}_\theta(u_{\text{train}}(t))). \quad (4.2)$$

où N_T est le nombre de pas de temps. La fonction du readout F , qui minimise l'erreur moyenne, est typiquement obtenue avec une simple *régression linéaire*.

Après cette procédure, le RC peut être utilisé pour l'inférence afin de faire des prédictions. Lors de l'alimentation d'un nouveau signal d'entrée $u(t)$ au réservoir, les vecteurs d'état du réservoir $x_i(t)$ sont observés et utilisés pour calculer le signal de sortie prédit $\hat{y} = F(x_1(t), \dots, x_N(t))$.

Il est important de rappeler que ce principe général est valable pour les deux types d'implémentations que nous avons évoquées avant (ESN et LSM). Néanmoins, la principale différence entre les LSM et les ESN provient de leurs origines et concerne les propriétés de leurs neurones et leurs fonctions d'activation. Les LSM ont été créés dans une perspective de neurosciences et, par conséquent, sont plus motivés biologiquement. En d'autres termes, leur liquide (ou réservoir) contient plus de neurones biologiquement plausibles que les ESN. Ils sont généralement mis en œuvre à l'aide de neurones à pointes (Spiking) pour créer des microcircuits neuronaux. Les LSM ont cependant été quelque peu restreints en ce qui concerne les applications pratiques en raison de trop nombreuses conditions requises lors de la construction du réservoir [257], ce qui rend difficile l'obtention d'un réservoir avec les caractéristiques souhaitées, alors que les ESN ont une plage beaucoup plus large d'hyperparamètres sur laquelle ils peuvent fonctionner [265].

Les ESN ont été développés pour le traitement du signal et les applications de contrôle [266] et sont au centre de ce travail. Dans le reste de cette thèse, les termes RC et ESN seront utilisés de manière interchangeable, sauf indication contraire.

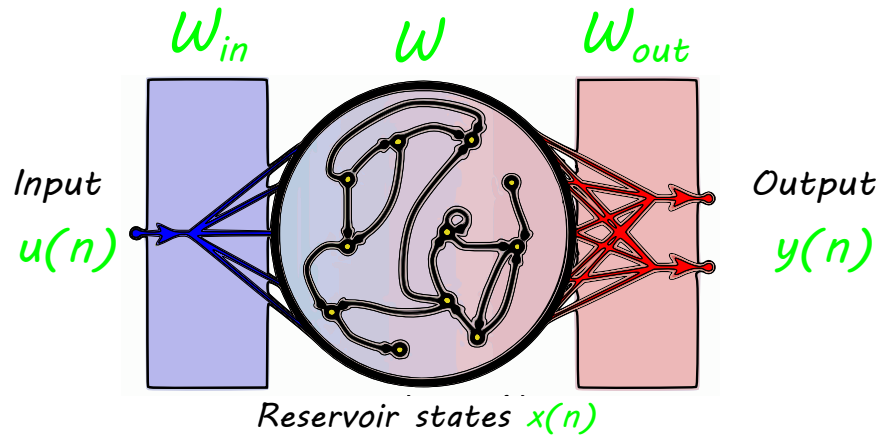


FIGURE 4.4 – L’anatomie d’un ESN est composée de trois parties, une couche d’entrée \mathbf{W}_{in} , une couche réservoir \mathbf{W} et une couche de sortie \mathbf{W}_{out} .

4.3.2 Echo State Networks

Développés par JAEGER [20, 258], l’architecture typique d’un ESN se compose de trois couches : une couche d’entrée, une couche de réservoir et une couche de sortie. Chaque neurone de la couche d’entrée a des connexions pondérées avec chaque neurone de la couche réservoir (Fig. 4.4) (Des connexions directes optionnelles à la couche de sortie sont également possibles.). A l’intérieur de la couche réservoir, des connexions récurrentes entre chaque neurone sont présentes. Chaque neurone du réservoir a également des connexions pondérées à chaque neurone de la couche de sortie. Les neurones de sortie peuvent avoir des connexions de rétroaction récurrentes facultatives vers les neurones du réservoir et des connexions récurrentes vers eux-mêmes et les autres neurones de sortie.

La figure 4.4 montre une configuration ESN typique. Notez que seules les connexions des premiers nœuds d’entrée, de réservoir et de sortie sont illustrées par souci de simplicité. Dans le travail présenté ici, aucune connexion entrée-sortie, sortie-réservoir ou sortie-sortie n’est utilisée.

Le modèle de base

Comme expliqué précédemment, les ESN sont appliqués à des tâches temporelles supervisées où, pour un signal d’entrée d’entraînement donné $\mathbf{u}(n) \in \mathbb{R}^{N_u}$ un signal de sortie cible souhaité $\mathbf{y}_{target}(n) \in \mathbb{R}^{N_y}$ est connu.

Ici $n = 1, \dots, T$ est le temps discret et T est le nombre de points de données dans l’ensemble de données d’apprentissage. En pratique, le jeu de données peut être constitué

de plusieurs séquences de longueurs variables, mais cela ne change pas les principes. La tâche consiste à apprendre un modèle avec la sortie $\mathbf{y}(n) \in \mathbb{R}^{N_y}$, où $\mathbf{y}(n)$ correspond aussi bien que possible à $\mathbf{y}_{target}(n)$, en minimisant une mesure d'erreur $E(\mathbf{y}, \mathbf{y}_{target})$, et plus important encore, qui se généralise bien à des données invisibles. La mesure d'erreur E est généralement une erreur quadratique moyenne (MSE), par exemple la (RMSE).

$$E(\mathbf{y}, \mathbf{y}_{target}) = 1/N_y \sum_{i=1}^{N_y} \sqrt{1/T \sum_{n=1}^T (y_i(n) - y_i^{target}(n))^2} \quad (4.3)$$

qui est également moyennée sur les N_y dimensions de la sortie ici.

Dans ce travail, nous allons utiliser une implémentation assez spéciale des ESNs avec des neurones à fuite (Leaky integrator neurons LIN) proposée par [267]. Aujourd'hui, un certain nombre de versions de LIN sont souvent utilisées dans les ESN. La dynamique de ces neurones en temps discret, est donnée par l'équation de mise à jour typique suivante :

$$\mathbf{x}(n+1) = (1 - \alpha)\mathbf{x}(n) + \alpha\tilde{\mathbf{x}}(n+1) \quad (4.4)$$

$$\tilde{\mathbf{x}}(n+1) = \tanh(\mathbf{W}_{in}[bias; \mathbf{u}(n+1)] + \mathbf{W}\mathbf{x}(n) + \mathbf{W}_{fb}\mathbf{y}(n) + \nu), \quad (4.5)$$

où $\mathbf{x}(n) \in \mathbb{R}^{N_x}$ est un vecteur d'activations de neurones du réservoir et $\tilde{\mathbf{x}}(n) \in \mathbb{R}^{N_x}$ est sa mise à jour, le tout au pas de temps n , $\tanh(\cdot)$ est appliqué élément par élément, $[\cdot; \cdot]$ représente une concaténation verticale de vecteurs (ou matrices), $\mathbf{W}_{in} \in \mathbb{R}^{N_x \times (1+N_u)}$ et $\mathbf{W} \in \mathbb{R}^{N_x \times N_x}$ sont respectivement les matrices de poids d'entrée et de récurrence, et $\alpha \in (0, 1]$ est le taux de fuite (Leaking Rate). Pour certaines tâches, nous incluons également la matrice de connexion du feedback $\mathbf{W}_{fb} \in \mathbb{R}^{N_x \times N_y}$, mais pas toujours. D'autres sigmoïdes peuvent être utilisés que la "tanh" comme fonction d'activation, qui est cependant le choix le plus courant. *bias* est un terme constant, généralement l'unité. Le terme de biais constant sert à pousser la fonction d'activation dans les régimes de saturation non-linéaires [214]. Un petit terme de bruit constant ν , généralement dans la plage de $\nu = 0,00001 \rightarrow 0,01$, est souvent ajouté à l'entrée pendant la phase d'entraînement. Le terme de bruit aide à "immuniser" le réservoir contre les changements inattendus de l'entrée $\mathbf{u}(n)$, puisqu'il permet d'éviter le surajustement à un ensemble de données spécifique.

De manière similaire à la section.4.3.1, les étapes majeures pour utiliser un ESN sont :

- **Création** : Générer un grand réservoir aléatoire RNN ($\mathbf{W}_{in}, \mathbf{W}, \alpha$). Les paramètres qui contrôlent la dynamique du réservoir doivent également être spécifiés (voir la section 4.3.2 et 4.3.2 ci-dessous pour plus de détails).

- **Training** : Exécuter à l'aide de l'entrée $\mathbf{u}(n)$ l'apprentissage et collecter les états d'activation de réservoir correspondants $\mathbf{x}(n)$ dans une matrice \mathbf{M} . Les données de la sortie $\mathbf{y}(n)$ utilisées pendant le training sont collectées dans une matrice \mathbf{T} . Calculer les poids du readout linéaire \mathbf{W}_{out} du réservoir à l'aide d'une régression linéaire, minimiser le RMSE entre $\mathbf{y}(n)$ et $\mathbf{y}_{target}(n)$;

$$\mathbf{y}(n) = \mathbf{W}_{out}[1; \mathbf{u}(n); \mathbf{x}(n)], \quad (4.6)$$

La méthode de régression linéaire utilisée dans notre travail est la méthode de régression de *ridge* (également connue sous le nom de régularisation de TIKHONOV) pour calculer les poids de sortie du réservoir \mathbf{W}_{out} et qui est donnée par [258] :

$$\mathbf{W}_{out} = \mathbf{T}^T \mathbf{M} (\mathbf{M}^T \mathbf{M} + \beta \mathbb{I})^{-1} \quad (4.7)$$

où β est une constante, \mathbb{I} est la matrice d'identité. La régression *ridge* est considérée comme la méthode préférée pour calculer les poids de la matrice de sortie du réservoir \mathbf{W}_{out} , pour plusieurs raisons. Premièrement, elle est plus efficace dans le calcul que la méthode de *pseudo-inverse*. L'autre avantage, et sans doute le plus important, de la méthode d'apprentissage par régression *ridge* est la régularisation β . Ce terme de décalage constant dans la matrice inversée sert à diminuer la taille des entrées de \mathbf{W}_{out} . Cela rend le réseau moins sensible au bruit et réduit tout problème de surajustement [256, 258].

- **Testing** : Utiliser le réseau *entraîné* sur de nouvelles données d'entrée $\mathbf{u}(n)$ dans le calcul de $\mathbf{y}(n)$ et en employant les poids *entraînés* de sortie \mathbf{W}_{out} .

Les hyperparamètres du réservoir et de sa dynamique

Étant donné le modèle d'ESN (section. 4.3.2), le réservoir est défini par le tuple $(\mathbf{W}_{in}, \mathbf{W}, \alpha)$. Les matrices d'entrée \mathbf{W}_{in} et celle de connexion récurrente \mathbf{W} , sont générées aléatoirement selon certains paramètres qui seront discutés dans les sections suivantes, et le taux de fuite α est sélectionné comme un paramètre libre qui influence la dynamique du réservoir.

Par analogie avec d'autres approches du Machine learning, et en particulier les réseaux de neurones, ce que nous appelons ici des *paramètres* pourraient également être appelés *méta-paramètres* ou **hyper-paramètres**, car ce ne sont pas des poids de connexion concrets mais des paramètres régissant leurs distributions. Nous les appellerons **paramètres globaux** pour mieux refléter leur nature, ou simplement *paramètres* par souci de brièveté.

Les paramètres globaux définissant le réservoir sont : la *taille* N_x , la *sparsité*, la *distribution des éléments non nuls* et le *rayon spectral* ρ de \mathbf{W} , le *input scaling* de \mathbf{W}_{in} , le *taux de fuite* α , Le *bruit* et le *biais*, la *fonction d'activation* et la *régularisation*. Nous allons maintenant procéder dans cet ordre pour donner plus de détails sur chacun de ces choix de conception, et des intuitions sur la manière de les réaliser.

- **La taille du réservoir** Un paramètre évidemment crucial du modèle Eq. (4.4) est N_x , le nombre d'unités dans le réservoir.

En général, plus le réservoir est grand, meilleures sont les performances pouvant être obtenues, à condition que des mesures de régularisation appropriées soient prises contre le surajustement. Étant donné que l'entraînement et l'exécution d'un ESN sont peu coûteuses en termes de calcul par rapport à d'autres approches RNN, des tailles de réservoir d'ordre 10^4 ne sont pas rares [268].

Plus l'espace des signaux du réservoir $\mathbf{x}(n)$ est grand, plus il est facile de trouver une combinaison linéaire des signaux pour approximer la cible $\mathbf{y}_{target}(n)$. En effet, N_x correspond au nombre effectif de paramètres du modèle, car il y a exactement un poids de sortie entraînable associé à chaque neurone réservoir. Par conséquent, un N_x plus élevé augmente la capacité du réseau à modéliser des systèmes plus complexes. Les tâches difficiles ont tendance à nécessiter des réservoirs plus grands. D'autre part, N_x contraint la *capacité de mémoire* maximale du réseau (alors que les effets de la mémoire dynamique sont régis par le *rayon spectral* comme nous le verrons plus tard). Cela pourrait indiquer que des réservoirs plus grands sont toujours bénéfiques. L'ajout de neurones supplémentaires connectés au hasard enrichit le seau de transformations non-linéaires des signaux d'entrée, qui peuvent être utilisés pour construire le signal de sortie.

- **La sparsité du réservoir** Un nombre croissant de travaux qui analysent le rôle de la sparsité dans la conception de systèmes de réseaux de neurones (dynamiques) [269, 270]. La caractérisation qui ressort de ces études est que le fait d'avoir des connexions clairsemées entre les neurones n'est pas seulement avantageux en terme de calcul - car il permet des calculs rapides de matrice creuse - mais peut également être bénéfique pour obtenir une meilleure performance dans la pratique. De plus, dans le contexte des systèmes de traitement de l'information d'inspiration neurobiologique, un faible degré de connectivité entre les neurones a été montré capable d'améliorer la qualité des représentations internes développées [271].

Dans le cadre du Reservoir computing, une stratégie typique dans sa conception consiste à configurer la couche récurrente de manière clairsemée. L'intuition initiale

était que la rareté dans la couche récurrente \mathbf{W} non entraînée, pourrait permettre un découplage des variables d'état et donc permettre d'avoir des représentations plus riches [258].

Récemment dans [272], les résultats indiquent que sous les configurations de réservoir, le nombre de connexions non nulles peut jouer un rôle déterminant au niveau des performances. En particulier, alors qu'un nombre modeste de connexions récurrentes est déjà suffisant pour atteindre de bonnes performances, ils ont constaté que l'entrée la plus parcimonieuse pour les connexions de réservoir conduisent aux meilleurs résultats en termes de mémoire à court terme et en termes de dimension effective de la matrice collectant les états.

- **La topologie du réservoir** De nombreux travaux ont étudié la distribution optimale des éléments non nuls de la couche récurrente du réservoir \mathbf{W} . En effet, la topologie utile pour améliorer significativement les performances de l'ESN, est encore méconnue. L'approche classique proposée par le pionnier du domaine des ESNs, consiste à produire des connexions reliées aléatoirement (Random) [258] (Fig. 4.5). Un réseau ERDŐS-RÉNYI est utilisé comme matrice d'adjacence du réservoir. Ensuite, les chercheurs ont exploré une variété de topologies en adoptant la théorie des réseaux complexes.

Comme son nom l'indique, les réseaux complexes font référence à un graphe constitué d'un ensemble de sommets et d'un ensemble d'arêtes, où l'organisation et la topologie ne sont pas triviales. Ces graphes peuvent être utilisés pour représenter et étudier plusieurs organisations, telles que l'internet, les relations sociales entre les individus, les réseaux de neurones biologiques, les chaînes métaboliques, les chaînes alimentaires [273-275].

Les principales propriétés de ces réseaux sont [276] :

- La complexité structurelle (difficile à visualiser le réseau).
- L'évolution : les changements constants dans la structure du réseau en raison de l'inclusion et de la suppression des sommets et des connexions.
- La diversité des connexions.
- Dynamique complexe : au-delà de la structure, ce qui affecte à grande échelle l'état du réseau est sa dynamique, qui est le flux d'informations, défauts de communication, synchronisation et corrélation entre les sommets et autres.

La façon dont les arêtes sont corrélées définit le type de réseau. Il existe trois grandes catégories de réseaux complexes, qui sont :

- **Réseaux Erdős-Rényi (Aléatoires)** : C'était le premier type de réseau com-

plexe mathématiquement étudié. Proposé dans [277], ce modèle considère que les arêtes sont ajoutées aléatoirement étant donné un nombre fixe de sommets N . Le degré moyen de chaque sommet est donné par l'équation :

$$\langle k \rangle = p(N - 1) \quad (4.8)$$

où (p) est la probabilité qu'un sommet donné se connecte à n'importe quel autre sommet du réseau, suivant une distribution de Poisson. De cette manière, tous les sommets ont sensiblement le même nombre de connexions, ainsi que la même chance de recevoir de nouvelles connexions (Fig. 4.5).

- **Réseaux scale-free** : La principale caractéristique de ce type de réseau est que plus un nœud a de connexions, plus ce nœud est susceptible de recevoir de nouvelles connexions. Proposé par [278], les auteurs ont observé qu'une telle propriété est présente dans plusieurs réseaux du monde réel, tels qu'Internet, les réseaux métaboliques, les réseaux de neurones biologiques, les réseaux de citation d'articles, etc. Par conséquent, les auteurs ont montré que la distribution des degrés de plusieurs de ces réseaux suit une loi de puissance, qui est donnée par l'équation suivante :

$$P(k) \sim k^{-\lambda} \quad (4.9)$$

où k est le nombre de connexions d'un sommet donné et λ est l'exposant de l'échelle. Ce type de réseau (Fig. 4.5) aura la majorité des sommets avec un faible nombre de connexions, et seulement quelques-uns avec un nombre élevé de connexions (ces sommets sont appelés *hubs*).

- **Réseaux small-world** : Dans [279], les auteurs ont proposé un modèle où les sommets créent des connexions avec les sommets les plus proches avec une probabilité plus élevée, comme observé dans les réseaux sociaux, par exemple. La majorité des sommets créent des connexions avec d'autres sommets tout au long du chemin le plus court (le nombre minimum d'arêtes entre deux nœuds). Dans le même temps, le coefficient de clustering est relativement élevé, ce qui signifie que les nœuds voisins ont tendance à être connectés (Fig. 4.5).
- **Réseaux en cluster** : Notez que les réseaux complexes peuvent avoir une topologie en cluster (Fig. 4.6), où des nœuds similaires sont connectés et il y a moins de connexions entre les nœuds de différents clusters, comme on peut l'observer sur le World Wide Web, dans les réseaux de citation d'articles, dans les réseaux routiers de trafic, les réseaux sociaux et autres [281]. Le clustering (également

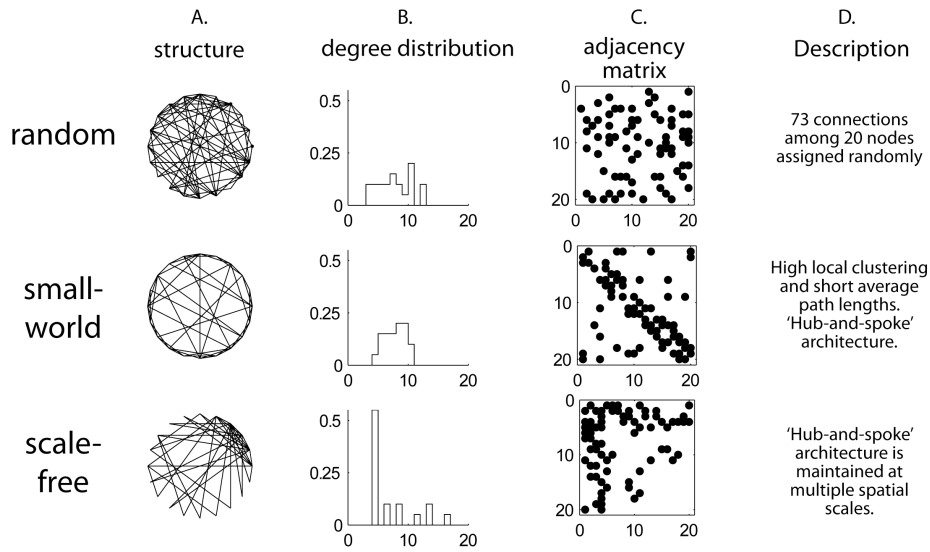


FIGURE 4.5 – Extrait de [280] : Des réseaux aléatoires, small-world et scale-free contenant 20 nœuds et 73 connexions ont été générés par des algorithmes informatiques. Le nombre de nœuds et de connexions a été choisi pour être petit afin de mettre en évidence les différences entre les graphiques. A. Une représentation structurelle de chaque graphique. B. Répartition des distributions. C. Matrices d'adjacence. D. Description des modèles de connectivité.

connu sous le nom de détection de communauté) est actuellement l'un des plus grands défis de l'apprentissage automatique, en raison de la grande complexité des algorithmes traditionnels lors de la gestion de grandes quantités de sommets et d'arêtes. Ce fait rend ce domaine de recherche pertinent et prometteur, comme l'indiquent les travaux [281-284], et autres.

Les spécialistes du domaine des ESNs, ont mené une multitude d'études en utilisant ces différentes topologies dans le but d'optimiser la dynamique collective interne des ESNs [215, 285-288].

- **Le rayon spectral** Le rayon spectral (ρ) est la valeur absolue de la valeur propre la plus élevée de la matrice des poids de réservoir qui est utilisée pour redimensionner les poids de \mathbf{W} , afin de satisfaire la propriété d'état d'écho. La propriété d'état d'écho permet essentiellement au réservoir de posséder une mémoire d'évanouissement, ou "échos", de ses entrées passées (la propriété d'état d'écho est discutée plus en détail dans la section 4.3.2 ci-dessous). Après l'initialisation des poids du réservoir, ces poids sont remis à l'échelle par le rayon spectral :

$$\mathbf{W}' = \frac{\rho * \mathbf{W}}{\lambda_{max}}. \quad (4.10)$$

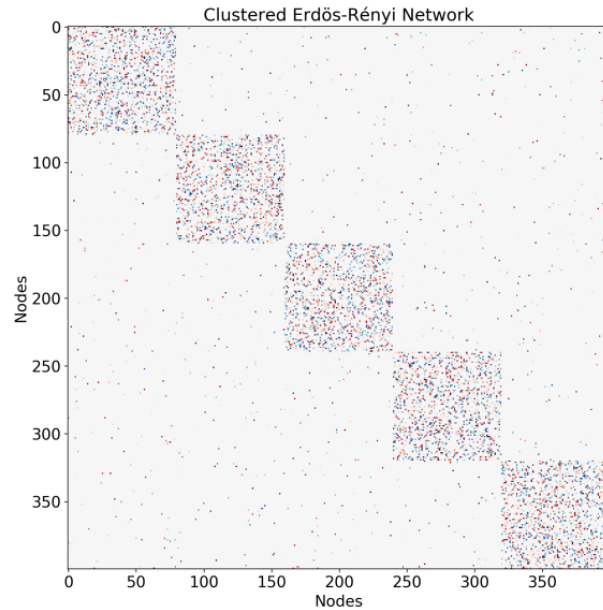


FIGURE 4.6 – Extrait de [289] : La matrice adjacente d’un réseau ERDÖS RENYI en cluster.

où λ_{max} est la valeur propre absolue maximale de la matrice de poids de réservoir *initialisée*.

La variation du rayon spectral affecte la capacité de mémoire qu’un réservoir contient : des valeurs bien inférieures à l’unité diminuent la capacité de mémoire du réservoir car les activations des neurones du réservoir deviennent très contractiles autour de leur origine 4.10. (a), ce qui signifie que toute activité d’activation s’éteint rapidement après une période sans entrée. Des valeurs plus grandes proches de l’unité augmentent la capacité de mémoire car les neurones restent actifs plus longtemps après aucune entrée [290]. Les valeurs supérieures à l’unité créent un réservoir très non-linéaire qui à son tour a une faible capacité de mémoire conduisant à la saturation des activations des neurones du réservoir, alors qu’une petite valeur de rayon spectral diminue la capacité de mémoire du réservoir car ses activations s’éteignent rapidement après aucune entrée.

Idéalement, les valeurs autour de l’unité créent généralement un réservoir linéaire avec une capacité de mémoire élevée, où les neurones du réservoir sont répartis sur des parties non-linéaire et linéaires de leurs activations (Fig. 4.7).

Comme principe directeur, $\rho(\mathbf{W})$ devrait être plus grand pour les tâches où un historique plus étendu de l’entrée est nécessaire pour l’exécuter. Le rayon spectral détermine la vitesse à laquelle l’influence d’un apport s’éteint dans un réservoir avec

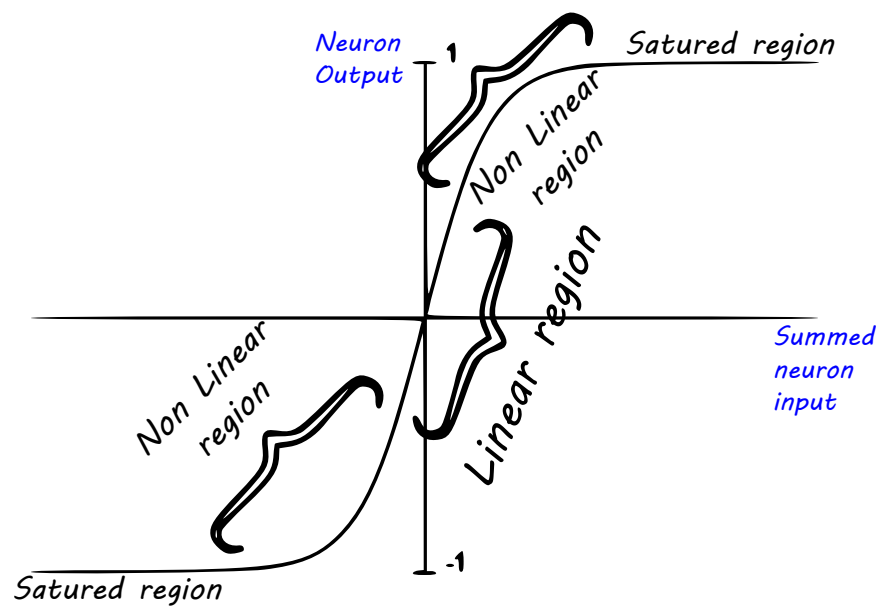


FIGURE 4.7 – La fonction d’activation de \tanh et les différentes régions où se trouve son point de fonctionnement. Le gain de la non-linéarité sigmoïde est le plus grand autour de l’origine. Une fois que le neurone est piloté par un signal externe ou un biais constant, le point de travail se déplace vers le haut ou vers le bas et le gain diminue, résultant en un réservoir moins dynamique.

le temps et assure la stabilité des activations du réservoir [291].

- **La mise à l’échelle des entrées** La mise à l’échelle de la matrice de poids d’entrée \mathbf{W}_{in} est un autre paramètre clé à optimiser dans un ESN. Pour \mathbf{W}_{in} uniformément distribué, nous définissons généralement la mise à l’échelle d’entrée comme la plage de l’intervalle $[-a; a]$ à partir duquel les valeurs de \mathbf{W}_{in} sont échantillonnées ; pour les poids d’entrée distribués normaux, on peut prendre l’écart-type comme mesure d’échelle. Afin d’avoir un petit nombre de paramètres librement réglables, souvent toutes les colonnes de \mathbf{W}_{in} sont mises à l’échelle ensemble en utilisant une seule valeur de mise à l’échelle. Il a été suggéré dans les travaux originaux de l’ESN [256, 258] de mettre à l’échelle et de déplacer les données d’entrée, en optimisant l’ampleur des deux. Mais le même effet peut être obtenu en mettant à l’échelle les poids d’entrée des entrées actives et le biais, respectivement.

Pratiquement, l’échelle d’entrée (*input scaling*) détermine la mesure dans laquelle les activations des neurones du réservoir sont pilotées par l’entrée, en d’autres termes la mesure dans laquelle l’entrée actuelle impacte l’état suivant du réservoir [290]. La valeur par défaut est généralement l’unité, mais la modification de la mise à l’échelle de l’entrée peut avoir un impact majeur sur les performances. En effet, une petite

échelle d'entrée crée généralement un réservoir linéaire (en supposant que les autres paramètres, tels que le rayon spectral, ont des valeurs non-extrêmes) car la plupart des points de travail des neurones du réservoir restent dans la plage linéaire de leurs fonctions d'activation.

D'autre part, des échelles d'entrée plus grandes déplacent le point de travail des neurones du réservoir vers les parties non-linéaires de leurs activations. Une très grande échelle d'entrée conduit à la saturation de l'ensemble du réservoir, où toutes les activations de ses neurones sont proches de -1 ou $+1$ et restent à ce point de saturation ou alternent entre les extrêmes pour la majorité de l'ensemble de données. Dans cette situation, le gain de la fonction d'activation est le plus faible, car il s'est déplacé vers le haut ou vers le bas le long de la fonction d'activation et les neurones du réservoir peuvent ne plus être capables de discriminer entre de petits changements dans les données d'entrée. Pour les données hautement non-linéaires, dont les valeurs changent radicalement d'un pas de temps à l'autre et qui nécessitent très peu de mémoire à court terme, une échelle d'entrée élevée vers un réservoir peut offrir les meilleures performances, car il est nécessaire de cartographier les données d'entrée d'une région de son espace d'états de grande dimension vers une autre région afin de classer correctement les données, ce qui est aidé par un espace d'états plus grand. Si un réservoir saturé peut offrir les meilleures performances, il peut également être instable du fait que ses activations neuronales se déplacent d'un extrême à l'autre.

- **Le taux de fuite** Le taux de fuite α détermine la mesure dans laquelle l'activation d'un neurone diminue sur une période de temps. Cela a été introduit à l'origine dans le but de modifier la dynamique du réservoir pour correspondre à l'échelle de temps (c'est-à-dire le taux de variation d'un signal) d'un ensemble de données, par exemple une onde sinusoïdale très lente. En utilisant des neurones de type (tanh) standard, ces données à évolution lente sont difficiles à modéliser car les neurones ne sont influencés qu'indirectement par leur valeur précédente et n'ont pas de mémoire réelle. Cette idée a été introduite à l'origine par Jaeger [267] dans le cadre de la fonction d'activation d'un neurone intégrateur à fuite (LIN) qui, grâce à l'utilisation d'un taux de décroissance de fuite, possédait une certaine mémoire interne. Il a été montré que faire correspondre la vitesse de la dynamique interne du réservoir avec l'échelle de temps de certains ensembles de données peut entraîner une amélioration des performances [292]. La diminution du taux de fuite augmente la capacité des neurones du réservoir à se souvenir des entrées présentées plus loin dans le passé, mais diminue leur capacité à se souvenir des entrées les plus récentes, alors qu'un taux

de fuite important (notez que le taux de fuite le plus élevé ne peut pas être supérieur à l'unité) a l'effet contraire où le réservoir est capable de se souvenir d'entrées plus récentes que celles présentées plus loin dans le passé [293]. Par conséquent, le taux de fuite optimal dépend des besoins en mémoire de la tâche à accomplir.

- **Le bruit et le biais** L'ajout de bruit à l'ensemble de données peut améliorer les performances pour certains domaines problématiques en régularisant le réservoir et en aidant à éviter le surajustement. L'ajout de bruit entraîne généralement une augmentation de l'erreur d'apprentissage, mais une diminution de l'erreur de test. Il a été rapporté par *Wyffels et al* [294] que l'insertion de bruit obtenait des performances similaires à la régression de *ridge* car elle régularisait le réservoir contre le surajustement, mais la nature aléatoire du bruit rendait difficile la reproduction fiable des résultats pour chaque architecture de réservoir.

D'autre part, l'utilisation d'un biais vers le réservoir peut aider à éloigner les activations des neurones du réservoir de leur origine lors de l'application du réseau à des tâches génératives (c'est-à-dire en exécution libre). Le biais peut également être utilisé pour stabiliser le réseau en utilisant une grande valeur pour déplacer les activations des neurones réservoirs vers la partie non linéaire de leur fonction d'activation, ce qui peut être utile avec des ensembles de données hautement non linéaires [258, 290]. Appliquer un biais au *readout* peut également améliorer les performances lorsque la sortie cible d'un ensemble de données a une moyenne qui n'est pas égale à zéro [258].

L'effet de la modification du biais sur la dynamique du réservoir a également été étudié par Verstraeten [291] qui ont montré que de grandes valeurs de biais peuvent augmenter la quantité de non-linéarité asymétrique qu'un neurone possède en appliquant un décalage par rapport à son origine.

- **Fonction d'activation** L'utilisation de différentes fonctions d'activation peut avoir un impact sur les performances d'un réservoir. La fonction d'activation utilisée a une grande influence sur la capacité de mémoire que peut posséder un neurone. Pour les neurones linéaires, ils possèdent une capacité mémoire maximale égale au nombre de neurones à l'intérieur du réservoir. Bien que les neurones linéaires aient une plus grande capacité de mémoire que les neurones avec des fonctions d'activation sigmoïde, ils ont du mal avec des données hautement séparables non-linéairement. SCHRAUWEN [295] ont étudié des réservoirs qui ont été adaptés à l'aide de la plasticité intrinsèque. Dans une expérience utilisant une tâche qui nécessite de la mémoire, il a été constaté que les neurones sigmoïdes avaient une plus grande capacité

de mémoire que les neurones avec une fonction d'activation de Fermi. VERSTRAETEN [257] ont étudié les performances de différentes fonctions d'activation sur différentes tâches. Pour une tâche de reconnaissance des chiffres parlés, les neurones linéaires se sont révélés les moins performants avec le taux d'erreur le plus élevé, tandis que les neurones sigmoïdes offraient des performances légèrement meilleures.

- **La régularisation** L'utilisation de la régression *ridge* pour former les poids de sortie \mathbf{W}_{out} a l'avantage de régulariser la lecture du réservoir pendant la formation, ce qui trouve le compromis optimal entre la complexité du modèle et ses capacités de généralisation, évitant le surajustement tout en offrant de bonnes performances sur des données invisibles. Un exemple de ceci a été montré par WYFFELS et SCHRAUWEN [296] où les erreurs de réseau augmentaient lorsqu'elles étaient formées sans régression *ridge*. En revanche, la régularisation conduit à un réseau plus robuste au bruit présent dans les données invisibles, ce qui se traduit par de meilleures performances.

Pratiquement, la fonction objective du terme de régularisation dans l'éq. 4.7, $(\beta\mathbb{I})$, est de pénaliser les grandes valeurs de \mathbf{W}_{out} à l'erreur quadratique entre $\mathbf{y}(n)$ et $\mathbf{y}_{cible}(n)$. Les valeurs optimales de β peuvent varier de plusieurs ordre de grandeurs, en fonction de l'instance exacte du réservoir et de la longueur des données d'apprentissage. Dans le cas d'une recherche simple exhaustive, il est conseillé de rechercher sur une grille logarithmique.

Dynamique et propriétés souhaitables d'un bon réservoir

La section précédente a décrit les hyperparamètres d'un réservoir qui peuvent être modifiés pour améliorer les performances. Cette section donne un aperçu de la dynamique souhaitable d'un réservoir qu'il convient d'obtenir pour donner de bonnes performances lorsqu'il est appliqué à une tâche donnée.

- **La propriété de l'état *Echo*** Afin d'obtenir un réservoir qui fonctionne bien sur une tâche donnée, JAEGER [258] a souligné qu'un réservoir de type ESN doit satisfaire la propriété d'état d'*écho* (**ESP**). Ce signifie essentiellement que le réservoir doit avoir une mémoire qui s'estompe (*fading memory*), où il se souvient juste assez des entrées précédentes pour que cette mémoire ne perturbe pas la représentation du réservoir de l'entrée actuelle. Cela signifie également que les entrées récentes ont une plus grande influence sur l'état actuel des neurones du réservoir que les entrées présentées dans le passé et que tout apport devrait éventuellement disparaître dans le réservoir après une période, permettant au réservoir de revenir à un état stable.

Un réservoir qui satisfait la propriété d'état d'écho (ESP) devrait, pour toute entrée, "supprimer" avec une vitesse adéquate toutes les conditions initiales du réservoir créées à l'initialisation qui ont un impact sur les activations du réservoir[267].

Afin de satisfaire la propriété d'état d'écho, il est recommandé que la plus grande valeur absolue des valeurs propres de la matrice de connexions récurrentes \mathbf{W} du réservoir, connue sous le nom de rayon spectral (ρ) (comme introduit dans la section précédente), soit proche mais pas égale à l'unité [258]. Le rayon spectral peut être interprété comme le facteur d'oubli du réseau, car il détermine dans quelle mesure le réseau oublie ses entrées précédentes. Pour des valeurs supérieures à l'unité, les réservoirs classiques deviennent asymptotiquement instables et la propriété d'état d'écho disparaît.

En règle générale, JAEGER [258] recommande d'utiliser des valeurs proches de l'unité pour les tâches qui nécessitent une longue mémoire et une valeur plus petite lorsque moins de mémoire est requise.

- **Calcul à l'*edge of stability*** Les réservoirs (\mathbf{W}_{in} et particulièrement \mathbf{W}) sont des systèmes dynamiques, excitables et généralement non-linéaires. La couche de lecture \mathbf{W}_{out} , d'autre part, est linéaire et sans mémoire, et - du point de vue de l'apprentissage - pas très puissante en termes de calcul. La fonctionnalité d'un réservoir dans ce contexte peut ainsi être décrite comme "offrant un riche ensemble de transformations temporelles non-linéaires des entrées actuelles et passées" à la couche de lecture. Il s'ensuit que le régime dynamique dans lequel le réservoir fonctionne est crucial pour la quantité d'informations que la lecture peut extraire des états du réservoir. En plus de remplir la propriété d'état d'écho qui permet au réservoir d'avoir une mémoire qui s'estompe (*fading*), le réservoir doit contenir des neurones qui mappent l'entrée sur un état riche et dynamique spatialement tout en restant entre un régime trop stable et un régime instable.

Il s'ensuit que le réservoir doit fonctionner dans un régime dynamique qui se situe entre un comportement hautement stable (où le réservoir reste dans la même région, largement non influencé par l'entrée externe) et un comportement chaotique, où chaque changement du signal d'entrée est amplifié de façon exponentielle et la plupart des informations sur l'entrée sont perdues. De plus, pour les systèmes chaotiques, la propriété d'état d'écho ne tient plus, puisque de légers changements dans les conditions initiales sont amplifiés de façon exponentielle. Ainsi, le réservoir devrait réagir juste «assez dynamiquement» pour séparer les entrées de différentes classes.

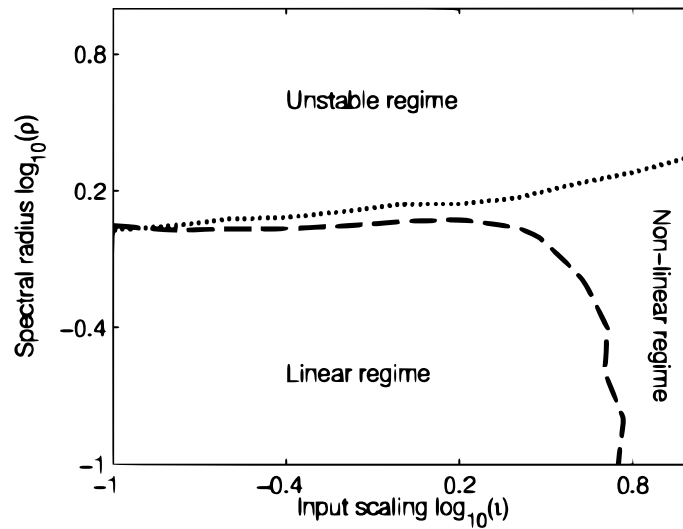


FIGURE 4.8 – De [291], différents régimes dynamiques pour les neurones réservoirs standards avec une fonction d’activation de \tanh , en fonction du rayon spectral et l’*input scaling*.

Essentiellement, un réservoir qui est trop stable ou instable aura souvent de moins bonnes performances qu’un réservoir qui contient les deux caractéristiques, car il est incapable de donner une bonne représentation des données d’entrée dans son espace d’état interne, ce qui réduit ses capacités computationnelles. Dans l’ouvrage de BERTSCHINGER et NATSCHLÄGER [22], il a été démontré que les réservoirs qui fonctionnent à la limite de la stabilité (*edge of stability*) sont plus puissants, Computationnellement, que ceux qui sont trop stables ou trop instables. Cela a été proposé à la fois pour les LSM [21, 297] et les ESN [298].

- **Cartographie non-linéaire et capacité de mémoire** De tous les paramètres du réservoir qui peuvent être modifiés, les deux paramètres qui ont la plus grande influence sur la dynamique du réservoir, sont l’échelle d’entrée (*input scaling*) et le rayon spectral. Comme indiqué dans la section précédente, l’échelle d’entrée contrôle l’impact de l’entrée sur les activations du réservoir : plus l’échelle est élevée, plus l’influence de l’entrée est élevée, tandis que le rayon spectral contrôle l’impact des états précédents du réservoir sur l’état actuel du réservoir où des valeurs élevées augmentent cet impact.

Typiquement, on souhaite obtenir un ESN dont le réservoir a une capacité de mémoire aussi élevée que possible, tout en conservant la possibilité d’effectuer des mapping non-linéaires des signaux d’entrée afin de bien séparer les données fortement non-linéaires. Il s’agit de garantir un optimum entre deux propriétés fondamentales qui résident en : l’*echo state property* et la *separation property*.

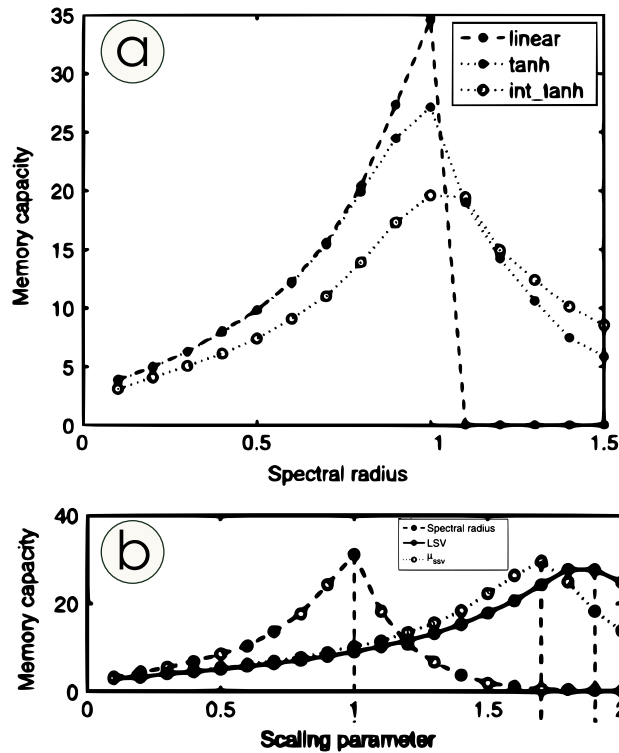


FIGURE 4.9 – De [257]. (a) : Influence de la complexité du nœud sur la capacité de mémoire, pour les nœuds *linéaires*, *tanh* et *intégrateurs tanh* en fonction du rayon spectral de la matrice de connexion. (b) : Performances sur la tâche nécessitant une capacité de mémoire en fonction de différentes métriques de réservoir (notamment l'*input scaling* et le rayon spectral).

Ce compromis entre les capacités de mapping non-linéaire et la capacité de mémoire d'un réservoir a ensuite été étudié par VERSTRAETEN et *al* [291] où les exigences de mémoire d'une tâche se sont avérées dominantes sur le mapping non-linéaire. Cela est dû au fait qu'un réservoir, qui a une quantité de capacité de mémoire plus élevée du fait que les points de travail de ses neurones sont dans leurs régions linéaires, offre de meilleures performances qu'un réservoir qui contient plus de neurones non-linéaires et perd par conséquent une partie de cette mémoire [291].

La figure 4.8 montre le comportement dynamique d'un réservoir avec différentes valeurs d'échelle d'entrée et de rayon spectral. Cela montre qu'avec des valeurs élevées pour l'échelle d'entrée et le rayon spectral, on obtient un réservoir **instable**. Avec des valeurs relativement faibles des deux paramètres, un réservoir qui fonctionne dans la région linéaire est obtenu, où le réservoir a une capacité de mémoire maximale (comme le montre JAEGER [298]), tandis que des valeurs légèrement plus élevées de l'échelle d'entrée et du rayon spectral déplacent le réservoir vers son régime non

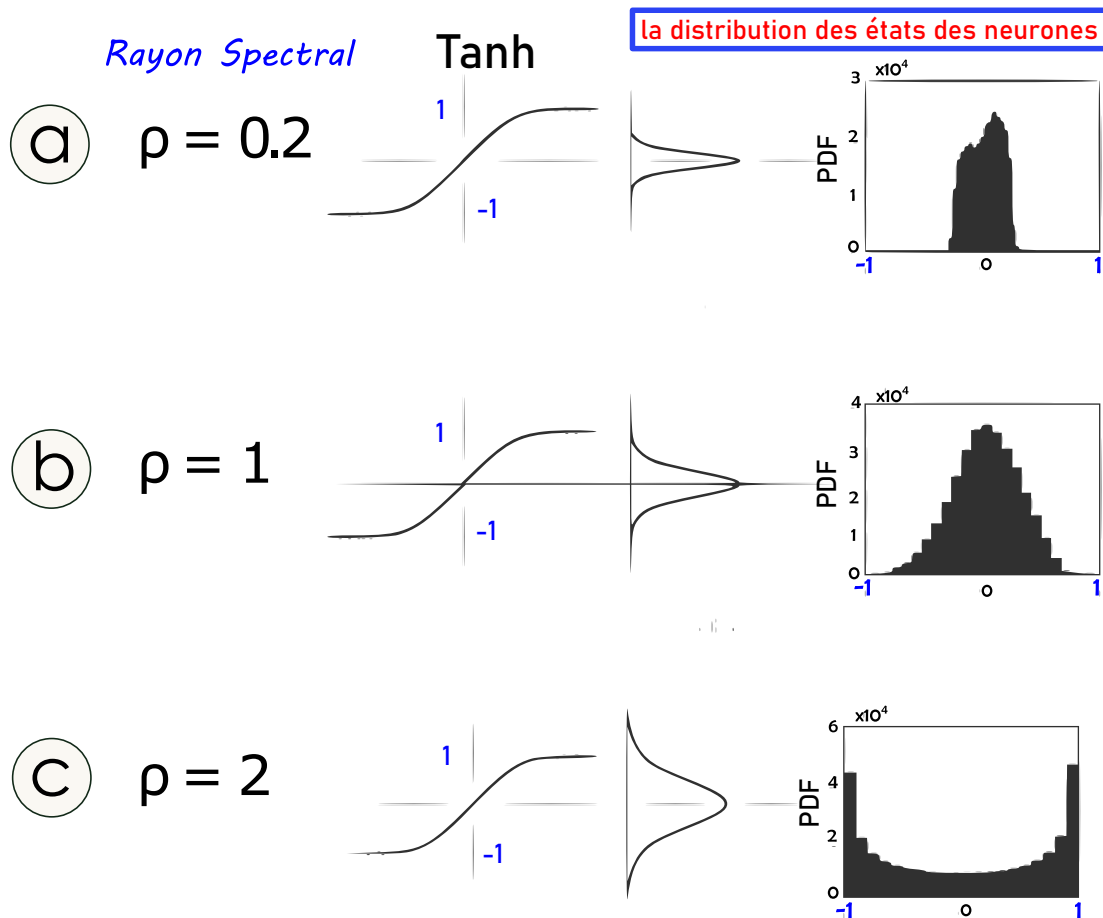


FIGURE 4.10 – De [293]. Impact des distributions d'état des réservoir : sur la figure du haut, la dynamique est trop stable pour la plupart des tâches et la distribution est assez étroite. Dans la figure au milieu, l'équilibre entre les composants linéaires et non-linéaires est juste, et dans la figure en bas, le réservoir est trop dynamique, ce qui provoque la saturation des nœuds et la perte de puissance expressive du réservoir.

linéaire.

Sur la figure 4.9, on peut remarquer qu'en prenant des valeurs d'*input scaling* et de *rayon spectral* autour de 1, les réservoirs maximisent leur capacité de mémoire tout en achevant des performances optimales. Les auteurs de ces travaux ont renforcé leur exploration visant à situer la région de paramètres qui donne accès à l'*edge of stability*. En effet, les mesures introduites ci-dessus quantifient le régime dynamique du réservoir dans son point de fonctionnement actuel. Cependant, cette dynamique peut également être liée à la distribution des états du réservoir. Ceci est illustré dans la figure 4.10 pour le cas de base des réservoirs *tanh*. La dynamique est contrôlée par le rayon spectral avec des valeurs de **.2**, **1** et **2** (pour (a),(b) et (c) respectivement).

Nous pouvons voir cela se refléter dans la distribution des états des neurones : pour un petit rayon spectral, la dynamique est trop stable et les distributions d'état sont centrées autour de la zone linéaire de la non-linéarité. Pour un paramètre de rayon spectral trop grand, les distributions d'état sont centrées autour des zones de saturation non-linéaires de la non-linéarité, ce qui se traduit par un réservoir instable et une perte d'expressivité à cause de la saturation. Sur la figure 4.10.(b), le régime dynamique est "juste", ce qui se reflète dans la distribution des états : la plupart des états se trouvent dans la zone linéaire de la fonction tanh, mais il y a toujours une bonne quantité d'états non-linéaires. Ce lien entre la dynamique et les distributions d'état renforce la localisation de la région du *edge of stability* vers des valeurs autour de l'unité et pour le *input scaling* et pour le *rayon spectral*.

Implémentation physique du *Reservoir Computing*

Le nombre d'études sur le RC a augmenté rapidement en ces dernières années. Dans cette section, nous discutons des tendances récentes dans les études autour du RC, notamment, les réalisations physiques qui ont attiré des attentions considérables. Une autre méthode consiste à utiliser d'autres systèmes dynamiques au lieu des RNN. Tout système dynamique a le potentiel de servir de réservoir s'il peut présenter des réponses dynamiques aux entrées. Ces réservoirs étaient auparavant considérés comme *exotiques* [256], mais le nombre d'études sur le RC physique a rapidement augmenté. Divers réservoirs physiques ont été proposés en utilisant différents types de systèmes physiques, de substrats et de dispositifs. Ils visent à développer du matériel d'apprentissage automatique économe en énergie.

Les réponses des systèmes RC physiques sont utilisées pour former une lecture qui est réalisée à l'aide de dispositifs physiques ou de calculs basés sur des logiciels. La régression linéaire ou un autre algorithme d'apprentissage automatique simple est utilisé dans la lecture des RC de type ESN [256].

Un exemple intrigant de RC physique dans une première étude est un RC fluide avec de l'eau dans un seau pour la reconnaissance de formes comme illustré à la figure 4.11(a)[299]. Les signaux d'entrée sont transmises aux moteurs électriques qui génèrent des ondulations sur la surface de l'eau, et les ondulations sont enregistrées à l'aide d'une caméra vidéo. À partir des signaux enregistrés, une lecture basée sur un logiciel est entraînée. Les performances de l'*ordinateur liquide* sont démontrées dans une tâche XOR et une tâche de reconnaissance de chiffres. Par la suite, de nombreux modèles de systèmes dynamiques et systèmes physiques ont été utilisés comme réservoirs potentiels.

En effet, plusieurs types de réservoirs basés sur des modèles de systèmes dynamiques

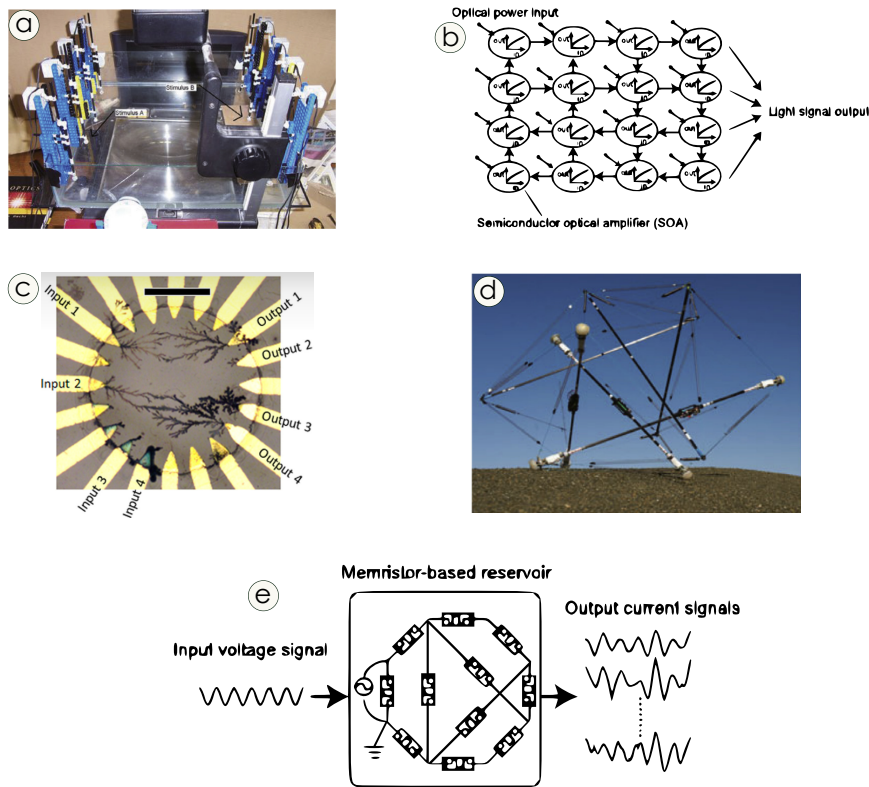


FIGURE 4.11 – (a) LSM physiquement mis en œuvre dans un seau [299]. (b) Un réservoir optique à réseau de nœuds 4×4 avec un motif de connexion en tourbillon [300]. La mémoire à long terme est fournie par les connexions de rétroaction. (c) Application d'un réservoir sur un substrat biocompatible [301]. (d) Rovers planétaires réalisés avec des réservoirs mécaniques [302]. (e) Illustration schématique d'un système RC basé sur un réservoir à base de memristor [303].

non-linéaires bien connus ont été étudiés, y compris les systèmes dynamiques retardés [304], les automates cellulaires [305] et les oscillateurs couplés [306].

Dans une autre étude, les réservoirs ont été configurés avec des substrats physiques constitués de nanotubes de carbone et de mélanges de polymères [307, 308]. Un algorithme évolutif assisté par ordinateur détermine les signaux de tension de commande et les emplacements des électrodes d'entrée/sortie pour trouver une configuration optimale du réservoir. Les résultats expérimentaux ont montré que le RC basé sur les matériaux avec la configuration de réservoir optimisée réussissait à prédire les séries chronologiques. Une autre idée récemment proposée est d'utiliser un système quantique complexe dynamique pour le RC [309]. Plusieurs réservoirs basés sur l'électronique de spin (spintronique) ont été proposés. La spintronique est un domaine de recherche émergent l'électronique à l'échelle nanométrique impliquant à la fois la charge et le spin des électrons pour développer de nouveaux

dispositifs électroniques, tels que le stockage non-volatile. Les systèmes de spin sont des candidats potentiels pour les dispositifs de réservoir à faible puissance et à petite échelle [310].

Un RC photonique, avec les avantages potentiels de faible consommation et de calcul extrêmement rapide, a été proposé pour la première fois en 2008, puis affiné sur la base de simulations numériques [300, 311, 312]. En tant que telle, une puce RC photonique en silicium est un système attrayant pour le calcul optique à ultra haute vitesse et à faible consommation d'énergie. Ensuite, une autre approche a été proposée, visant la mise en œuvre d'un réservoir photonique spatialement étendu basée sur l'imagerie diffractive en utilisant un élément optique diffractif (DOE) standard. Ainsi, BRUNNER et FISCHER [313] ont mis en évidence le couplage à l'intérieur d'un réseau de lasers à cavité verticale émettant par la surface (VCSEL). Une approche différente de la mise en œuvre photonique de RC consiste à exploiter les propriétés d'excitabilité de dispositifs photoniques spécifiques. Le comportement de pointe des dispositifs photoniques excitables, qui peuvent être mis en œuvre par la technologie des semi-conducteurs, ressemble aux propriétés des neurones biologiques [314-316]. Le concept de RC basé sur une ligne à retard, n'utilisant qu'un seul nœud non linéaire avec rétroaction retardée, a été introduit il y a quelques années par APPELTANT [304] et PACQUOT [317] comme moyen de minimiser la complexité matérielle attendue dans les systèmes photoniques. Les premières implémentations matérielles optoélectroniques de RC ont été développées indépendamment par LARGER [318] et PAQUOT [319]. Les deux réalisations expérimentales étaient basées sur l'implémentation optoélectronique d'une cavité en anneau optique de type IKEDA [320, 321].

La recherche dans ce domaine ne cesse de croître puisque la combinaison de la nanophotonique et du paradigme RC a le potentiel de révolutionner davantage le traitement de l'information photonique dans les années à venir.

4.4 Conclusion

D'abord, nous avons mis en évidence l'intersection entre le *machine learning* et les systèmes complexes afin de bien situer la deuxième partie de ce travail de thèse. Ensuite, nous avons défini le principe général du *Reservoir Computing* et introduit le modèle théorique de base d'un *Echo State Network*. Nous avons énuméré les hyperparamètres qui définissent un ESN et influencent dramatiquement la dynamique du réservoir. Plus particulièrement, nous avons discuté l'importance de la topologie du réservoir, du rayon spectral et du *input scaling* dans la maximisation des performances des ESNs. Nous avons exposé la dynamique

ainsi que le *trade-off* souhaitables dans un réservoir performant. Enfin, nous avons scanné la multitude des implémentations physiques de ce nouveau paradigme de *Reservoir Computing*.

Reservoir Computing à base d'état chimère optique

5.1 Introduction

Le reservoir computing est une approche paradigmatique du traitement de l'information basée sur des réseaux de neurones récurrents. Dans le chapitre précédent, nous avons détaillé l'anatomie ainsi que les hyper-paramètres qui influencent la dynamique collective des neurones de ce type de réseau. Également, nous avons rapporté les propriétés souhaitables dans un bon réservoir afin de générer des dynamiques internes non-linéaires qui maximisent l'apprentissage des tâches complexes. D'ailleurs, une panoplie d'investigations a prouvé l'excellence du reservoir computing dans le traitement des systèmes complexes et les dynamiques non-linéaires. De telles études se sont concentrées sur l'identification de signaux chaotiques [322, 323], l'inférence de variables partielles [214, 324], l'observation dynamique de systèmes excitables [325] et la prédiction des séries chaotiques [326].

Cependant, le lien entre structure/fonction reste toujours un enjeu contesté au sein de la communauté du reservoir computing. Inspirés principalement par le cerveau biologique, les chercheurs explorent l'efficacité d'une variété de topologies qui calquent la connectivité cérébrale, dans les tâches cognitives complexes [215, 285-288]. En d'autres termes, le reservoir computing représente un point d'agrégation qui offre une interaction interdisciplinaire entre la théorie d'information, la théorie du chaos et des systèmes complexes, la dynamique non-linéaire et les neurosciences. Par conséquent, l'implémentation de nos états chimères optiques dans l'architecture du reservoir computing, présente un examen approfondi de leurs capacités computationnelles étudiées dans la première partie de cette thèse.

Dans ce chapitre, nous allons évaluer les performances de nos réservoirs à base d'état chimère dans l'observabilité, l'inférence et la prédiction des séries chaotiques ainsi que dans la réplication totale d'attracteur chaotique. Ensuite, nous allons approfondir notre investigation à travers une analyse structurelle et dynamique, tout en adoptant une boîte d'outillage qui rassemble les neurosciences et la physique des systèmes complexes. En d'autres termes, le but de ce chapitre est de vérifier les capacités computationnelles de nos états chimères, ainsi que de dévoiler les propriétés fondamentales offertes par les patterns de type SIERPINSKI qui propulsent l'apprentissage de nos réservoirs à base d'état chimère.

5.2 La prédiction du système de Lorenz

5.2.1 Système de Lorenz

L'attracteur de LORENZ apparaît dans un système non-linéaire écrit pour la première fois par EDWARD LORENZ en 1963 pour décrire un écoulement dissipatif non-périodique dans un système hydrodynamique [35]. LORENZ motive le système comme une tentative de décrire un fluide avec une sorte de forçage externe et de dissipation thermique et visqueuse. La forme canonique du système de LORENZ est donnée par :

$$\begin{aligned}\dot{x} &= a(y - x) \\ \dot{y} &= bx - y - xz \\ \dot{z} &= xy - cz\end{aligned}\tag{5.1}$$

où a, b et c sont des paramètres et \dot{x} désigne la dérivée temporelle de x . Des valeurs différentes des paramètres amèneront le système de LORENZ à afficher un comportement périodique ou chaotique. La figure 5.1 montre l'attracteur chaotique pour le système de LORENZ, ainsi que les composants x , y et z de l'attracteur tracés en fonction du temps.

Le système de LORENZ est déterministe, ce qui signifie que si vous connaissez les valeurs de départ exactes de vos variables, vous pouvez en théorie déterminer leurs valeurs futures à mesure qu'elles changent avec le temps. LORENZ a démontré que si vous initiez ce modèle en choisissant certaines valeurs pour x , y et z , puis recommencez avec des valeurs légèrement différentes, vous arriverez rapidement à des résultats fondamentalement différents. Dans la pratique, vous ne pouvez jamais connaître la valeur exacte d'une mesure physique, même si vous pouvez vous en approcher. Avec ces résultats, LORENZ a étonné la communauté scientifique [35] en montrant qu'un système d'équations simplifié pouvait défier les méthodes

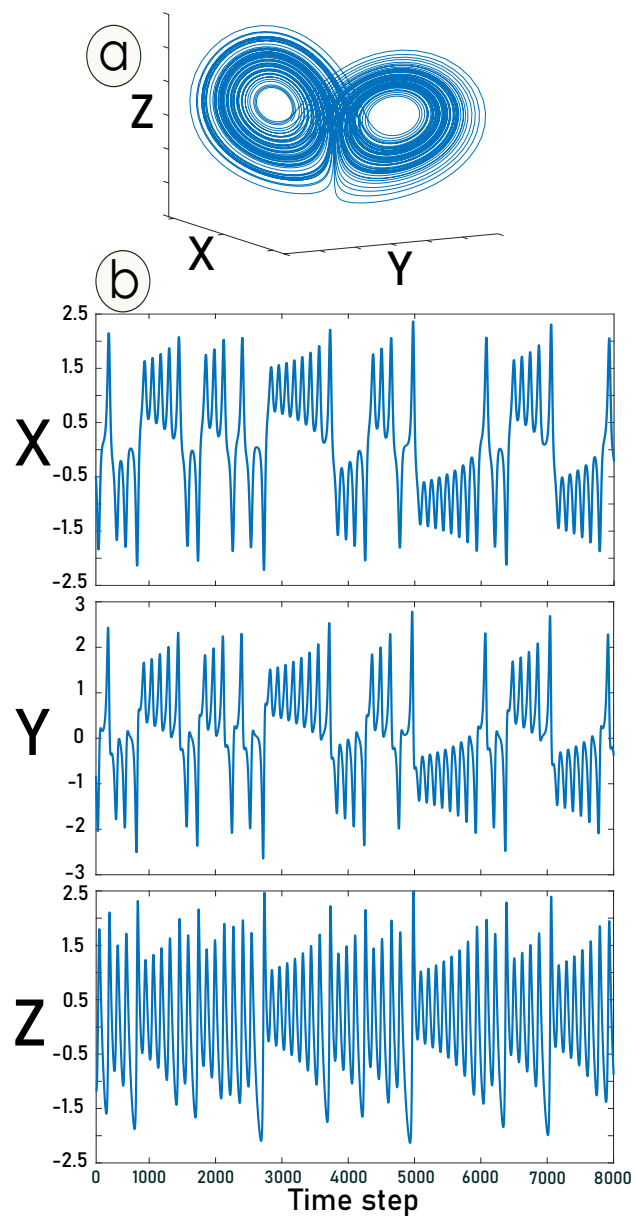


FIGURE 5.1 – L'Attracteur de LORENZ. (a) Les "ailes de papillon" caractéristiques de l'attracteur de LORENZ en 3 dimensions. (b) Les composantes temporelles $x(t)$, $y(t)$ et $z(t)$ de l'attracteur de LORENZ. La simulation numérique est établie avec $a = 10$, $b = 28$ et $c = 8/3$ (qui placent les systèmes dans le régime chaotique) et $dt = 0.01$.

conventionnelles de prédiction. C'est ce qu'on appelle le *chaos*, et ses implications sont considérables, en particulier dans le domaine de la prévision météorologique.

5.2.2 Observabilité des réservoirs à base d'état chimère

Fréquemment, lorsqu'on étudie la dynamique d'un système physique, on a accès qu'à un ensemble limité de mesures des variables d'état et on souhaite en déduire des valeurs de variables d'état non mesurées. En principe, il pourrait être possible d'atteindre cet objectif si, en plus des mesures, on a également une connaissance de la dynamique du système. Dans la théorie du contrôle, une méthode de déduction réussie de ce type est appelée un *observateur*. La question centrale de l'observabilité est : quand avons-nous suffisamment d'informations sur un système pour pouvoir reconstruire toutes les variables du système à partir d'un sous-ensemble d'entre elles seulement. Les observateurs sont d'une grande utilité pour le contrôle et la prédiction de la dynamique. Le problème de l'observateur pour le cas où le système dynamique est linéaire a été entièrement résolu dans le travail classique de KALMAN, qui a également formulé des conditions d'observabilité sous lesquelles il est possible d'atteindre le but de déduire l'état complet d'un système linéaire à partir d'un ensemble partiel donné de mesures d'état (Voir les manuels sur la théorie du contrôle, par exemple [327]).

Les observateurs et l'observabilité ont également été largement étudiés pour les systèmes dynamiques non-linéaires (par exemple [328]). Dans ce travail, nous considérons le problème de l'observateur pour les situations dans lesquelles on ne dispose pas d'un modèle mathématique suffisamment précis du système non-linéaire d'intérêt. Au lieu d'un tel modèle, nous supposons qu'il existe une période initiale pour laquelle les mesures de toutes les variables souhaitées du système sont disponibles, et nous cherchons à utiliser ces mesures dans la période de temps initiale pour déduire l'ensemble complet des variables souhaitées pour la période suivante, pour laquelle nous supposons que seules des mesures d'un sous-ensemble limité des variables souhaitées sont possibles. Les Reservoirs computing ont été utilisés avec succès comme observateurs pour les systèmes non-linéaires. Plus précisément, LU *et al* [214] ont utilisé le reservoir computing en tant qu'observateur pour déduire les composantes y et z du système RÖSSLER (un attracteur chaotique qu'on va étudier dans la section 5.3) ainsi que celui du système de LORENZ fournissant uniquement la composante x au réservoir (cela a également fonctionné pour les autres permutations où le réservoir n'a reçu que y ou z). ZIMMERMANN *et al* [325] ont démontré l'efficacité de cette approche pour *observer* la dynamique chaotique spatio-temporelle dans les milieux excitables (par exemple, le tissu cardiaque) en utilisant les réseaux d'états d'écho (ESN) pour apprendre et approximer les relations requises entre les quantités observées et non observées.

Dans cette partie, nous soulignons que l'objectif que nous abordons est l'inférence de

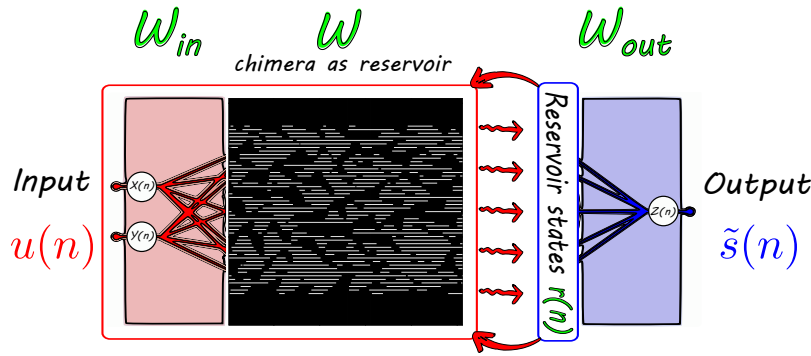


FIGURE 5.2 – L’anatomie d’un réservoir à base d’état chimère est composée de trois parties, une couche d’entrée \mathbf{W}_{in} , une couche réservoir \mathbf{W} qui est constituée d’un état chimère et une couche de sortie \mathbf{W}_{out} .

variable d’état non mesurée plutôt que sa prédiction. Pour ce faire, notre méthode utilise le reservoir computing à base d’état chimère afin de prouver sa propriété d’observabilité.

Anatomie des réservoirs à base d’état chimère

Nous nous intéressons à la situation dans laquelle \mathbf{u} et \mathbf{s} peuvent tous deux être mesurés sur une période spécifique, $[\mathbf{0}, \mathbf{T}]$, mais que seul \mathbf{u} de dimension $N_{\mathbf{u}}$ peut être mesuré à partir de ce moment ; notre méthode est d’utiliser la connaissance continue de \mathbf{u} pour estimer \mathbf{s} en fonction du temps lorsque la mesure directe de \mathbf{s} n’est pas disponible (en $t < T$).

À cette fin, nous utilisons le reservoir computing à base d’état chimère. L’implémentation spéciale des ESNs avec des neurones à fuite (Leaky integrator neurons LIN) nous aidera dans notre étude. Comme expliqué dans le chapitre précédent, la dynamique du réservoir dans ce cas est donnée par :

$$\mathbf{r}(n+1) = (1 - \alpha)\mathbf{r}(n) + \alpha \tanh(\mathbf{W}_{in}[\text{bias}; \mathbf{u}(n+1)] + \mathbf{W}\mathbf{r}(n) + \nu) \quad (5.2)$$

Afin de construire notre réservoir, nous avons préparé les dynamiques des états chimères sous forme de matrice carrée ($N \times N$ avec N est le nombre de neurones). La matrice a ensuite été remise à l’échelle au rayon spectral ρ spécifié par l’utilisateur. Pour remettre à l’échelle la matrice, nous suivons la procédure spécifiée par JAEGER [298] (voir le chapitre précédent). Finalement, nous avons implémenté cette matrice comme couche réservoir \mathbf{W} (Fig. 5.2).

La matrice \mathbf{W}_{in} de dimension $N * (1 + N_{\mathbf{u}})$, est créée d’une façon random. Dans le tableau 5.1, nous présentons les hyper-paramètres utilisés dans la construction de notre réservoir à base d’état chimère.

TABLEAU 5.1 – Les hyper-paramètres utilisés dans notre réservoir à base d'état chimère pour *observer* la dynamique de la variable z du système de LORENZ.

Paramètre	Valeur
Taille du réservoir (N)	256
Rayon spectral de la matrice W (ρ)	1
Bruit (ν)	0.0001
Taux de fuite (α)	0.3
Input scaling (\mathbf{W}_{in})	1
Bias	1
Durée de l'entraînement	20985
Méthode de régression	Ridge
Paramètre de régularisation (β)	1e-7

Inférence sans-modèle de variable non mesurée

Le but est de tester les performances du reservoir computing à base d'état chimère, en tant qu'observateur, de la boîte noire des systèmes dynamiques non-linéaires sur l'exemple du système chaotique de LORENZ (introduit dans la section précédente). Nous choisissons deux des variables du système de LORENZ comme mesure "continuellement disponible" $u(t)$. L'autre coordonnée est prise comme la composante cible $s(t)$ que le réservoir est entraîné à reproduire, et dont la connaissance n'est disponible que pour la phase d'apprentissage $0 \leq t \leq T$ ($T = 8455$).

Nous continuons ensuite à intégrer, à piloter le système avec les deux variables $u(t)$ et à tester à quel point la sortie générée $\tilde{s}(t)$ par le reservoir computing à base d'état chimère est proche de la variable non observée.

Pour faciliter la comparaison mutuelle entre les trois variables, nous utilisons des versions à échelle unitaire de x , y et z , que nous notons \tilde{x} , \tilde{y} et \tilde{z} . Par exemple, \tilde{z} est définie par :

$$\tilde{z}(t) = \frac{z(t) - \langle z(t) \rangle}{\sqrt{\langle [z(t) - \langle z(t) \rangle]^2 \rangle}} . \quad (5.3)$$

Dans notre cas d'étude, les données d'entrée sont $u(t) = [\tilde{x}(t), \tilde{y}(t)]$ (disponibles dans les phases d'entraînement et de test) or la variable cible non mesurée est $s(t) = [\tilde{z}(t)]^T$ (disponible que dans la phase d'entraînement).

Nous constatons que notre réservoir à base d'état chimère entraîné (Avec $\tilde{x}(t)$ et $\tilde{y}(t)$) a pu générer une sortie $\tilde{z}(t)$ qui est très proche de celle du système de LORENZ (Fig. 5.3a). Dans la Fig. 5.3b, nous présentons un test plus précis que l'inspection visuelle de la Fig. 5.3a

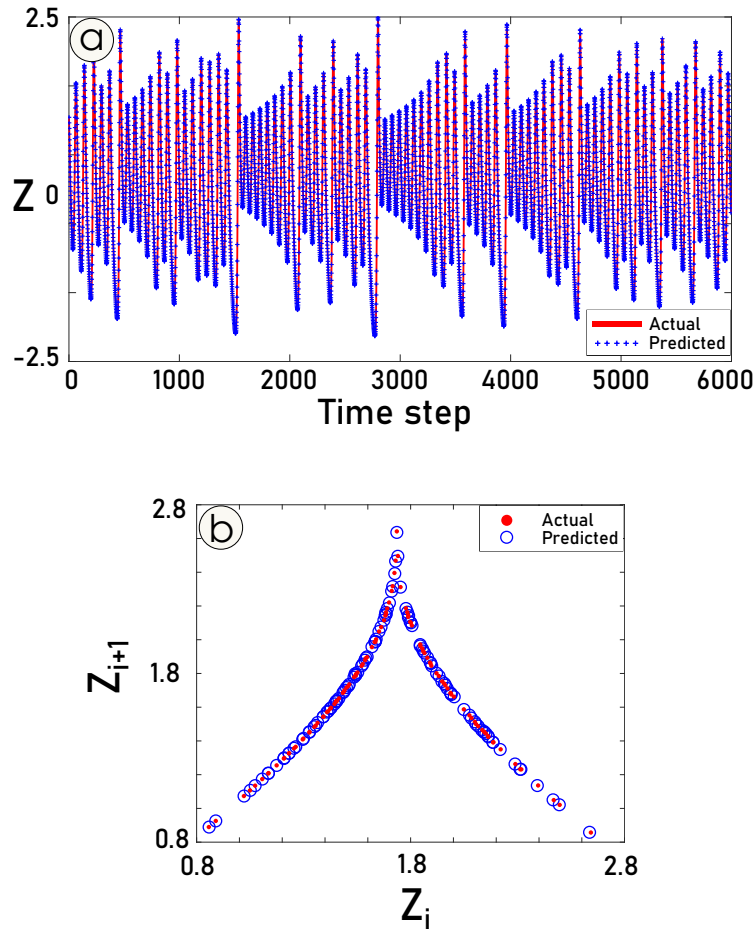


FIGURE 5.3 – Observabilité des réservoirs à base d’état chimère. (a) : La composante $\tilde{z}(t)$ du système de LORENZ générée par le réservoir à base d’état chimère (en +), déduite des entrées $\tilde{x}(t)$ et $\tilde{y}(t)$ tandis que la ligne rouge est la valeur réelle obtenue par intégration numérique de l’équation de LORENZ eq. (5.1). (b) : La carte de 1^{er} retour de la coordonnée $\tilde{z}(t)$ réelle et inférée du système de LORENZ. Ce graphique est réalisé avec la série chronologique en (a), où les points rouges proviennent du système LORENZ réel, tandis que les cercles en bleu superposés aux points rouges proviennent de l’inférence établie par notre réservoir.

afin de vérifier la précision du résultat obtenu. Pour ce faire, nous suivons la procédure de LORENZ consistant à tracer la carte de retour des maxima successifs de $\tilde{z}(t)$. Nous obtenons d’abord $\tilde{z}(t)$ pour une longue période de temps, pour les séries temporelles réelles et inférées. Nous localisons ensuite tous les maxima locaux des $\tilde{z}(t)$ réels et inférés dans l’ordre temporel et les notons $[\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_m]$. Ensuite, nous traçons des paires consécutives de ces maxima $[\tilde{z}_i, \tilde{z}_{i+1}]$ pour $i = 1, \dots, m - 1$ sous forme de points sur la Fig. 5.3b. Les points rouges proviennent de l’intégration numérique du système de LORENZ réel, tandis que les cercles bleus proviennent de l’inférence de sortie de notre réservoir à base d’état chimère. Comme

le confirme la figure Fig. 5.3b, Les deux résultats sont en très bon accord tout au long de l'exécution. Ainsi, notre réservoir réussit très précisément à reproduire une des propriétés fondamentales de l'attracteur chaotique de LORENZ.

Dans cette section, nous avons montré que le reservoir computing peut être un outil très efficace et polyvalent pour reconstruire de manière robuste un système dynamique non linéaire dont une des variables est non mesurée. D'autre part, l'implémentation de nos états chimères optiques dans l'architecture du reservoir computing comme étant une matrice de connexion, a maintenue un très bon niveau de performances ainsi que la propriété d'observabilité. Ce premier résultat nous a permis de confirmer le pouvoir calculatoire et les capacités de traitement d'information de nos états chimères optiques. Cependant, il est plus judicieux d'explorer les propriétés fondamentales qui ont permis à nos états chimères optiques d'effectuer avec succès l'inférence d'une variable chaotique non mesurée. Dans la section suivante, nous allons essayer de pousser plus loin nos états chimères optiques tout en effectuant une analyse approfondie afin de comprendre les mécanismes contribuant à ce résultat.

5.2.3 Analyse structurelle et dynamique des réservoirs à base d'état chimère

Les ESNs classiques ont un réservoir interne \mathbf{W} de neurones connectés de manière dispersée et aléatoire. Le réservoir fonctionne comme une projection non-linéaire interne des entrées, générant une riche variété de caractéristiques, de sorte que la couche de lecture \mathbf{W}_{out} peut séparer linéairement les motifs plus facilement. Ainsi, la performance d'un reservoir computing est liée à la richesse des représentations générées par les neurones du réservoir, elles-mêmes liées au schéma de connectivité entre les neurones du réservoir \mathbf{W} . Comme indiqué dans le chapitre précédent, les chercheurs dans ce domaine ont exploré l'utilité d'une panoplie de topologies inspirées nécessairement par le cerveau humain [215, 285-288]. Ils ont étudié la relation entre la structure et la fonctionnalité du réservoir, car elle dépend essentiellement de sa topologie (sa connectivité).

Dans la même ligne directrice, nous investiguons, dans cette section, les éléments cruciaux qui caractérisent la topologie de nos réservoirs à base d'état chimère. Nous proposons une analyse structurelle et dynamique dans l'optique de comprendre les mécanismes contribuant au traitement optimal d'information, menant ainsi à des excellentes performances dans l'inférence et la prédiction des séries chaotiques.

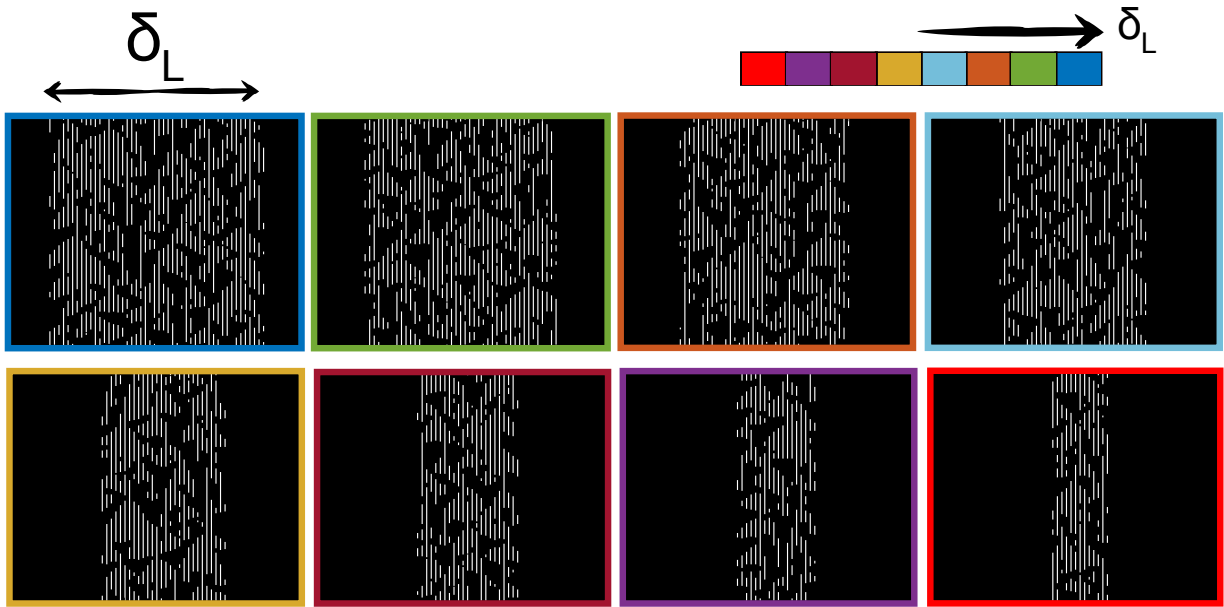


FIGURE 5.4 – Les 8 états chimères utilisés dans ce travail comme des matrices de connexion dans l’architecture de nos réservoirs. Chaque couleur dans la barre de couleurs indique une taille spécifique δ_L de la partie incohérente des états chimères.

Inférence multi-pas en avant en fonction de la taille des états chimères

Dans les trois premiers chapitres de cette thèse, nous avons montré que nos états chimères optiques partagent un ensemble de propriétés computationnelles avec la classe complexe des automates cellulaires. A travers une analogie qualitative et quantitative, nous avons conclu que les patterns triangulaires de type SIERPINSKI (présents dans la partie incohérente de nos états chimères optiques) jouent un rôle fondamentale dans le traitement d’information.

Dans le cadre de notre analyse structurale et dynamique des réservoirs à base d’état chimère, nous allons approfondir notre compréhension du potentiel calculatoire des patterns auto-organisés de nos états chimères. Pour ce faire, nous menons une étude statistique des performances de nos réservoirs en augmentant la taille de la partie incohérente (qui accueille les domaines triangulaires), tout en gardant fixe le nombre de 256 neurones dans la matrice de connexion \mathbf{W} (Fig. 5.4).

La tâche demandée aux réservoirs sera l’inférence d’une variable non mesurée du système de LORENZ. En effet, les composants d’entrée sont $u(t) = [\tilde{y}(t - n_{sh}), \tilde{z}(t - n_{sh})]$ (disponibles dans les phases d’entraînement et de test, et n_{sh} étant le nombre de pas de temps du décalage temporel) alors que la variable cible non mesurée est $s(t) = [\tilde{x}(t)]^T$ (disponible que dans la phase d’entraînement). La durée d’entraînement est de 8445 et celle du test est de 3620

TABLEAU 5.2 – La grille de recherche utilisée dans l'inférence multi-pas d'une variable non mesurée du système de LORENZ par les réservoirs à base d'état chimère.

Hyperparamètre	Interval des valeurs	Nombre de points
Rayon spectral de la matrice W (ρ)	[0.1,1.1]	51
Taux de fuite (α)	[0,1]	100
Paramètre de régularisation (β)	[1e-10,1e-1]	100

(avec $dt = 0.01$). Dans l'optique d'avoir une vision plus globale sur les performances de nos réservoirs à base d'état chimère, nous avons étudié l'effet des trois hyper-paramètres majeurs dans l'architecture du concept du reservoir computing. Il s'agit du rayon spectral (ρ), le taux de fuite (α) et le paramètre de régularisation (β). Par conséquent, notre calcul se base sur une grille de recherche tridimensionnelle très fine (510000 manipulations). Les détails des valeurs sont donnés le Tab.5.2.

Cette grille est appliquée à chaque réservoir constitué d'une taille d'état chimère δ_L comme étant sa matrice de connexion \mathbf{W} (Fig. 5.5). En utilisant la même méthodologie expliquée dans la section précédente, nous calculons la racine de l'erreur quadratique moyenne **RMSE** entre la variable cible non mesurée $s(t)$ et la série produite par le réservoir à base d'état chimère $\tilde{s}(t)$:

$$\text{RMSE}(s(t), \tilde{s}(t)) = \sqrt{\frac{\sum_{i=0}^{N-1} (s_i - \hat{s}_i)^2}{N}} \quad (5.4)$$

Les grilles de recherche 'tridimensionnelles' (Fig. 5.5) nous permettront, par la suite, d'établir une analyse fonctionnelle et dynamique de nos réservoirs tout en explorant l'effet de la taille de la partie incohérente δ_L des états chimères optiques sur les performances.

Dans la figure 5.5, nous avons présenté les résultats du cas où $n_{sh} = 15$ (le décalage temporel de $\tilde{y}(t - n_{sh})$ et $\tilde{z}(t - n_{sh})$). Nous remarquons qu'en augmentant la taille de la partie incohérente δ_L de nos états chimères optiques, on élargit l'intervalle des hyper-paramètres menant vers de très bonnes performances (des erreurs minimales en bleu foncé). Nous constatons, également, un point commun entre toutes les tailles qui se réside dans la présence d'une 'Vallée de stabilité' autour des petites valeurs du taux de fuite α . La largeur de cette 'Vallée' augmente avec la taille δ_L . En outre, nous observons qu'en se mettant dans cette 'Vallée', les performances de nos réservoirs deviennent quasi-indépendantes aux rayon spectral ρ et paramètre de régularisation β .

Ces remarques sont confirmées dans la figure 5.6, où on présente le RMSE en fonction de la taille δ_L des états chimères, en traçant la moyenne et l'écart-type des meilleures performances (les minimums) de chaque grille. Nous avons remarqué que l'observation reste

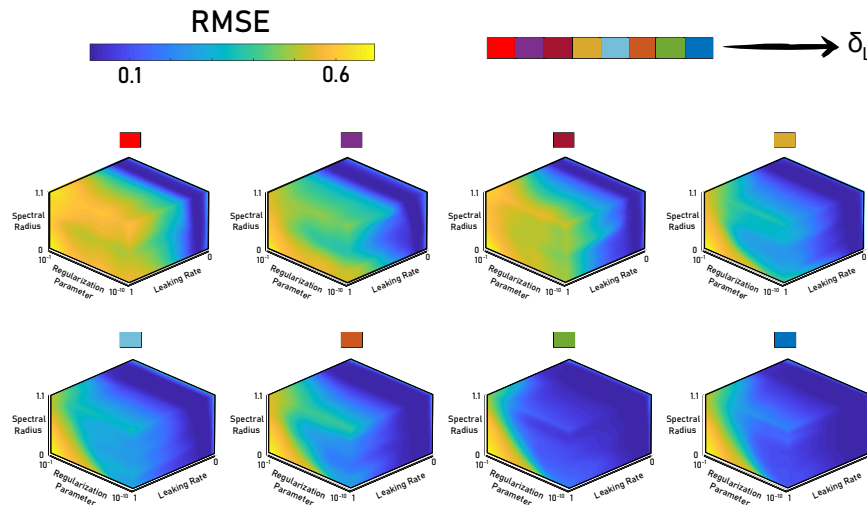


FIGURE 5.5 – Les grilles de sondage des différentes tailles δ_L pour le cas de $n_{sh} = 15$. L'apparition d'une 'Vallée de stabilité' (en bleu) est un point en commun pour toutes les tailles d'état chimère.

identique pour les différentes valeurs de n_{sh} (Fig. 5.6(a),(b),(c) et (d)). Plus on augmente δ_L , plus les réservoirs excellent dans l'inférence de la variable non mesurée (RMSE devient petit) et plus la robustesse des résultats est renforcée (l'écart-type devient petit).

Comme nous avons vu dans le chapitre précédent, le rôle de la couche cachée \mathbf{W} dans le réservoir computing est de transformer non linéairement les entrées séquentielles dans un espace de grande dimension tel que les caractéristiques des entrées peuvent être efficacement lues par un simple algorithme d'apprentissage. En se basant sur cette définition fondamentale, nous pouvons conclure que nos états chimères optiques à travers les patterns triangulaires possèdent les ingrédients de base pour améliorer l'architecture computationnelle du réservoir computing. En augmentant la taille δ_L , la présence des domaines type SIERPINSKI garantissent la haute dimensionnalité et la non-linéarité nécessaires pour cartographier les entrées dans un grand espace afin de faciliter la séparation et de lire les dépendances spatio-temporelles des entrées dans les tâches d'inférence et de prédiction. D'autre part, la présence d'une "Vallée de stabilité" vers des taux de fuites petits (Fig. 5.5) est justifié par la nature chaotique du système de LORENZ. La dépendance non-linéaire à long et court terme dans la série chaotique impose une dynamique lente au réservoir afin d'optimiser son apprentissage. En effet, l'augmentation de δ_L a pour effet d'améliorer la mémoire interne individuelle des neurones dans le réservoir \mathbf{W} , et par conséquent, on observe un élargissement de la 'Vallée de stabilité'.

En conclusion, nos réservoirs à base d'états chimères ont garanti un optimum entre le

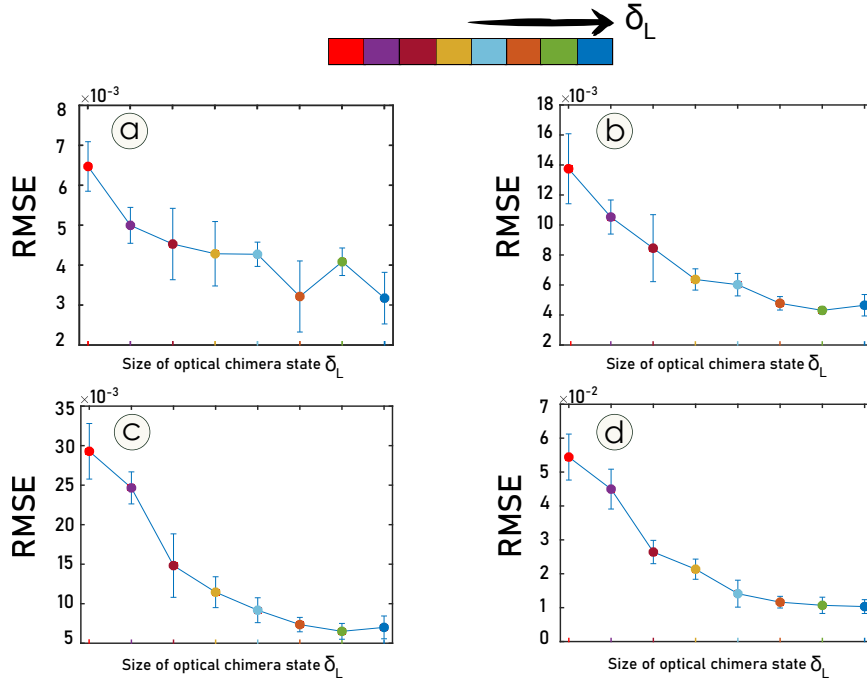


FIGURE 5.6 – L'erreur en fonction de la taille δ_L . (a) : $n_{sh}=10$, (b) : $n_{sh}=15$, (c) : $n_{sh}=20$ et (d) : $n_{sh}=25$.

mappage non-linéaire et la capacité de mémoire. Cet optimum est traduit par des performances maximisées en augmentant la taille δ_L qui est accompagnée par l'augmentation du nombre des patterns triangulaires. Dans la section suivante, nous allons explorer minutieusement le rôle de la complexité des domaines SIERPINSKI dans la création de la 'Vallée de stabilité'. En d'autres termes, nous effectuons une analyse de la dynamique de nos réservoirs à base d'état chimère tout en mettant le doigt sur les propriétés fondamentales qui améliorent l'apprentissage.

Analyse dynamique

Nous avons énuméré dans le chapitre précédent les caractéristiques souhaitables dans la dynamique du réservoir afin d'obtenir des performances optimales dans le cadre d'une tâche donnée. Nous avons vu que la dynamique générée par le réservoir doit satisfaire un optimum entre deux propriétés fondamentales 'antagonistes' : il s'agit de l'état d'écho et celle de la séparation (approximation non-linéaire).

La propriété d'état d'écho est un aspect crucial dans le traitement séquentiel de l'information. Elle reflète la capacité du réservoir à se souvenir juste assez des entrées passées pour les utiliser dans le traitement et dans la représentation des entrées actuelles et futures.

En d'autres termes, il s'agit d'une mémoire qui s'estompe avec le temps menant vers une dynamique stable. Le type de cette mémoire est lié à la nature de la tâche demandée ainsi que la complexité des données temporelles.

D'autre part, la propriété de séparation est étroitement attribuée à la capacité du réservoir à discriminer les différents motifs et caractéristiques non-linéaires cachées dans les entrées. Ce pouvoir du mapping non-linéaire résulte en une dynamique riche et diversifiée des activations neuronales du réservoir, menant à une expressivité renforcée (La capacité à approximer et à présenter différentes relations entre les entrées et les sorties).

Le mariage équilibré des deux propriétés nécessite une recherche fastidieuse de l'espace d'hyper-paramètres qui permet de positionner la dynamique du réservoir dans la région du bord de la stabilité (edge of stability). Il s'agit d'une région, où la dynamique du réservoir n'est ni trop stable ni chaotique. Autrement dit, le réservoir parvient à maintenir une mémoire à court et long terme des entrées tout en étant capable de générer des motifs dynamiques diversifiés en réponse à des entrées différentes.

Le rayon spectral est le paramètre le plus important dans la conception du reservoir computing et le contrôle de sa dynamique [258, 267, 290, 291]. Il constitue une voie pertinente permettant d'explorer la région du bord de la stabilité, connue par sa puissance computationnelle. Cependant, le choix optimal du rayon spectral reste un énorme défi et il nécessite souvent des ajustements minutieux en fonction de la nature de la tâche demandée et les données disponibles [258, 267]. En réalité, le rayon spectral doit, d'une part, assurer la propriété d'écho en contrôlant la vitesse et le temps d'influence des entrées passées sur les états internes actuels du réservoir. D'autre part, ce paramètre reste crucial dans la stabilité de la dynamique du réservoir, puisque des valeurs inappropriées peuvent la déstabiliser et nuire ainsi à la propriété d'écho. Finalement, le rayon spectral doit garantir une distribution diversifiée des activations des neurones du réservoir, en couvrant les points critiques de travail de la fonction d'activation (dans notre cas la tanh), à savoir ses parties linéaires et non-linéaires.

Dans la même ligne directrice, nous avons investigué l'effet du rayon spectral sur les performances des réservoirs à base d'état chimères tout en se mettant dans la 'Vallée de stabilité' (vu dans la Fig. 5.5). Dans la figure 5.7, nous avons tracé les erreurs RMSE en fonction du rayon spectral tout en changeant la taille de la partie incohérente δ_L des états chimères (pour les différentes valeurs de n_{sh}). Les erreurs sont moyennées sur les 100 valeurs de paramètre de régularisation (β) utilisées dans la grille de recherche (Tab.5.2).

Supposant que la région du bord de stabilité est définie par une excellente performance computationnelle qui se traduit par un **RMSE** ≤ 0.1 (Fig. 5.7). Nous constatons que le

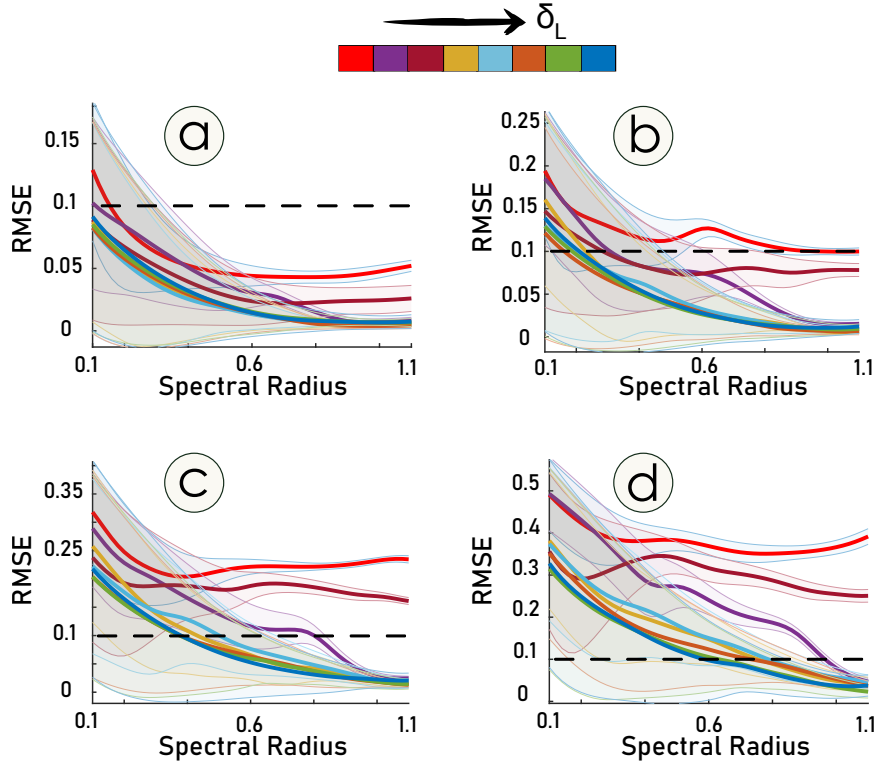


FIGURE 5.7 – L'erreur des différentes tailles δ_L en fonction du rayon spectral. La ligne en pointillés détermine la région de 'stabilité' avec un $\mathbf{RMSE} \leq 0.1$. (a) : $n_{sh}=10$, (b) : $n_{sh}=15$, (c) : $n_{sh}=20$ et (d) : $n_{sh}=25$.

point commun entre les sous-figures (Fig. 5.7(a),(b),(c) et (d)) est que plus δ_L augmente plus les courbes qui représentent les variations des RMSE en fonction du rayon spectral appartiennent à la région du bord de stabilité. En d'autres termes, l'augmentation de la partie incohérente des états chimères dans nos réservoirs a pour conséquence d'élargir l'intervalle des rayons spectraux qui placent les performances dans la région optimale. D'autre part, quand on augmente le décalage temporel des composantes observées ($\tilde{y}(t - n_{sh})$, $\tilde{z}(t - n_{sh})$) la tâche d'inférence devient au fur et à mesure complexe et par conséquent la région du bord de stabilité rétrécit tout en confirmant davantage la conclusion précédente. Remarquablement, à partir de la taille δ_L illustrée en courbe jaune (5^{ème} grande taille), les performances de nos réservoirs stagnent indépendamment du rayon spectral choisis ou de l'augmentation de δ_L (Fig. 5.7). Cette robustesse indique une saturation du seuil d'apprentissage de nos réservoirs à base d'état chimères. Nos réservoirs maintiennent une stabilité dans la dynamique interne, illustrée par des RMSE minimales même si on dépasse des valeurs de rayon spectral de 1. En revanche, les réservoirs classiques (Random) présentaient une dynamique instable ou chaotique lorsque le rayon spectral était supérieur à 1, ce qui pourrait fortement

dégrader leurs performances d'apprentissage [258, 267, 290, 291, 298].

En collectant ces remarques ensemble, on peut dire que la présence des patterns triangulaires de type SIERPINSKI dans nos réservoirs à base d'état chimères, est étroitement liée au maintien d'une dynamique neuronale adéquate pour une région étendue des valeurs optimales de rayon spectral. En effet, ces domaines auto-organisés se comportent comme étant des clusters qui assurent une connectivité locale et globale entre les différents neurones, constituant ainsi un réservoir avec une dynamique interne collective riche. Localement, les connexions intra-clusters forment des boucles de rétraction, ce qui permet un traitement récurrent des entrées. Cette connectivité locale entre les neurones d'un cluster donné, favorise une mémoire à court terme qui permet au réservoir d'utiliser des informations récentes dans la génération de sa dynamique future. Globalement, les connexions inter-clusters favorise une propagation de l'information à travers toutes les régions du réseau. Cette connectivité globale entre des neurones appartenant à des clusters différents, contribue à une intégration à longue portée de l'information permettant ainsi à capturer des dépendances temporelles à long terme.

Par conséquent, la distribution spatiale des domaines triangulaires de nos états chimères fournit un mélange des deux types de connectivités à l'architecture de nos réservoirs. Ce mélange alimente la dynamique générée avec une mémoire à court et long terme nécessaire pour maintenir la propriété d'écho. Particulièrement, cette combinaison justifie les performances élevées de nos réservoirs dans l'inférence et la prédiction des séries chaotiques dont les dépendances temporelles sont multi-échelles et complexes.

D'autre part, les clusters peuvent être considérés comme des sous-réservoirs impliquant un parallélisme fructueux dans le traitement des informations multi-caractéristiques. Cette vision explique le maintien de la stabilité dans nos réservoirs ainsi que la robustesse des résultats obtenus sur un intervalle large de rayon spectral. En fait, les clusters jouant le rôle des centres de relais de l'information (Hubs), renforcent la résilience et l'immunité de la dynamique de nos réservoirs contre les perturbations et les blocages. En cas de la défaillance d'un seul sous-réservoir par exemple, le flux d'information n'est pas nécessairement affecté et la dynamique du réservoir résiste en restant stable.

Par ailleurs, la communication entre ces sous-réservoirs consolide la profondeur computationnelle de nos réservoirs, en améliorant leur expressivité. Cette dernière est garantie par la réponse non-linéaire de la dynamique spatio-temporelle du réservoir aux données d'entrées. Finalement, la richesse des représentations générées par nos réservoirs contribue à la séparation et à la détection des relations complexes dans les données chaotiques justifiant ainsi l'insensibilité des performances à la variation du paramètre de régularisation

notamment à partir d'un rayon spectrale de 0.6 (Fig. 5.7).

Dans cette section, nous avons confirmé à travers une analyse dynamique, le lien étroit entre la connectivité offerte par nos états chimères et la capacité d'apprentissage de nos réservoirs. A l'inverse des architectures classiques (Random), nos réservoirs à base d'état chimère réussissent à générer une dynamique neuronale adéquate tout en évitant des optimisations fastidieuses des hyper-paramètres. En outre, la connectivité sous forme de clusters est capable d'imiter des mécanismes d'apprentissage plus réels et proches du cerveau biologique. En effet, de nombreuses études en neurosciences, ont trouvé des preuves d'attributs topologiques non-aléatoires qui améliorent théoriquement la capacité du cerveau humain à traiter l'information et stabiliser de manière robuste la dynamique neuronale [329]. Notamment un regroupement (clustering) élevé [330-333], des communautés ségréguées [334-336] et un noyau de nœuds centraux densément interconnectés [337, 338]. Dans le cadre du reservoir computing, les spécialistes se sont inspirés des attributs topologiques du connectome cérébral humain. Jarvis et ses co-auteurs [339] ont examiné l'effet des clusters, de la hiérarchie et de la connectivité intercluster sur le pouvoir prédictif du rayon spectral. Zhang et Deng [340] ont proposé un modèle étendu avec un réservoir ayant une topologie hybride (small-world et scale free). Najibi [288] a suggéré une structure communautaire et une architecture distribuée. KAWAI [286], DAMICELLI [341] et SUAREZ [17], ont utilisé directement une connectivité cérébrale réelle dans l'architecture du réservoir. La conclusion commune entre l'ensemble de ces recherches se réside en l'importance de la topologie en cluster dans l'élargissement de la portée de la propriété d'état d'écho et la stabilité de la dynamique du réservoir vis à vis du rayon spectral. Cependant, il reste toujours des efforts d'optimisation afin de contrôler l'interaction très sensible entre la taille des clusters, le nombre des intra/inter-connexions, le ratio entre la taille des clusters et la taille totale du réservoir et le nombre de pivots par cluster.

L'originalité de notre type de connectivité proposée tire son envergure de la quasi-indépendance des performances aux hyper-paramètres les plus influenceurs tels que le rayon spectral, le paramètre de régularisation et la taille du réseau. La forme géométrique triangulaire des patterns contribue automatiquement à l'optimisation du flux d'information à travers les différentes régions du réservoir. La saturation de l'apprentissage à partir d'une certaine taille δ_L de partie incohérente d'état chimère, signale probablement l'achèvement d'un seuil qui concerne le nombre ou la taille des clusters ou d'une propriété inconnue. Dans la section suivante, nous essayerons de répondre à ces deux derniers points énigmatiques en effectuant une analyse structurelle. La théorie d'information nous servira comme un chantier didactique afin de quantifier et d'illustrer les opérations basiques utilisées dans le

traitement de l'information. Une approche pluridisciplinaire combinant les neurosciences, la thermodynamique hors équilibre et les automates cellulaires, nous permettront de dévoiler la propriété fondamentale responsable des performances élevées de nos réservoirs à base d'état chimère ainsi que la saturation observée de l'apprentissage.

Analyse structurelle

La relation entre structure et fonction est un concept central dans les sciences naturelles et ingénierie. Les fonctionnalités du système nerveux sont façonnées de manière analogue par la structure et les dispositions des populations nerveuses. Le réseau complexe de projections synaptiques forme une hiérarchie de circuits neuronaux imbriqués et polyfonctionnels qui soutiennent la perception, la cognition et l'action. Ces schémas de câblage complets du système nerveux, appelés réseaux de connectivité structurelle ou connectomes, représentent les connexions physiques entre les éléments neuronaux. L'émergence des neurosciences offre l'opportunité de quantifier et d'articuler le lien entre les caractéristiques organisationnelles des réseaux neuronaux et le spectre des fonctions corticales. L'architecture des connectomes confère une signature distincte aux modèles de co-activation neuronale [17, 286, 341]. Les projections interrégionales favorisent la signalisation et la synchronisation entre les populations neuronales distantes, donnant lieu à une dynamique neuronale cohérente.

Tout au long de ces décennies d'efforts pour étudier les fonctions des composants individuels du cerveau, les neurosciences ont relevé ce défi grâce à un éventail éclectique de curiosités adoptées de la physique : mécanique statistique des réseaux complexes, thermodynamique et la théorie des systèmes dynamiques non-linéaires [342].

D'autre part, le cerveau est un système de traitement de l'information, qui a évolué au fil des millions d'années pour coder et manipuler un flux continu de signaux sensoriels. En tant que telle la théorie de l'information, fournit une perspective convaincante à travers laquelle on peut découvrir les relations causales du flux d'informations entre les éléments neuronaux du cerveau [342].

Inspirés par cette philosophie, nous adoptons, d'abord, le chantier de la théorie d'information afin de quantifier les opérations basiques qui calquent les activations neuronales dans nos réservoirs à base d'état chimère. Ensuite, nous utilisons l'intersection spectaculaire entre certaines notions physiques et les neurosciences dans le but d'investiguer les caractéristiques computationnelles offertes par la connectivité structurelle sous forme des patterns SIERPINSKI.

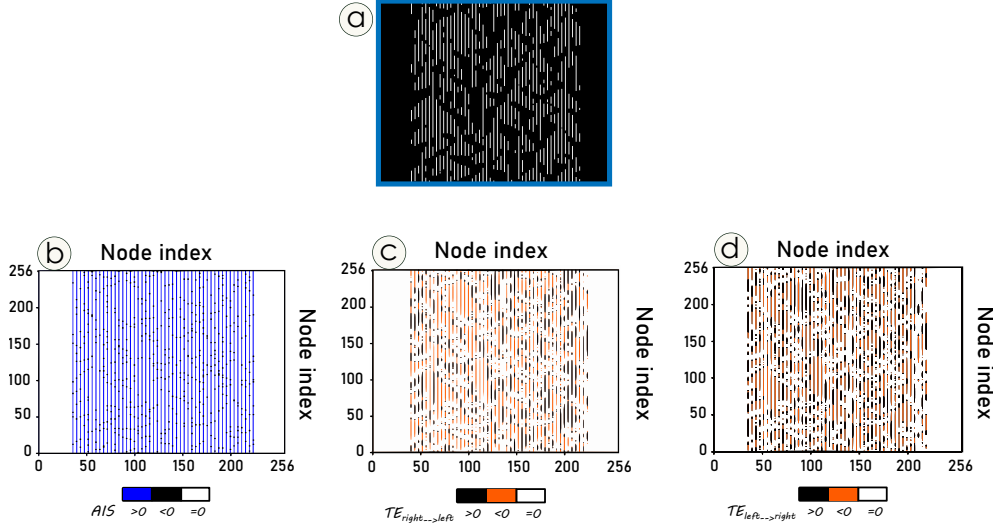


FIGURE 5.8 – Analyse locale du traitement d’information dans nos états chimères optiques. (a) : la matrice de connexion avec δ_L la plus grande. (b) : Son profil local du stockage d’information **AIS**. (c) : Son profil local de transfert d’information **TE** de droite vers la gauche . (d) : Son profil local de transfert d’information **TE** de gauche vers la droite. **AIS** est calculée pour un $k = 5$. **TE** est calculée pour un $k = 8$ et vers les neuds $(i + 4)$ et $(i - 4)$.

Théorie d’information Les mesures adoptées pour quantifier le flux de l’information sont principalement : l’entropie de transfert (TE) et le stockage actif de l’information (AIS). Comme nous avons vu dans le chapitre 3, elles ont été utilisées pour quantifier les propriétés dynamiques non-linéaires intrinsèques de systèmes complexes, en particulier les systèmes effectuant du calcul distribué [343, 344]. Ces deux grandeurs sont, particulièrement, utiles pour mesurer la manière dont les informations stockées et transportées par les signaux se propagent à travers un réseau et peuvent aider à identifier les composants les plus importants impliqués dans la dynamique d’un réseau [345, 346], et les performances du réseau peuvent ainsi être optimisées en conséquence [347].

D’abord, nous avons calculé ces entropies locales dans la partie constituante de la connectivité structurelle de nos réservoirs à base d’état chimère (Fig. 5.8). Dans la figure 5.8.b, nous représentons le profil local du stockage d’information **AIS**. Comme prévu, les entités dominantes de stockage d’information sont les domaines triangulaires ($\mathbf{AIS} > 0$). Les valeurs négatives de **AIS** mettent en lumière l’existence des ‘Walls’ qui séparent les domaines (l’arrêt du stockage) et elles tracent ainsi les frontières de chaque ‘cluster’. Les valeurs nulles des **AIS** concernent nécessairement la partie cohérente de nos états chimères, car le processus est déterministe (sans aucune variation). Dans les figures (Fig. 5.8(c) et

(d)), nous avons calculé les entropies locales de transfert d'information **TE** dans les deux directions. Les deux profils (Fig. 5.8(c) et (d)) confirment que les entités dominantes en transfert d'information sont les 'Walls' qui séparent les domaines triangulaires ($\mathbf{TE} > 0$). Les valeurs négatives de **TE** retracent les clusters assurant le stockage d'information. Les valeurs nulles de **TE** montrent l'absence d'un flux directionnel de l'information et concernent majoritairement la partie cohérente de nos états chimères.

Ces profils des entropies locales permettent de filtrer les rôles des clusters dans la propagation et la répartition de l'information dans nos réservoirs. En effet, les résultats de la figure 5.8 s'alignent avec nos explications rapportées dans la section précédente(5.2.3). D'une part, le stockage actif d'information **AIS** est directement lié à la structure regroupée (clusters) localement [348], en particulier les motifs de boucles de rétroaction et d'anticipation [349]. Par conséquent, la connectivité locale offerte par les domaines triangulaires favorise le stockage d'information au sein des clusters. D'autre part, **TE** est attribuée aux liens longs qui semblent être un facilitateur crucial du transfert d'information à travers le réseau [348]. Ainsi, les 'Walls' des **sierpinski** jouent également le rôle des connexions inter-clusters qui favorisent une connectivité globale.

Maintenant, dans l'optique d'approfondir notre compréhension des résultats de cette section, nous avons étudié la variation de la quantité du flux d'information dans nos réservoirs en fonction de la taille des états chimères δ_L .

Dans la figure 5.9, nous avons tracé la **somme** des **AIS** et des **TE** locales pour chaque taille δ_L (Fig. 5.9(a) et (b), respectivement). Nous observons un comportement général de mise à l'échelle pour les deux grandeurs. En effet, les opérations basiques de stockage et transfert d'information se multiplient en augmentant la taille δ_L . Évidemment, cette mise à l'échelle n'est qu'une signature de l'augmentation du nombre de patterns **sierpinski** présents dans la partie incohérente δ_L . Par conséquent, nous concluons que la forme géométrique des 'clusters' qui constituent la connectivité structurelle de nos réservoirs, joue un rôle fondamental dans l'optimisation du flux d'information et la montée en performances. Remarquablement, la théorie d'information à travers le calcul de **AIS** et **TE**, nous a permis de filtrer les sentiers ainsi que les chemins adoptés par l'information dans nos réservoirs à base d'état chimère. Autrement dit, nous avons mis en évidence les mécanismes responsables de la réconciliation requise entre la propriété d'écho, la stabilité et la richesse dans la dynamique de nos réservoirs. Ces conclusions forment un pas significatif vers la compréhension de l'interaction énigmatique entre la partie structurelle et dynamique du reservoir computing.

De manière surprenante, la variation de **TE** (Fig. 5.9b) présente un minimum 'local' à

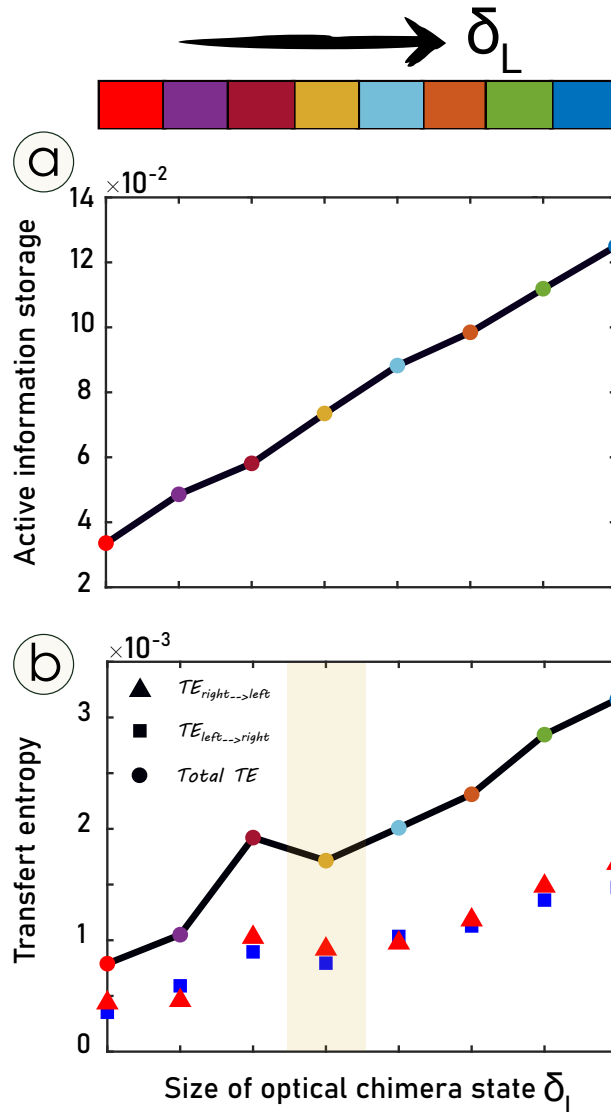


FIGURE 5.9 – La mise à l'échelle du traitement de l'information en fonction de δ_L . (a) : la somme des **AIS** locaux pour chaque taille δ_L . (b) : la somme des **TE** locaux pour chaque taille δ_L représentée en cercle. En marqueur triangle rouge, on illustre la somme des **TE** de droite à gauche. En marqueur carré bleu, on représente la somme des **TE** de gauche à droite.

partir de la taille δ_L illustrée par la couleur jaune. De manière surprenante, il s'agit de la taille minimale qui annonce la saturation de l'apprentissage (Fig. 5.7). Il s'avère judicieux d'investiguer cette observation critique qui est omniprésente et dans le volet structurel (connectivité) et dans le volet fonctionnel (performance).

Les propriétés physiques Les approches systémiques de la théorie de la complexité reposent sur la preuve que certains des phénomènes les plus fascinants des systèmes vivants (tels que la mémoire et la capacité à résoudre des problèmes) sont des phénomènes collectifs, issus des interactions de nombreuses unités de base, et pourraient ne pas être réductibles à la compréhension de composants élémentaires sur une base individuelle. Les théoriciens ont longtemps été séduits par l'idée d'adapter les concepts et méthodes issus de la mécanique statistique pour éclairer l'organisation à grande échelle des systèmes complexes biologiques [350-354]. L'une des conséquences les plus frappantes des interactions entre constituants élémentaires de la matière (atomes, molécules, électrons...) est l'émergence de phases diverses dont le comportement n'a que peu de ressemblance avec celui de leurs constituants de base ou de petits groupes d'entre eux [351, 355, 356]. Les systèmes constitués de très nombreux composants (microscopiques) peuvent présenter des types assez divers de comportements collectifs (macroscopiques), c'est-à-dire des phases, avec différents niveaux d'ordre interne. De plus, de légers changements dans les conditions externes (par exemple température, pression...) ou dans la force des interactions peuvent induire des réarrangements structurels spectaculaires, c'est-à-dire des **transitions de phase**.

On a supposé que, dans certaines circonstances, les systèmes vivants pourraient tirer d'importants avantages fonctionnels en opérant juste à la frontière entre les phases ordonnées et désordonnées, c'est-à-dire au bord même d'une transition de phase (continue) ou point critique [123, 357-360]. Par exemple, de manière plutôt générique, les systèmes vivants doivent parvenir à un compromis entre robustesse (Résilience de l'état du système aux perturbations externes ; qui est une propriété d'une phase ordonnée) et flexibilité (réactivité aux stimuli environnementaux, qui est une caractéristique des phases désordonnées). Un équilibre optimal entre ces deux tendances contradictoires peut être atteint en maintenant l'état du système à la limite d'une transition de phase ordre-désordre, c'est-à-dire à la **criticalité**.

Crutchfield, Langton et Packard [7, 8, 110] sont les leaders incontestés dans l'édification de l'hypothèse du **edge of chaos**. Ils ont montré que les réseaux complexes fonctionnant dans le régime critique peuvent avoir des capacités de calcul exceptionnellement élevées. En effet, cette hypothèse stipule que le point critique confère aux systèmes complexes distribués, un compromis optimal entre stockage et transmission d'informations, deux des ingrédients clés proposés par Turing comme indispensables aux machines informatiques universelles [361]. Les précurseurs visionnaires de l'hypothèse (edge of chaos) ont observé la maximisation de la complexité des réponses des systèmes situés à la criticalité. Cette propriété est traduite par la génération spontanée et riche des patterns spatio-temporels

qui sont des outils sophistiqués pour garantir un traitement optimal d'information.

Jusqu'à présent, les preuves disponibles soutiennent systématiquement la conjecture selon laquelle le cerveau est au bord du chaos, ou presque [11-14]. En effet, la criticalité est très prometteuse pour les neurosciences afin de comprendre la balance requise entre ségrégation/intégration de l'information [362, 363]. Cet équilibre équivalent à celui du stockage/transfert [12, 364] justifie le pouvoir de flexibilité et d'adaptation du cerveau aux tâches cognitives demandées. Dans le cadre du reservoir computing, une série d'ouvrages fondateurs ont montré que de telles machines peuvent effectuer des calculs en temps réel – en répondant rapidement aux signaux d'entrée variables dans le temps – de manière cohérente mais flexible s'ils opèrent à proximité d'un point critique [19, 21, 22, 25, 365].

Revenons à nos états chimères optiques, nous pouvons leur attribuer cet équilibre entre ordre-désordre en regardant les caractéristiques de la partie cohérente stable et de la partie incohérente intermittente. Comme nous l'avons expliqué tout au long de ce travail, la complexité de nos états chimères est renforcée en augmentant la taille de la partie incohérente δ_L , grâce à la présence des patterns SIERPINSKI. Dans la section précédente, nous avons mis en évidence le rôle computationnel remarquable des patterns triangulaires émergents en calculant le stockage actif d'information **AIS** et l'entropie de transfert **TE**. Nous avons constaté un comportement de mise à l'échelle des grandeurs calculées en augmentant la taille δ_L (Fig. 5.9). Il s'agit d'une conséquence directe du renforcement de la complexité qui accompagne l'augmentation du nombre des patterns triangulaires. Ce résultat, concernant le flux d'information dans nos réservoirs, est en accord avec le concept central de l'hypothèse du bord du chaos qui connecte la complexité aux capacités computationnelles [7, 205, 366].

Maintenant, nous nous intéressons à une caractéristique importante susceptible d'être exploitée par les systèmes vivants qui se situent dans la phase critique, afin d'induire un comportement coordonné des unités individuelles dans l'espace et dans le temps. Il s'agit de l'émergence spontanée d'un comportement coordonné collectif auto-organisé. **L'auto-organisation** est un phénomène impressionnant, observé dans une variété de systèmes complexes naturels ou artificiels, dans lesquels des structures ou des dynamiques globales cohérentes émergent d'interactions locales entre des parties individuelles. Si le degré d'intégration collective d'un système ne dépend pas de sa taille, son niveau de robustesse et d'adaptabilité est généralement augmenté et nous l'appelons **invariant à l'échelle**. En tant que signature mathématique de l'invariance d'échelle, les lois de puissance sont omniprésentes dans les systèmes critiques (la physique [367], la biologie [368], l'économie [369]). Les neurosciences ont prouvé que l'invariance d'échelle a de nombreux effets surprenants qui

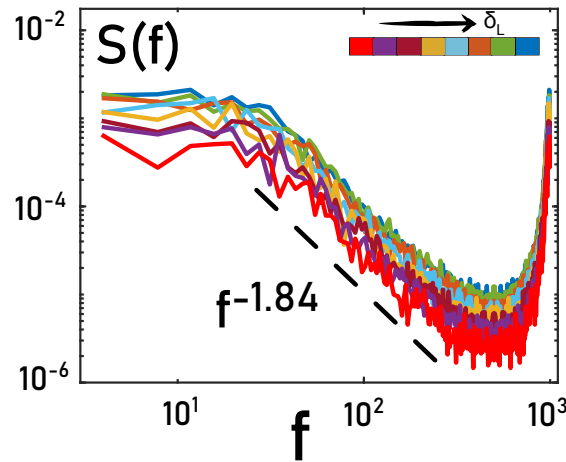


FIGURE 5.10 – Analyse spectrale de nos états chimères optiques pour les différentes δ_L . La ligne en pointillés représente l’ajustement log-log du spectre de puissance dans la gamme des basses et moyennes fréquences montrant une loi de puissance.

sont pertinents pour le traitement de l’information. Plusieurs études ont lié la répartition uniforme et à différentes échelles des corrélations spatio-temporelles dans le cerveau humain à la création d’un équilibre optimal entre intégration et ségrégation des informations [11-15, 362, 363].

Dans la figure 5.10, nous avons calculé le spectre de puissance $S(f)$ des états chimères pour les différentes tailles de partie incohérente δ_L . Une décroissance spectrale suivant une loi de puissance vers les basses et moyennes fréquences, présente un point commun entre tous les spectres illustrés. Il s’agit de la propriété du bruit $1/f^\alpha$, qui est considérée comme l’une des empreintes de la complexité des systèmes proches de la criticalité. Nous avons vu dans notre premier travail [366] que l’existence des patterns auto-organisés **sierpinski** dans nos états chimères optiques, est responsable de l’apparition de ce phénomène du bruit $1/f^\alpha$. En effet, Ninagawa [207, 370] a conjecturé une relation étroite entre l’universalité computationnelle et le $1/f^\alpha$ illustré par les patterns présents dans les automates cellulaires. Dans le cadre de nos réservoirs, cette loi de puissance présentée par les clusters triangulaires semble être responsable des performances excellentes de nos états chimères dans l’inférence/prédiction des séries chaotiques.

Nos analyses montrent que l’auto-organisation et l’invariance d’échelle incarnées dans la répartition spatiale des clusters à travers le réseau, améliorent l’efficacité de la connectivité topologique du réservoir computing dans les tâches difficiles.

Cependant, ce type d’invariance d’échelle reste incapable de nous expliquer la saturation de l’apprentissage dans nos réservoirs à partir d’une certaine taille δ_L . Il s’avère judicieux

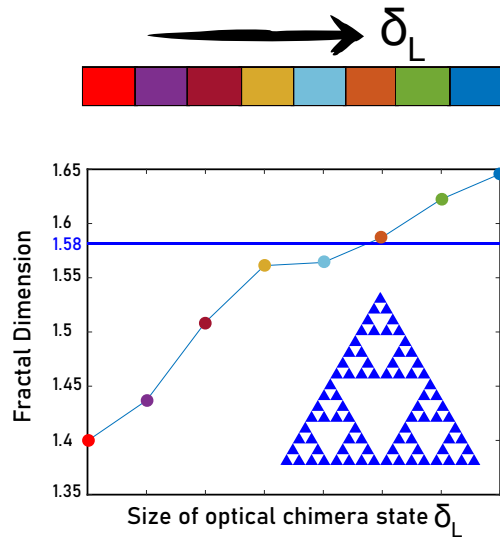


FIGURE 5.11 – La dimension fractale en fonction de δ_L . La ligne en bleu représente la valeur exacte du célèbre **sierpinski** géométrique.

de creuser notre analyse vers la forme géométrique de la notion d'invariance d'échelle, car la saturation des performances de nos réservoirs semble avoir une relation étroite avec l'achèvement éventuel d'un nombre seuil de clusters et donc de triangles **sierpinski**. Il s'agit de la notion d'**auto-similarité** statistique qui est attribuée aux objets physiques fractals. Les structures invariantes à l'échelle ont été largement étudiées dans différents domaines scientifiques depuis plus de 50 ans. Les origines remontent à une publication phare de Mandelbrot, où il a inventé le terme 'auto-similarité statistique' en observant que pour certains objets "chaque [...] portion peut – au sens statistique – être considérée comme une image à échelle réduite de l'ensemble" [371]. Mandelbrot a utilisé la *dimension Hausdorff-Besicovitch* afin de mesurer la façon dont la géométrie d'un objet se comporte lors du redimensionnement.

Dans la figure 5.11, nous avons calculé les dimensions fractales des états chimères pour les différentes tailles δ_L . Plus de détails théoriques sur la méthode de calcul sont rapportés dans l'annexe A. Nous remarquons que nos calculs de la dimension fractale des états chimères montrent une augmentation d'auto-similarité au sein de nos états chimères en fonction de δ_L . Ce résultat confirme le rôle fondamentale de la propriété d'invariance d'échelle dans les performances exceptionnelles et la robustesse de nos réservoirs à base d'états chimères. En réalité, le traitement fractal de l'information permet d'extraire des caractéristiques complexes et non-linéaires à différentes échelles cachées dans les données. D'ailleurs, des chercheurs de Google et Stanford ont montré que l'auto-similarité renforce la profondeur des

réseaux de neurones ainsi que leurs expressivités [372]. D'autre part, la nature récursive des fractals contribue à l'amélioration de la récurrence de nos réservoirs en leur permettant de détecter des propriétés fractales présentes dans les séries chaotiques entraînées. Par ailleurs, Jaeger [298] a suggéré que N soit dans la plage ($T/10 \preceq N \preceq T/2$) avec T comme longueur des données d'entraînement (dans notre cas d'étude $T = 8446$). La topologie auto-similaire de nos états chimères a fourni une optimisation de la taille du réservoir et une économie irréprochable des ressources de calcul ($N = 256$ neurones étaient suffisants pour un bon apprentissage). En conclusion, la propriété d'auto-similarité garantit à nos réservoirs à base d'état chimère, le compromis optimal entre la flexibilité/la robustesse requises pour améliorer les capacités computationnelles des systèmes critiques.

La fractalité est une propriété omniprésente dans la nature et l'architecture des systèmes du vivant. Elle reste le choix géométrique optimal pour un transport efficace d'information (les éléments nutritifs des racines vers les branches de l'arbre [373], le sang du cœur vers les organes à travers le réseau des vaisseaux sanguins [374], l'oxygène dans le réseau de bronchioles des poumons [375]...). L'auto-similarité dans les systèmes biologiques comme les poumons ou les réseaux capillaires sanguins, augmente la surface efficace d'échange d'information dans une surface limitée tout en minimisant l'énergie (les poumons humains assurent $150m^2$ des échanges respiratoires grâce à leur géométrie fractale [376]). Finalement, la fractalité consolide la résilience face aux perturbations en adoptant des voies saines de transport, ce qui justifie le maintien des fonctionnalités respiratoires ou de la circulation sanguine dans certains cas d'infection et d'inflammation [377].

Concernant les neurosciences, les chercheurs ciblent la compréhension de la pertinence de la fractalité et de l'invariance d'échelle dans le cerveau. D'un point de vue structurel, l'arborisation fractale se manifeste à la fois au niveau microscopique (dendrites et axones) et au niveau macroscopique (les fibres de connexion à longue portée inter-zones.). Les études [212, 213, 330, 378-380] ont montré que l'auto-similarité offre un équilibre entre un coût de câblage minimal et une maximisation de transfert d'information. Elles ont aussi montré [379, 381] une perte d'auto-similarité dans le connectome structurel des patients présentant des troubles de la conscience et des fonctions cognitives.

Jusqu'à maintenant, nous avons décortiqué les relations structure/dynamique et structure/fonction au sein de nos réservoirs à base d'état chimère. Nous avons exploré les propriétés physiques fondamentales offertes (auto-organisation, invariance d'échelle et auto-similarité) par les patterns triangulaires et avons fait le lien avec le flux optimum d'information au sein de nos réservoirs. Cependant, l'énigme autour de la saturation de l'apprentissage à partir d'une certaine taille δ_L persiste et laisse la résolution du puzzle incomplète.

En revenant à la figure 5.11, nous remarquons que la dimension fractale confirme davantage les constatations issues des calculs de l'entropie de transfert (Fig. 5.9) et celles des performances en fonction du rayon spectral (Fig. 5.7). Une stagnation des valeurs autour de 1.57 est partagée par la taille en question (en jaune) et celle d'après (en cyan), montrant un 'minimum local' qui indique une transition subie par l'auto-similarité de nos états chimères. Il s'agit des tailles critiques dont les auto-similarités statistiques s'approchent de l'auto-similarité exacte du triangle de **sierpinski** (Fig. 5.11). Ce résultat est en très bon accord avec les conclusions de notre premier travail motivé par la présence des patterns triangulaires [366], où nous avons montré que nos états chimères optiques partagent un ensemble de propriétés computationnelles avec la classe complexe des automates cellulaires. En effet, WOLFRAM [382] a connecté la maximisation des capacités computationnelles dans les automates cellulaires avec une dimension fractale dans l'intervalle [1.58, 1.69]. Nous constatons que la saturation de l'apprentissage correspond aux cas où la dimension fractale des états chimères est quasiment équivalente à celle des automates cellulaires.

La réplication de l'attracteur chaotique de Lorenz

Contrairement aux sections précédentes dans lesquelles nous utilisons nos réservoirs à base d'état chimère pour prédire/inférer certains composants du système chaotique de LORENZ, nous essayons ici de faire en sorte que le réservoir génère l'attracteur par lui-même, sans entraînement externe. Bien que nos réservoirs entraînés puissent facilement produire des composants de l'attracteur chaotique de LORENZ qui lui sont alimentés, il n'est pas clair s'il 'apprennent' réellement l'attracteur. Autrement dit, nous allons examiner nos réservoirs à base d'état chimère afin de vérifier s'ils sont capables de maintenir l'évolution de la trajectoire sur un attracteur chaotique.

La dynamique du réservoir dans le cas du mode génératif est donnée par :

$$\mathbf{r}(n+1) = (1 - \alpha)\mathbf{r}(n) + \alpha \tanh(\mathbf{W}_{in}[bias; \mathbf{u}(n+1)] + \mathbf{W}\mathbf{r}(n) + \mathbf{W}_{fb}\mathbf{y}(n) + \nu), \quad (5.5)$$

où le terme de la matrice de connexion du feedback \mathbf{W}_{fb} est ajouté. Les paramètres du réseau sont donnés dans le tableau 5.3. Le réservoir a été entraîné à l'attracteur de LORENZ en utilisant le forçage sans entrée (c'est-à-dire simplement avec le forçage de sortie).

Dans la figure 5.12, nous montrons l'attracteur exact de LORENZ (en bleu) ainsi que ses composants individuels, et les mêmes informations générées par notre réservoir à base d'état chimère (en orange). Nous remarquons que notre réservoir génère des prédictions correctes à court terme (la partie ombrée en vert), puis s'écartent des trajectoires réelles de LORENZ,

TABLEAU 5.3 – Les hyper-paramètres utilisés dans notre réservoir à base d'état chimère pour générer l'attracteur de LORENZ.

Paramètre	Valeur
Taille du réservoir (N)	256
Rayon spectral de la matrice W (ρ)	10
Bruit (ν)	0.0001
Taux de fuite (α)	0.1
Input scaling (\mathbf{W}_{in})	1
feed-back scaling (\mathbf{W}_{fb})	1
Bias	1
Durée de l'entraînement	6000
Méthode de régression	Ridge
Paramètre de régularisation (β)	$0.5 \cdot 10^{-8}$

ce qui est attendu puisque toute petite erreur augmente de façon exponentielle en raison de la dynamique chaotique du système de Lorenz. Il semble qu'après un certain temps, bien que la prédiction de notre réservoir s'écarte de la trajectoire réelle, la dynamique à long terme semble ressembler à celle du système Lorenz original. (Fig. 5.12a).

L'objectif principal de la prédiction de séries temporelles chaotiques est consacré à l'augmentation de la fenêtre de prédiction. Cette dernière est mesurée par l'exposant de LYAPUNOV maximum qui présente alors une référence pour indiquer un comportement chaotique et pour calculer le temps de LYAPUNOV. Dans le cas du système de LORENZ (avec les paramètres choisis dans notre étude), l'exposant maximum de LYAPUNOV est $\lambda_{max} = 0.89$ [383] et donc le temps de LYAPUNOV est donné par l'inverse de λ_{max} . Dans la figure (5.12.b), nous remarquons que notre réservoir imite parfaitement jusqu'à 7 temps de LYAPUNOV. Pour les réservoirs classiques [214, 384], la fenêtre de prédiction ne peut atteindre que 5 à 6 fois le temps de LYAPUNOV. Nos réservoirs ne nécessitent aucun ensemble d'optimisation en terme d'hyper-paramètres. D'ailleurs, le résultat illustré dans la figure 5.12 est optimal bien que nous ayons pris un rayon spectral très grand qui est égale à 10. D'autre part, nos réservoirs montrent une robustesse statistique au niveau des performances à l'inverse des réservoirs classiques dont les calculs divergent dans la plupart du temps.

En conclusion, nos réservoirs à base d'état chimère réussissent, mieux par rapport aux réservoirs classiques, dans l'apprentissage et la réplique totale de l'attracteur chaotique de LORENZ. Nos réservoirs offrent une large fenêtre de prédiction tout en évitant une optimisation fastidieuse d'hyper-paramètre. Ce nouveau challenge montre l'efficacité et la pertinence de notre analyse structurelle, dynamique et fonctionnelle que nous avons établies dans la section précédente. En d'autres termes, ces résultats confirment l'importance

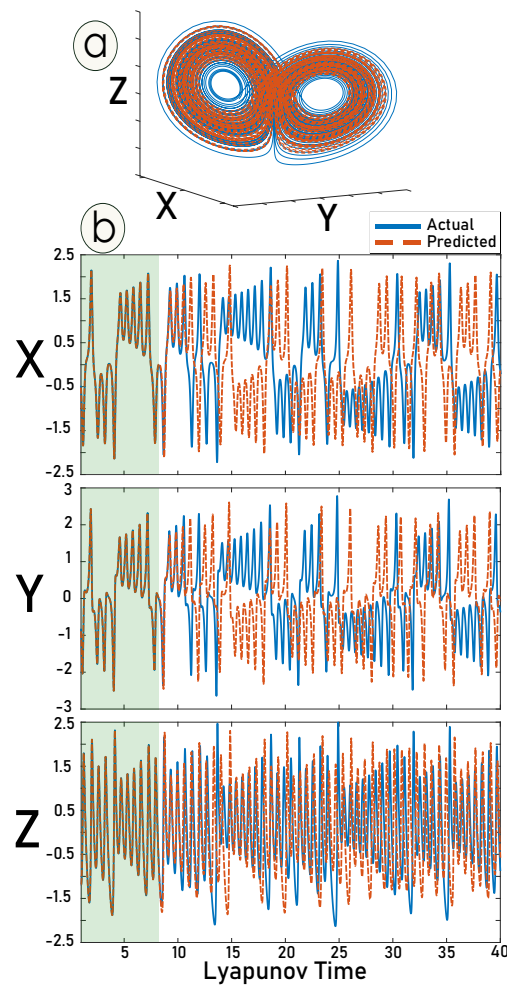


FIGURE 5.12 – La réplique totale de l'attracteur de LORENZ. (a) : Les "ailes de papillon" caractéristiques de l'attracteur de LORENZ en 3 dimensions (en bleu : exact, en orange : généré par nos réservoirs). (b) : Ensembles de données réels (bleu) et prédits par nos réservoirs pendant la phase de prévision.

fondamentale du traitement fractale de l'information, notamment, dans le cas des séries complexes chaotiques.

5.3 La prédiction du système de Rössler

5.3.1 La réplique de l'attracteur chaotique de Rössler

Les équations de RÖSSLER ont été décrites pour la première fois dans les années 1970 par OTTO RÖSSLER comme un système explicitement conçu pour produire un chaos temporel continu. Bien que RÖSSLER ait écrit plusieurs systèmes capables de créer un chaos continu,

le plus le système couramment utilisé est celui donné par [385] :

$$\begin{aligned}\dot{x} &= -y - z, \\ \dot{y} &= x + ay, \\ \dot{z} &= b + z(x - c).\end{aligned}\tag{5.6}$$

Les paramètres a , b , c déterminent si le système évolue périodiquement, de manière chaotique ou converge vers une solution statique. Encore une fois, \dot{x} désigne la dérivée temporelle de x . Nous avons utilisé les valeurs des paramètres $a = 0,15$, $b = 2,0$ et $c = 10$, ce qui installe une évolution chaotique du système [385]. Nous avons utilisé nos réservoirs à base d'état chimère dans le but de reconstruire, autonomement, l'attracteur chaotique de RÖSSLER.

Dans la figure 5.13, nous montrons l'attracteur exacte de RÖSSLER (en bleu) ainsi que ses composants individuels, et les mêmes informations générées par notre réservoir à base d'état chimère (en orange). Notre réservoir génère des prédictions correctes à court terme (la partie ombrée en vert), puis s'écartent des trajectoires réelles de Rössler, ce qui est attendu puisque toute petite erreur augmente de façon exponentielle en raison de la dynamique chaotique du système de RÖSSLER. Nous constatons qu'après un certain temps, bien que la prédiction de notre réservoir s'écarte de la trajectoire réelle, la dynamique à long terme reproduit celle du système RÖSSLER original (Fig. 5.13a).

Dans le cas du système de RÖSSLER, l'exposant maximum de LYAPUNOV est $\lambda_{max} = 0.09$ [383] et donc le temps de LYAPUNOV est donné par l'inverse de λ_{max} . Dans la sous-figure 5.13.b, nous remarquons que notre réservoir imite parfaitement jusqu'à 7 temps de LYAPUNOV. Nos réservoirs ne nécessitent aucune procédure d'optimisation en terme d'hyper-paramètres. D'ailleurs, le résultat illustré dans la figure 5.13 est optimal bien que nous ayons pris un rayon spectral très grand qui égale à 10. D'autre part, nos réservoirs montrent une robustesse statistique remarquable au niveau des performances à l'inverse des réservoirs classiques dont les calculs divergent dans la plupart du temps. Cependant, nous avons remarqué une légère sensibilité à la variable z dont la dynamique est régie par la troisième équation du système de RÖSSLER, qui contient le terme non-linéaire et qui induit la réinjection au début du mouvement de spirale.

En conclusion, nos réservoirs à base d'état chimère réussissent encore une fois, dans l'apprentissage et dans la réplique totale d'un autre attracteur chaotique. Ce nouveau résultat confirme l'importance fondamentale du traitement auto-similaire de l'information, notamment, dans le cas des séries complexes chaotiques.

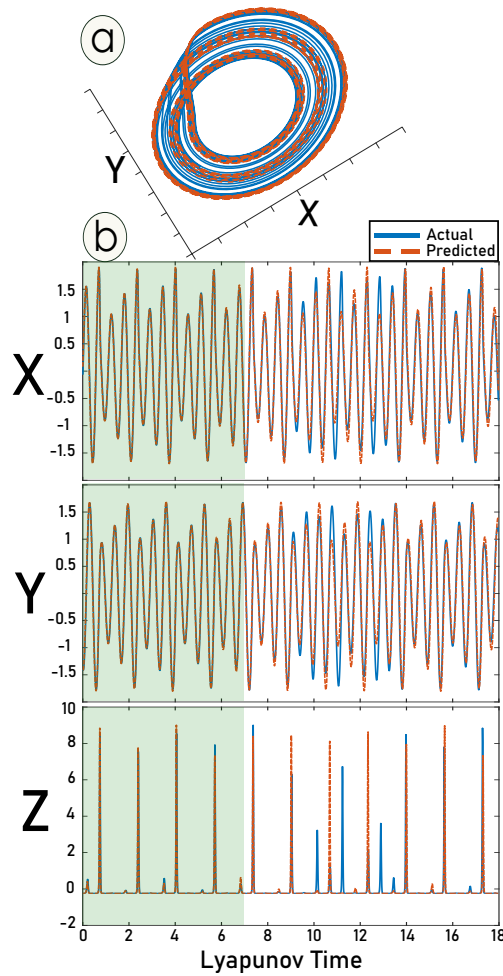


FIGURE 5.13 – La réplique totale de l’attracteur de RÖSSLER. (a) : La forme pliée caractéristique de l’attracteur de RÖSSLER en 2 dimensions(en bleu : exact, en orange : généré par nos réservoirs). (b) : Ensembles de données réels (bleu) et prédits par nos réservoirs pendant la phase de prévision.

5.3.2 Inférence multi-pas en avant d’une variable non mesurée de l’attracteur de Rössler

Afin de vérifier nos résultats sur le système de LORENZ, nous avons établi une étude statistique des performances de nos réservoirs dans l’inférence d’une variable non mesurée du système de RÖSSLER en fonction de la taille de la partie incohérente δ_L . En suivant la même procédure suivie dans la section(5.2.3), nous avons collecté les résultats illustrés dans la figure 5.14.

Nos conclusions restent toujours valables concernant la montée en performances grâce à l’augmentation du nombre des patterns SIERPINSKI. En réalité, ces résultats confirment

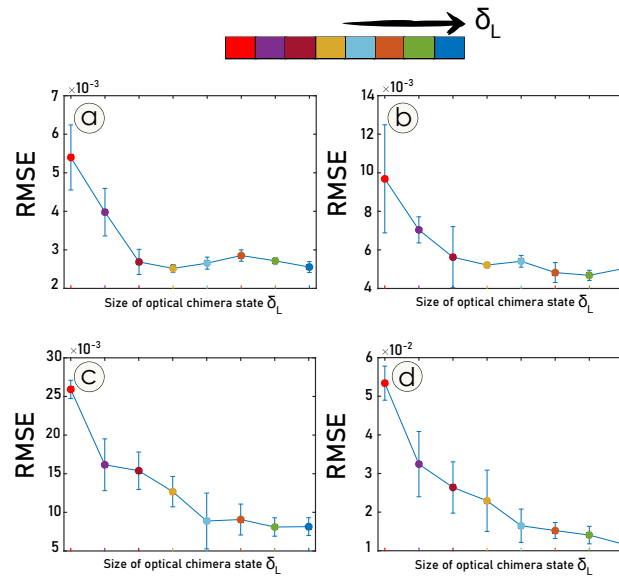


FIGURE 5.14 – L’erreur en fonction de la taille δ_L . (a) : $n_{sh}=5$, (b) : $n_{sh}=10$, (c) : $n_{sh}=15$ et (d) : $n_{sh}=20$. les données ont été intégrés avec $dt = 0.1$

davantage le rôle crucial de la propriété d’auto-similarité dans le traitement des séries chaotiques. En d’autres termes, la répartition des clusters triangulaires dans la connectivité garantit la réconciliation entre la propriété d’écho, la stabilité et la propriété de séparation nécessaires pour une dynamique riche et expressive des données apprises.

5.4 Conclusion

Nos réservoirs à base d’état chimère ont montré des performances élevées dans différents challenges d’apprentissage des attracteurs chaotiques (inférence, prédiction et réplique totale). Notre analyse fonctionnelle/dynamique a mis en lumière l’importance de la connectivité en clusters (patterns de SIERPINSKI) dans l’élargissement de la portée de la propriété d’état d’écho, la stabilité ainsi que la richesse de la dynamique interne de nos réservoirs. L’originalité de notre type de topologie proposée réside dans la quasi-indépendance des performances vis à vis des hyper-paramètres les plus critiques tels que le rayon spectral, le paramètre de régularisation et la taille du réseau. D’autre part, notre analyse structurelle/dynamique a montré le rôle crucial de la géométrie triangulaire des clusters dans l’optimisation du flux d’information au sein de nos réservoirs à travers une maximisation des opérations du stockage/transfert de l’information traitée. En outre, cette analyse nous a permis de dévoiler l’importance des propriétés d’auto-organisation, d’invariance d’échelle et d’auto-similarité dans la connectivité structurelle du reservoir computing. Le traitement

fractal de l'information a garanti une consolidation de l'expressivité/la robustesse à nos réservoirs à base d'état chimère. Finalement, cet ensemble de propriétés critiques de nos états chimères optiques offre à nos réservoirs une architecture/dynamique optimale pour accomplir des tâches complexes d'apprentissage.

Conclusion générale

Les travaux de recherche menés au cours de cette thèse ont porté sur l'étude des capacités computationnelles des états chimères optiques. En effet, l'auto-organisation spontanée et les comportements collectifs complexes qui distinguent la dynamique spatio-temporelle de nos états chimères optiques, ont motivé notre cheminement à travers les différents domaines scientifiques intéressés par le calcul distribué.

Tout d'abord, nous avons examiné la capacité des états chimères optiques à imiter les possibilités calculatoires remarquables des automates cellulaires élémentaires. Pour ce faire, nous avons mis en évidence une analogie qualitative et quantitative entre les dynamiques supportées par les deux systèmes concernés. Sur le volet qualitatif, nous avons adopté des notions de la mécanique computationnelle afin de comprendre la nature du calcul distribué. Cette approche stipule que la formation des patterns auto-organisés dans la dynamique spatio-temporelle d'un automate cellulaire constitue le mécanisme crucial de leur calculabilité émergente. Également, nous avons emprunté le filtrage local de la dynamique d'information suggéré par LIZIER [115, 188, 189], qui permet de décortiquer le rôle des patterns émergents en termes des trois opérations fondamentales acquises pour traiter l'information (stockage-transfert-modification). Nos résultats s'alignent avec ceux établis sur les automates cellulaires concernant les rôles computationnels des patterns émergents. En effet, cette analogie qualitative était confirmée grâce à la ressemblance phénotypique et fonctionnelle des *domaines réguliers* de type SIERPINSKI et des *particules embarquées* (walls des domaines) dans nos états chimères optiques, aux *blinkers* et aux *gliders* dans les automates cellulaires élémentaires. Sur le volet quantitatif, nous avons adopté, d'abord, l'approche statistique de LANGTON [7], qui connecte le comportement complexe des automates cellulaires élémentaires à l'amélioration des capacités calculatoires. Notre méthode de clustering basée sur des métriques entropiques, a prouvé son efficacité dans la classification des automates cellulaires tout en positionnant le comportement complexe dans la région critique située entre l'ordre et le désordre. Nous avons remarqué que la complexité des dynamiques de nos états chimères optiques appartient au nuage de la classe complexe d'automates

cellulaires. L'occupation de cette classe complexe d'automate cellulaire d'une position centrale dans le cadre de l'hypothèse du "bord du chaos" et dans la théorie de l'information, a motivé un ensemble de comparaisons minutieuses. Nos résultats montrent que nos états chimères optiques partagent un ensemble de propriétés calculatoires sophistiquées avec la classe complexe des automates cellulaires élémentaires [366].

La suite logique de ce premier travail est d'explorer ces capacités computationnelles dans le cadre d'un dispositif dédié au traitement d'information. Par conséquent, nous avons choisi une approche paradigmatique basée sur des réseaux de neurones récurrents connue sous le nom de *Reservoir Computing*. Il s'agit d'un carrefour interdisciplinaire qui réunit la théorie d'information, la théorie du chaos et des systèmes complexes, la dynamique non-linéaire et les neurosciences. L'implémentation de nos états chimères optiques dans l'architecture du reservoir computing, nous a permis d'examiner les conclusions du premier travail. Nos résultats montrent que nos réservoirs à base d'état chimère assurent de bonnes performances dans différents challenges concernant la prédiction et la réplication des attracteurs chaotiques. Notre analyse fonctionnelle/dynamique a révélé l'importance de la connectivité en clusters (patterns de SIEPRINSKI) dans l'élargissement de la portée de la propriété d'état d'écho, la stabilité ainsi que la richesse de la dynamique interne de nos réservoirs. En fait, ce qui rend notre topologie unique, c'est sa capacité à maintenir des performances presque indépendantes des paramètres influents tels que le rayon spectral, le paramètre de régularisation et la taille du réseau tout en évitant des optimisations laborieuses. Par ailleurs, notre analyse structurelle a révélé l'importance cruciale de la géométrie triangulaire des clusters dans l'optimisation du flux d'information à l'intérieur de nos réservoirs, ce qui maximise les opérations de stockage et de transfert de l'information traitée. De plus, cette analyse a mis en évidence l'importance des propriétés d'auto-organisation, d'invariance d'échelle et d'auto-similarité dans la connectivité structurelle du reservoir computing. En réalité, l'approche fractale du traitement de l'information renforce l'expressivité et la robustesse de nos réservoirs à base d'états chimères. En conclusion, cet ensemble de propriétés critiques présentes dans nos états chimères optiques confère à nos réservoirs une architecture et une dynamique optimales pour exécuter des tâches d'apprentissage complexes.

À notre connaissance, il s'agit de l'une des premières études suggérant que l'état chimère peut offrir des propriétés d'objet computationnel. L'originalité de ce travail prend son ampleur de l'approche multidisciplinaire adoptée dans le but d'explorer le rôle computationnel des patterns auto-organisés. En effet, notre travail offre une nouvelle interface entre les neurosciences, les réseaux de neurones artificiels et la physique des systèmes complexes, ouvrant la voie à de futures recherches hybrides d'exploration prometteuse menant

à une compréhension approfondie de la relation énigmatique fonction/structure dans le cerveau biologique. Une autre direction pourrait être adoptée, concerne l'implémentation de l'auto-organisation et l'invariance d'échelle dans la connectivité structurelle des neuromorphiques optiques [386]. Il s'agit d'un domaine en plein essor qui vise à s'inspirer du cerveau pour concevoir du matériel informatique efficace utilisant un substrat photonique. L'intérêt pour ce domaine vient du gain potentiel en rapidité et en parallélisme dû à l'utilisation de technologies optiques tout en réduisant la consommation énergétique.

D'autre part, les conclusions de cette thèse établissent les fondements pour des recherches futures, en mettant particulièrement l'accent sur le rôle d'auto-organisation et de l'auto-similarité dans l'optimization des performances du reservoir computing. En effet, on peut envisager le challenge de nos réservoirs à base d'état chimère dans d'autres tâches complexes comme la prédiction des dynamiques spatio-temporelles (des données 2D), la classification, la reconnaissance du langage naturel.

Bibliographie

- [1] Bruce EDMONDS. « What is Complexity ?-The philosophy of complexity per se with application to some examples in evolution ». In : *The evolution of complexity*. Kluwer, Dordrecht, 1995.
- [2] John H HOLLAND. « Studying complex adaptive systems ». In : *Journal of systems science and complexity* 19 (2006), p. 1-8.
- [3] Simon LEVIN. « Complex adaptive systems : exploring the known, the unknown and the unknowable ». In : *Bulletin of the American mathematical Society* 40.1 (2003), p. 3-19.
- [4] Ilya PRIGOGINE et Gregoire NICOLIS. « Self-organisation in nonequilibrium systems : towards a dynamics of complexity ». In : *Bifurcation analysis* (1985), p. 3-12.
- [5] Carlos GERSHENSON et Francis HEYLIGHEN. « How can we think the complex ». In : *Managing organizational complexity : philosophy, theory and application* 3 (2005), p. 47-62.
- [6] Mitchell M WALDROP. *Complexity : The emerging science at the edge of order and chaos*. Simon et Schuster, 1993.
- [7] Chris G LANGTON. « Computation at the edge of chaos : Phase transitions and emergent computation ». In : *Physica D : nonlinear phenomena* 42.1-3 (1990), p. 12-37.
- [8] Norman H PACKARD. « Adaptation toward the edge of chaos ». In : *Dynamic patterns in complex systems* 212 (1988), p. 293-301.
- [9] Stephen WOLFRAM. « Theory and applications of cellular automata ». In : *World Scientific* (1986).
- [10] John M BEGGS. *The cortex and the critical point : Understanding the power of emergence*. MIT Press, 2022.
- [11] John M BEGGS et Nicholas TIMME. « Being critical of criticality in the brain ». In : *Frontiers in physiology* 3 (2012), p. 163.
- [12] Jordan O'BYRNE et Karim JERBI. « How critical is brain criticality ? » In : *Trends in Neurosciences* (2022).
- [13] Dante R CHIALVO. « Critical brain networks ». In : *Physica A : Statistical Mechanics and its Applications* 340.4 (2004), p. 756-765.

- [14] Vincent ZIMMERN. « Why brain criticality is clinically relevant : a scoping review ». In : *Frontiers in neural circuits* 14 (2020), p. 54.
- [15] George F GROSU, Alexander V HOPP, Vasile V MOCA, Harald BÂRZAN, Andrei CIUPARU, Maria ERCSEY-RAVASZ, Mathias WINKEL, Helmut LINDE et Raul C MUREȘAN. « The fractal brain : scale-invariance in structure and dynamics ». In : *Cerebral Cortex* 33.8 (2023), p. 4574-4605.
- [16] Guillermo B MORALES et Miguel A MUÑOZ. « Optimal input representation in neural systems at the edge of chaos ». In : *Biology* 10.8 (2021), p. 702.
- [17] Laura E SUÁREZ, Blake A RICHARDS, Guillaume LAJOIE et Bratislav MISIC. « Learning function from structure in neuromorphic networks ». In : *Nature Machine Intelligence* 3.9 (2021), p. 771-786.
- [18] Ling FENG, Lin ZHANG et Choy Heng LAI. « Optimal machine intelligence near the edge of chaos ». In : *APS March Meeting Abstracts*. T. 2021. 2021, B60-003.
- [19] Wolfgang MAASS, Thomas NATSCHLÄGER et Henry MARKRAM. « Real-time computing without stable states : A new framework for neural computation based on perturbations ». In : *Neural computation* 14.11 (2002), p. 2531-2560.
- [20] Herbert JAEGER et Harald HAAS. « Harnessing nonlinearity : Predicting chaotic systems and saving energy in wireless communication ». In : *science* 304.5667 (2004), p. 78-80.
- [21] Robert LEGENSTEIN et Wolfgang MAASS. « Edge of chaos and prediction of computational performance for neural circuit models ». In : *Neural networks* 20.3 (2007), p. 323-334.
- [22] Nils BERTSCHINGER et Thomas NATSCHLÄGER. « Real-time computation at the edge of chaos in recurrent neural networks ». In : *Neural computation* 16.7 (2004), p. 1413-1436.
- [23] David SUSSILLO et Larry F ABBOTT. « Generating coherent patterns of activity from chaotic neural networks ». In : *Neuron* 63.4 (2009), p. 544-557.
- [24] Taro TOYOIZUMI et LF ABBOTT. « Beyond the edge of chaos : Amplification and temporal integration by recurrent networks in the chaotic regime ». In : *Physical Review E* 84.5 (2011), p. 051908.
- [25] Joschka BOEDECKER, Oliver OBST, Joseph T LIZIER, N Michael MAYER et Minoru ASADA. « Information processing in echo state networks at the edge of chaos ». In : *Theory in Biosciences* 131 (2012), p. 205-213.
- [26] MG CLERC, MA FERRÉ, Saliya COULIBALY, RG ROJAS et Mustapha TLIDI. « Chimera-like states in an array of coupled-waveguide resonators ». In : *Optics letters* 42.15 (2017), p. 2906-2909.
- [27] Eberhard BODENSCHATZ, Werner PESCH et Guenter AHLERS. « Recent developments in Rayleigh-Bénard convection ». In : *Annual review of fluid mechanics* 32.1 (2000), p. 709-778.

- [28] T YAMAGUCHI, L KUHNERT, Zs NAGY-UNGVARAI, SC MÜLLER et B HESS. « Gel systems for the Belousov-Zhabotinskii reaction ». In : *The Journal of Physical Chemistry* 95.15 (1991), p. 5831-5837.
- [29] Philip BALL. « Forging patterns and making waves from biology to geology : a commentary on Turing (1952)'The chemical basis of morphogenesis' ». In : *Philosophical Transactions of the Royal Society B : Biological Sciences* 370.1666 (2015), p. 20140218.
- [30] Yoram LOUZOUN, Sorin SOLOMON, Henri ATLAN et Irun R COHEN. « Modeling complexity in biology ». In : *Physica A : Statistical Mechanics and its Applications* 297.1-2 (2001), p. 242-252.
- [31] Loet LEYDESDORFF. *A sociological theory of communication : The self-organization of the knowledge-based society*. Universal-Publishers, 2001.
- [32] Y KURAMOTO et D BATTOGTOKH. « Coexistence of Coherence and Incoherence in Nonlocally Coupled Phase Oscillators. » In : *NONLINEAR PHENOMENA IN COMPLEX SYSTEMS* 5.4 (2002), p. 380-385.
- [33] Daniel M ABRAMS et Steven H STROGATZ. « Chimera states for coupled oscillators ». In : *Physical review letters* 93.17 (2004), p. 174102.
- [34] Ming LI, Paul VITÁNYI et al. *An introduction to Kolmogorov complexity and its applications*. T. 3. Springer, 2008.
- [35] Edward N LORENZ. « Deterministic nonperiodic flow ». In : *Journal of atmospheric sciences* 20.2 (1963), p. 130-141.
- [36] Yoshiki KURAMOTO. « Chemical turbulence ». In : *Chemical oscillations, waves, and turbulence*. Springer, 1984, p. 111-140.
- [37] Michael G ROSENBLUM, Arkady S PIKOVSKY et Jürgen KURTHS. « Phase synchronization of chaotic oscillators ». In : *Physical review letters* 76.11 (1996), p. 1804.
- [38] Juan A ACEBRÓN, Luis L BONILLA, Conrad J Pérez VICENTE, Félix RITORT et Renato SPIGLER. « The Kuramoto model : A simple paradigm for synchronization phenomena ». In : *Reviews of modern physics* 77.1 (2005), p. 137.
- [39] Hugues CHATÉ. « Spatiotemporal intermittency regimes of the one-dimensional complex Ginzburg-Landau equation ». In : *Nonlinearity* 7.1 (1994), p. 185.
- [40] Boris I SHRAIMAN, Alain PUMIR, Wim van SAARLOOS, Pierre C HOHENBERG, Hugues CHATÉ et Margaret HOLEN. « Spatiotemporal chaos in the one-dimensional complex Ginzburg-Landau equation ». In : *Physica D : Nonlinear Phenomena* 57.3-4 (1992), p. 241-248.
- [41] Neils C RATTENBORG, Charles J AMLANER et Steven L LIMA. « Behavioral, neurophysiological and evolutionary perspectives on unihemispheric sleep ». In : *Neuroscience & Biobehavioral Reviews* 24.8 (2000), p. 817-842.

- [42] Peter J UHLHAAS et Wolf SINGER. « Neural synchrony in brain disorders : relevance for cognitive dysfunctions and pathophysiology ». In : *neuron* 52.1 (2006), p. 155-168.
- [43] Soumen MAJHI, Bidesh K BERA, Dibakar GHOSH et Matjaž PERC. « Chimera states in neuronal networks : a review ». In : *Physics of life reviews* 28 (2019), p. 100-121.
- [44] Adilson E MOTTER. « Spontaneous synchrony breaking ». In : *Nature Physics* 6.3 (2010), p. 164-165.
- [45] Mark J PANAGGIO et Daniel M ABRAMS. « Chimera states : coexistence of coherence and incoherence in networks of coupled oscillators ». In : *Nonlinearity* 28.3 (2015), R67.
- [46] Daniel M ABRAMS et Steven H STROGATZ. « Chimera states in a ring of nonlocally coupled oscillators ». In : *International Journal of Bifurcation and Chaos* 16.01 (2006), p. 21-37.
- [47] E OMEL'CHENKO, Matthias WOLFRUM et Yuri L MAISTRENKO. « Chimera states as chaotic spatiotemporal patterns ». In : *Physical Review E* 81.6 (2010), p. 065201.
- [48] Shin-ichiro SHIMA et Yoshiaki KURAMOTO. « Rotating spiral waves with phase-randomized core in nonlocally coupled oscillators ». In : *Physical Review E* 69.3 (2004), p. 036213.
- [49] Erik A MARTENS, Carlo R LAING et Steven H STROGATZ. « Solvable model of spiral wave chimeras ». In : *Physical review letters* 104.4 (2010), p. 044101.
- [50] Mark J PANAGGIO et Daniel M ABRAMS. « Chimera states on a flat torus ». In : *Physical review letters* 110.9 (2013), p. 094102.
- [51] Iryna OMELCHENKO, Yuri MAISTRENKO, Philipp HÖVEL et Eckehard SCHÖLL. « Loss of coherence in dynamical networks : spatial chaos and chimera states ». In : *Physical review letters* 106.23 (2011), p. 234102.
- [52] Iryna OMELCHENKO, E OMEL'CHENKO, Philipp HÖVEL et Eckehard SCHÖLL. « When nonlocal coupling between oscillators becomes stronger : patched synchrony or multichimera states ». In : *Physical review letters* 110.22 (2013), p. 224101.
- [53] Hidetsugu SAKAGUCHI. « Instability of synchronized motion in nonlocally coupled neural oscillators ». In : *Physical Review E* 73.3 (2006), p. 031907.
- [54] R GOPAL, VK CHANDRASEKAR, A VENKATESAN et M LAKSHMANAN. « Observation and characterization of chimera states in coupled dynamical systems with nonlocal coupling ». In : *Physical review E* 89.5 (2014), p. 052914.
- [55] Gautam C SETHIA, Abhijit SEN et Fatihcan M ATAY. « Clustered chimera states in delay-coupled oscillator systems ». In : *Physical review letters* 100.14 (2008), p. 144102.
- [56] E OMEL'CHENKO, Yuri L MAISTRENKO et Peter A TASS. « Chimera states : The natural link between coherence and incoherence ». In : *Physical review letters* 100.4 (2008), p. 044105.

- [57] Gautam C SETHIA et Abhijit SEN. « Chimera states : the existence criteria revisited ». In : *Physical review letters* 112.14 (2014), p. 144101.
- [58] Azamat YELDESBAY, Arkady PIKOVSKY et Michael ROSENBLUM. « Chimeralike states in an ensemble of globally coupled oscillators ». In : *Physical review letters* 112.14 (2014), p. 144103.
- [59] Bidesh K BERA et Dibakar GHOSH. « Chimera states in purely local delay-coupled oscillators ». In : *Physical Review E* 93.5 (2016), p. 052223.
- [60] Carlo R LAING. « Chimeras in networks with purely local coupling ». In : *Physical Review E* 92.5 (2015), p. 050904.
- [61] MG CLERC, Saliya COULIBALY, MA FERRÉ, MA GARCÍA-ÑUSTES et RG ROJAS. « Chimera-type states induced by local coupling ». In : *Physical Review E* 93.5 (2016), p. 052204.
- [62] Jacques KENGNE, Zeric Tabekoueng NJITACKE, A NGUOMKAM NEGOU, M FOUODJI TSOSTOP et Hilaire Bertrand FOTSIN. « Coexistence of multiple attractors and crisis route to chaos in a novel chaotic jerk circuit ». In : *International Journal of Bifurcation and Chaos* 26.05 (2016), p. 1650081.
- [63] MG CLERC, Saliya COULIBALY, MA FERRÉ et RG ROJAS. « Chimera states in a Duffing oscillators chain coupled to nearest neighbors ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 28.8 (2018), p. 083126.
- [64] Zahra FAGHANI, Zahra ARAB, Fatemeh PARASTESH, Sajad JAFARI, Matjaž PERC et Mitja SLAVINEC. « Effects of different initial conditions on the emergence of chimera states ». In : *Chaos, Solitons & Fractals* 114 (2018), p. 306-311.
- [65] Sarbendu RAKSHIT, Bidesh K BERA, Matjaž PERC et Dibakar GHOSH. « Basin stability for chimera states ». In : *Scientific reports* 7.1 (2017), p. 1-12.
- [66] Mark R TINSLEY, Simbarashe NKOMO et Kenneth SHOWALTER. « Chimera and phase-cluster states in populations of coupled chemical oscillators ». In : *Nature Physics* 8.9 (2012), p. 662-665.
- [67] N LAZARIDES, G NEOFOTISTOS et GP TSIRONIS. « Chimeras in SQUID metamaterials ». In : *Physical Review B* 91.5 (2015), p. 054303.
- [68] Erik Andreas MARTENS, Shashi THUTUPALLI, Antoine FOURRIERE et Oskar HALLATSCHEK. « Chimera states in mechanical oscillator networks ». In : *Proceedings of the National Academy of Sciences* 110.26 (2013), p. 10563-10567.
- [69] Laurent LARGER, Bogdan PENKOVSKY et Yuri MAISTRENKO. « Virtual chimera states for delayed-feedback systems ». In : *Physical review letters* 111.5 (2013), p. 054103.
- [70] Laurent LARGER, Bogdan PENKOVSKY et Yuri MAISTRENKO. « Laser chimeras as a paradigm for multistable patterns in complex systems ». In : *Nature communications* 6.1 (2015), p. 1-7.

- [71] Evgeny A VIKTOROV, Tatiana HABRUSEVA, Stephen P HEGARTY, Guillaume HUYET et Bryan KELLEHER. « Coherence and incoherence in an optical comb ». In : *Physical review letters* 112.22 (2014), p. 224101.
- [72] Chi-Hak UY, Lionel WEICKER, Damien RONTANI et Marc SCIAMANNA. « Optical chimera in light polarization ». In : *APL Photonics* 4.5 (2019), p. 056104.
- [73] Bruno GARBIN, Andrus GIRALDO, Kevin JH PETERS, Neil GR BRODERICK, Anne SPAKMAN, Fabrice RAINERI, Ariel LEVENSON, Said RK RODRIGUEZ, Bernd KRAUSKOPF et Alejandro M YACOMOTTI. « Spontaneous symmetry breaking in a coherently driven nanophotonic Bose-Hubbard dimer ». In : *Physical Review Letters* 128.5 (2022), p. 053901.
- [74] OA EGOROV et F LEDERER. « Spontaneously walking discrete cavity solitons ». In : *Optics letters* 38.7 (2013), p. 1010-1012.
- [75] U PESCHEL, O EGOROV et F LEDERER. « Discrete cavity solitons ». In : *Optics letters* 29.16 (2004), p. 1909-1911.
- [76] Pedro PARRA-RIVAS, Damià GOMILA, Manuel A MATÍAS, Stéphane COEN et Lendert GELENS. « Dynamics of localized and patterned structures in the Lugiato-Lefever equation determine the stability and shape of optical frequency combs ». In : *Physical Review A* 89.4 (2014), p. 043813.
- [77] Zheng LIU, Mardia OUALI, Saliya COULIBALY, MG CLERC, M TAKI et Mustapha TLIDI. « Characterization of spatiotemporal chaos in a Kerr optical frequency comb and in all fiber cavities ». In : *Optics letters* 42.6 (2017), p. 1063-1066.
- [78] Luigi A LUGIATO et René LEFEVER. « Spatial dissipative structures in passive optical systems ». In : *Physical review letters* 58.21 (1987), p. 2209.
- [79] Miles ANDERSON, François LEO, Stéphane COEN, Miro ERKINTALO et Stuart G MURDOCH. « Observations of spatiotemporal instabilities of temporal cavity solitons ». In : *Optica* 3.10 (2016), p. 1071-1074.
- [80] Gábor FÁTH. « Propagation failure of traveling waves in a discrete bistable medium ». In : *Physica D : Nonlinear Phenomena* 116.1-2 (1998), p. 176-190.
- [81] Marcel G CLERC, Ricardo G ELÍAS et René G ROJAS. « Continuous description of lattice discreteness effects in front propagation ». In : *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences* 369.1935 (2011), p. 412-424.
- [82] Kuniyuki KANEKO. « Spatiotemporal intermittency in coupled map lattices ». In : *Progress of Theoretical Physics* 74.5 (1985), p. 1033-1044.
- [83] Hugues CHATÉ et Paul MANNEVILLE. « Collective behaviors in spatially extended systems with local interactions and synchronous updating ». In : *Progress of theoretical physics* 87.1 (1992), p. 1-60.
- [84] Mikhail PROKOPENKO, Fabio BOSCHETTI et Alex J RYAN. « An information-theoretic primer on complexity, self-organization, and emergence ». In : *Complexity* 15.1 (2009), p. 11-28.

- [85] Melanie MITCHELL. *Complexity : A guided tour*. Oxford university press, 2009.
- [86] James P CRUTCHFIELD. « The calculi of emergence : computation, dynamics and induction ». In : *Physica D : Nonlinear Phenomena* 75.1-3 (1994), p. 11-54.
- [87] David P FELDMAN, Carl S MCTAGUE et James P CRUTCHFIELD. « The organization of intrinsic computation : Complexity-entropy diagrams and the diversity of natural information processing ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 18.4 (2008), p. 043106.
- [88] Stephen WOLFRAM. « Cellular automata as models of complexity ». In : *Nature* 311.5985 (1984), p. 419-424.
- [89] Mikhail PROKOPENKO, Vadim GERASIMOV et Ivan TANEV. « Evolving spatiotemporal coordination in a modular robotic system ». In : *International conference on simulation of adaptive behavior*. Springer. 2006, p. 558-569.
- [90] K-I GOH et A-L BARABÁSI. « Burstiness and memory in complex systems ». In : *EPL (Europhysics Letters)* 81.4 (2008), p. 48002.
- [91] Rafael MORGADO, Michał CIEŚLA, Lech LONGA et Fernando A OLIVEIRA. « Synchronization in the presence of memory ». In : *EPL (Europhysics Letters)* 79.1 (2007), p. 10002.
- [92] JA BROWN et JA TUSZYNSKI. « A review of the ferroelectric model of microtubules ». In : *Ferroelectrics* 220.1 (1999), p. 141-155.
- [93] Jürgen PAHLE, Anne K GREEN, C Jane DIXON et Ursula KUMMER. « Information transfer in signaling pathways : a study using coupled simulated and experimental data ». In : *BMC bioinformatics* 9.1 (2008), p. 1-14.
- [94] Iain D COUZIN, Richard JAMES, David MAWDSLEY, Darren P CROFT et Jens KRAUSE. « Social organization and information transfer in schooling fishes ». In : *Fish and Aquatic Resources Series* (2006).
- [95] Alexander S KLYUBIN, Daniel POLANI et Chrystopher L NEHANIV. « All else being equal be empowered ». In : *European Conference on Artificial Life*. Springer. 2005, p. 744-753.
- [96] Max LUNGARELLA et Olaf SPORNS. « Mapping information flow in sensorimotor networks ». In : *PLoS computational biology* 2.10 (2006), e144.
- [97] Osame KINOCHI et Mauro COPELLI. « Optimal dynamical range of excitable networks at criticality ». In : *Nature physics* 2.5 (2006), p. 348-351.
- [98] Joseph J ATICK. « Could information theory provide an ecological theory of sensory processing? » In : *Network : Computation in neural systems* 3.2 (1992), p. 213-251.
- [99] Manuel A SÁNCHEZ-MONTAÑÉS et Fernando J CORBACHO. « Towards a new information processing measure for neural computation ». In : *International Conference on Artificial Neural Networks*. Springer. 2002, p. 637-642.

- [100] T YAMADA et K AIHARA. « Spatio-temporal complex dynamics and computation in chaotic neural networks ». In : *ETFA'94. 1994 IEEE Symposium on Emerging Technologies and Factory Automation. (SEIKEN) Symposium)-Novel Disciplines for the Next Century-Proceedings*. IEEE. 1994, p. 239-244.
- [101] Mariusz H JAKUBOWSKI, Ken STEIGLITZ et Richard SQUIER. « Information transfer between solitary waves in the saturable Schrödinger equation ». In : *Physical Review E* 56.6 (1997), p. 7267.
- [102] Andrew ADAMATZKY. « New media for collision-based computing ». In : *Collision-Based Computing*. Springer, 2002, p. 411-442.
- [103] DE EDMUNDSON et RH ENNS. « Fully three-dimensional collisions of bistable light bullets ». In : *Optics letters* 18.19 (1993), p. 1609-1611.
- [104] Stephen WOLFRAM et M GAD-EL-HAK. « A new kind of science ». In : *Appl. Mech. Rev.* 56.2 (2003), B18-B19.
- [105] Pulin GONG et Cees VAN LEEUWEN. « Distributed dynamical computation in neural circuits with propagating coherent activity patterns ». In : *PLoS Computational Biology* 5.12 (2009), e1000611.
- [106] Eugene V KOONIN, Yuri I WOLF et Georgy P KAREV. *Power laws, scale-free networks and genome biology*. Springer, 2006.
- [107] Octavio MIRAMONTES. « Order-disorder transitions in the behavior of ant societies ». In : *Complexity* 1.3 (1995), p. 56-60.
- [108] Seth LLOYD. « The computational universe ». In : *Information and the nature of reality : From physics to metaphysics* (2010), p. 92-103.
- [109] Mile GU, Karoline WIESNER, Elisabeth RIEPER et Vlatko VEDRAL. « Quantum mechanics can reduce the complexity of classical models ». In : *Nature communications* 3.1 (2012), p. 1-5.
- [110] James P CRUTCHFIELD et Karl YOUNG. « Inferring statistical complexity ». In : *Physical review letters* 63.2 (1989), p. 105.
- [111] James P CRUTCHFIELD et David P FELDMAN. « Regularities unseen, randomness observed : Levels of entropy convergence ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 13.1 (2003), p. 25-54.
- [112] Alexander S KLYUBIN, Daniel POLANI et Chrystopher L NEHANIV. « Representations of space and time in the maximization of information flow in the perception-action loop ». In : *Neural computation* 19.9 (2007), p. 2387-2432.
- [113] Murray GELL-MANN. « Complex adaptive systems ». In : (1994).
- [114] Cosma Rohilla SHALIZI, Robert HASLINGER, Jean-Baptiste ROUQUIER, Kristina Lisa KLINKNER et Cristopher MOORE. « Automatic filters for the detection of coherent structure in spatiotemporal systems ». In : *Physical Review E* 73.3 (2006), p. 036104.

- [115] Joseph T LIZIER, Mikhail PROKOPENKO et Albert Y ZOMAYA. « Local measures of information storage in complex distributed computation ». In : *Information Sciences* 208 (2012), p. 39-54.
- [116] Andrei Nikolaevitch KOLMOGOROV. « Entropy per unit time as a metric invariant of automorphisms ». In : *Dokl. Akad. Nauk SSSR*. T. 124. 4. 1959, p. 754-755.
- [117] Yakov G SINAI. « On the notion of entropy of a dynamical system ». In : *Doklady of Russian Academy of Sciences*. T. 124. 3. 1959, p. 768-771.
- [118] Robert SHAW. « Strange attractors, chaotic behavior, and information flow ». In : *Zeitschrift für Naturforschung A* 36.1 (1981), p. 80-112.
- [119] JP CRUTCHFIELD et NH PACKARD. « Symbolic dynamics of one-dimensional maps : Entropies, finite precision, and noise ». In : *International Journal of Theoretical Physics* 21.6 (1982), p. 433-466.
- [120] AA BRUDNO. « Entropy and the complexity of the trajectories of a dynamic system ». In : *Trudy Moskovskogo Matematicheskogo Obshchestva* 44 (1982), p. 124-149.
- [121] James P CRUTCHFIELD. « Between order and chaos ». In : *Nature Physics* 8.1 (2012), p. 17-24.
- [122] Jan BOGG et Robert GEYER. *Complexity, science and society*. Radcliffe Abingdon, 2007.
- [123] Stuart A KAUFFMAN et al. *The origins of order : Self-organization and selection in evolution*. Oxford University Press, USA, 1993.
- [124] James P CRUTCHFIELD et Karl YOUNG. « Computation at the onset of chaos ». In : *The Santa Fe Institute, Westview*. Citeseer. 1988.
- [125] Kuniyiko KANEKO et Takashi IKEGAMI. « Homeochaos : dynamics stability of a symbiotic network with population dynamics and evolving mutation rates ». In : *Physica D : Nonlinear Phenomena* 56.4 (1992), p. 406-429.
- [126] Ricard V SOLÉ et Octavio MIRAMONTES. « Information at the edge of chaos in fluid neural networks ». In : *Physica D : Nonlinear Phenomena* 80.1-2 (1995), p. 171-180.
- [127] Melanie MITCHELL, Peter HRABER et James P CRUTCHFIELD. « Revisiting the edge of chaos : Evolving cellular automata to perform computations ». In : *arXiv preprint adap-org/9303003* (1993).
- [128] Ricard V SOLÉ et Sergi VALVERDE. « Information transfer and phase transitions in a model of internet traffic ». In : *Physica A : Statistical Mechanics and its Applications* 289.3-4 (2001), p. 595-605.
- [129] Hermann HAKEN. « Synergetics : an overview ». In : *Reports on Progress in Physics* 52.5 (1989), p. 515.
- [130] Brian GOODWIN. *How the leopard changed its spots : The evolution of complexity*. Princeton University Press, 2001.

- [131] Pierre C HOHENBERG et Bertrand I HALPERIN. « Theory of dynamic critical phenomena ». In : *Reviews of Modern Physics* 49.3 (1977), p. 435.
- [132] John VON NEUMANN, Arthur W BURKS et al. « Theory of self-reproducing automata ». In : *IEEE Transactions on Neural Networks* 5.1 (1966), p. 3-14.
- [133] Arthur Walter BURKS. *Essays on cellular automata*. Urbana : University of Illinois Press, 1970.
- [134] Martin GARDNER. « The Fantastic Combinations of Jhon Conway's New Solitaire Game'Life ». In : *Sc. Am.* 223 (1970), p. 20-123.
- [135] Wolfgang ALT. *Dynamics of cell and tissue motion*. Springer Science & Business Media, 1997.
- [136] Nicholas J SAVILL et Paulien HOGEWEG. « Modelling morphogenesis : from single cells to crawling slugs ». In : *Journal of theoretical biology* 184.3 (1997), p. 229-235.
- [137] Rob J De BOER, Pauline HOGEWEG et Alan S PERELSON. « Growth and recruitment in the immune network ». In : *Theoretical and Experimental Insights into Immunology*. Springer, 1992, p. 223-247.
- [138] Franco CELADA et Philip E SEIDEN. « A computer model of cellular interactions in the immune system ». In : *Immunology today* 13.2 (1992), p. 56-62.
- [139] Joana MOREIRA et Andreas DEUTSCH. « Cellular automaton models of tumor development : a critical review ». In : *Advances in Complex Systems* 5.02n03 (2002), p. 247-267.
- [140] JH MOORE et LW HAHN. « A cellular automata-based pattern recognition approach to identifying gene-gene and gene-environment interactions. » In : *American Journal of Human Genetics*. T. 67. 4. UNIV CHICAGO PRESS 5720 SOUTH WOODLAWN AVE, CHICAGO, IL 60637-1603 USA. 2000, p. 52-52.
- [141] Jason H MOORE et Lance W HAHN. « Multilocus pattern recognition using cellular automata and parallel genetic algorithms ». In : *Proceedings of the 3rd Annual Conference on Genetic and Evolutionary Computation*. 2001, p. 1452-1452.
- [142] Mitchel RESNICK. *Turtles, termites, and traffic jams : Explorations in massively parallel microworlds*. Mit Press, 1997.
- [143] Birgitt SCHÖNFISCH et Michael KINDER. « A fish migration model ». In : *International Conference on Cellular Automata*. Springer. 2002, p. 210-219.
- [144] Heiko BALZTER, Paul W BRAUN et Wolfgang KÖHLER. « Cellular automata models for vegetation dynamics ». In : *Ecological modelling* 107.2-3 (1998), p. 113-125.
- [145] Qing-cun ZENG et Xiao-dong ZENG. « An analytical dynamic model of grass field ecosystem with two variables ». In : *Ecological Modelling* 85.2-3 (1996), p. 187-196.
- [146] Michael F SHLESINGER. « New paths for random walkers ». In : *Nature* 355.6359 (1992), p. 396-397.

- [147] Christian BURKS et Doyne FARMER. « Towards modeling DNA sequences as automata ». In : *Physica D : nonlinear phenomena* 10.1-2 (1984), p. 157-167.
- [148] Steven A SMITH, Richard C WATT et Stuart R HAMEROFF. « Cellular automata in cytoskeletal lattices ». In : *Physica D : Nonlinear Phenomena* 10.1-2 (1984), p. 168-174.
- [149] David A YOUNG. « A local activator-inhibitor model of vertebrate skin patterns ». In : *Mathematical Biosciences* 72.1 (1984), p. 51-58.
- [150] N PACKARD. « Lattice models for solidification and aggregation ». In : *Proc. 1st International Symposium for Science on Form*. 1985, p. 95-101.
- [151] Bastien CHOPARD, M DROZ, S CORNELL et Laurent FRACHEBOURG. « Cellular automata approach to reaction-diffusion system : Theory and application ». In : *Cellular automata prospects in astrophysical applications* (1993), p. 157-186.
- [152] Bastien CHOPARD, Michael DROZ et Max KOLB. « Cellular automata approach to non-equilibrium diffusion and gradient percolation ». In : *Journal of Physics A : Mathematical and General* 22.10 (1989), p. 1609.
- [153] Richard J GAYLORD et Paul R WELLIN. *Computer simulations with Mathematica : explorations in complex physical and biological systems*. Springer-Verlag, 1995.
- [154] Fong LIU et Nigel GOLDENFELD. « Generic features of late-stage crystal growth ». In : *Physical Review A* 42.2 (1990), p. 895.
- [155] Birgitt SCHÖNFISCH. « Propagation of fronts in cellular automata ». In : *Physica D : Nonlinear Phenomena* 80.4 (1995), p. 433-450.
- [156] Michael CREUTZ. « Deterministic ising dynamics ». In : *Annals of physics* 167.1 (1986), p. 62-72.
- [157] Yves POMEAU. « Invariant in cellular automata ». In : *Journal of Physics A : Mathematical and General* 17.8 (1984), p. L415.
- [158] Barry F MADORE et Wendy L FREEDMAN. « Computer simulations of the Belousov-Zhabotinsky reaction ». In : *Science* 222.4624 (1983), p. 615-616.
- [159] Yoshitsugu OONO et M KOHMOTO. « Discrete model of chemical turbulence ». In : *Physical review letters* 55.27 (1985), p. 2927.
- [160] AT WINFREE, EM WINFREE et H SEIFERT. « Organizing centers in a cellular excitable medium ». In : *Physica D : Nonlinear Phenomena* 17.1 (1985), p. 109-115.
- [161] Uriel FRISCH, Brosl HASSLACHER et Yves POMEAU. « Lattice-gas automata for the Navier-Stokes equation ». In : *Lattice Gas Methods for Partial Differential Equations*. CRC Press, 2019, p. 11-18.
- [162] James M SAKODA. « The checkerboard model of social interaction ». In : *The Journal of Mathematical Sociology* 1.1 (1971), p. 119-132.
- [163] David GRIFFEATH et Cristopher MOORE. *New constructions in cellular automata*. Oxford University Press, 2003.

- [164] Giovanni A RABINO et Alessandra LAGHI. « Urban cellular automata : the inverse problem ». In : *International Conference on Cellular Automata*. Springer. 2002, p. 349-356.
- [165] Heinz MÜHLENBEIN et Robin HÖNS. « Stochastic analysis of cellular automata and the voter model ». In : *International Conference on Cellular Automata*. Springer. 2002, p. 92-103.
- [166] Rainer HEGSELMANN et Andreas FLACHE. « Understanding complex social dynamics : A plea for cellular automata based modelling ». In : *Journal of Artificial Societies and Social Simulation* 1.3 (1998), p. 1.
- [167] Karel CULIK et Sheng YU. « Undecidability of CA classification schemes ». In : *Complex Systems* 2.2 (1988), p. 177-190.
- [168] Wentian LI, Norman H PACKARD et Chris G LANGTON. « Transition phenomena in cellular automata rule space ». In : *Physica D : Nonlinear Phenomena* 45.1-3 (1990), p. 77-94.
- [169] Andrew WUENSCH. « Classifying cellular automata automatically : Finding gliders, filtering, and relating space-time patterns, attractor basins, and the Z parameter ». In : *Complexity* 4.3 (1999), p. 47-66.
- [170] Gina MB OLIVEIRA, Pedro PB DE OLIVEIRA et Nizam OMAR. « Definition and application of a five-parameter characterization of one-dimensional cellular automata rule space ». In : *Artificial life* 7.3 (2001), p. 277-301.
- [171] Arthur T WINFREE. *The geometry of biological time*. T. 2. Springer, 1980.
- [172] PE CLADIS et eds PALFFY-MUHORAY P. « Spatio-temporal patterns in nonequilibrium complex systems ». In : *SFI Studies in the Sciences of Complexity (Addison Wesley Longman, Reading, MA, 1995)* (1995).
- [173] H Frederick NIJHOUT. *Pattern formation in the physical and biological sciences*. CRC Press, 2018.
- [174] Eric BONABEAU. « Social insect colonies as complex adaptive systems ». In : *Ecosystems* 1.5 (1998), p. 437-443.
- [175] James P CRUTCHFIELD. « Is anything ever new? Considering emergence ». In : *SANTA FE INSTITUTE STUDIES IN THE SCIENCES OF COMPLEXITY-PROCEEDINGS VOLUME-*. T. 19. ADDISON-WESLEY PUBLISHING CO. 1994, p. 515-515.
- [176] Timothy O'CONNOR. « Emergent properties ». In : (2020).
- [177] James L MCCLELLAND, David E RUMELHART et Geoffrey E HINTON. « The appeal of parallel distributed processing ». In : *MIT Press, Cambridge MA* 3 (1986), p. 44.
- [178] Stephanie FORREST. « Emergent computation : self-organizing, collective, and cooperative phenomena in natural and artificial computing networks : introduction to the proceedings of the ninth annual CNLS conference ». In : *Physica D : Nonlinear Phenomena* 42.1-3 (1990), p. 1-11.

- [179] Lawrence GRAY, A NEW et al. « A mathematician looks at Wolfram's new kind of science ». In : *Notices of the American Mathematical Society* 50 (2)(2003) 200–211, URL <http://www.ams.org/notices/200302/fea-gray.pdf>. URL <http://www.ams.org/notices/200302/fea-gray.pdf>. Citeseer. 2003.
- [180] James E HANSON et James P CRUTCHFIELD. « The attractor—Basin portrait of a cellular automaton ». In : *Journal of statistical physics* 66.5 (1992), p. 1415-1462.
- [181] Melanie MITCHELL et al. « Computation in Cellular Automata : A Selected Review. » In : *Non-standard computation* (2005), p. 95-140.
- [182] Melanie MITCHELL, James P CRUTCHFIELD et Peter T HRABER. « Evolving cellular automata to perform computations : Mechanisms and impediments ». In : *Physica D : Nonlinear Phenomena* 75.1-3 (1994), p. 361-391.
- [183] Nino BOCCARA, Jamil NASSER et Michel ROGER. « Particlelike structures and their interactions in spatiotemporal patterns generated by one-dimensional deterministic cellular-automaton rules ». In : *Physical Review A* 44.2 (1991), p. 866.
- [184] Bruno MARTIN. « A Group Interpretation of Particles Generated by One-Dimensional Cellular Automaton, Wolfram's Rule 54 ». In : *International Journal of Modern Physics C* 11.01 (2000), p. 101-123.
- [185] Genaro Juárez MARTÍNEZ, Andrew ADAMATZKY et Harold V MCINTOSH. « Phenomenology of glider collisions in cellular automaton Rule 54 and associated logical gates ». In : *Chaos, Solitons & Fractals* 28.1 (2006), p. 100-111.
- [186] Wim HORDIJK, Cosma Rohilla SHALIZI et James P CRUTCHFIELD. « Upper bound on the products of particle interactions in cellular automata ». In : *Physica D : Nonlinear Phenomena* 154.3-4 (2001), p. 240-258.
- [187] Mariusz H JAKUBOWSKI, Ken STEIGLITZ et Richard SQUIER. « Computing with solitons : a review and prospectus ». In : *Collision-based computing* (2002), p. 277-297.
- [188] Joseph T LIZIER, Mikhail PROKOPENKO et Albert Y ZOMAYA. « Local information transfer as a spatiotemporal filter for complex systems ». In : *Physical Review E* 77.2 (2008), p. 026110.
- [189] Joseph T LIZIER, Mikhail PROKOPENKO et Albert Y ZOMAYA. « Information modification and particle collisions in distributed computation ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 20.3 (2010).
- [190] Stephen WOLFRAM. « Universality and complexity in cellular automata ». In : *Physica D : Nonlinear Phenomena* 10.1-2 (1984), p. 1-35.
- [191] Stephen WOLFRAM. « Computation theory of cellular automata ». In : *Communications in mathematical physics* 96.1 (1984), p. 15-57.
- [192] Joseph T LIZIER, Mikhail PROKOPENKO et Albert Y ZOMAYA. « A framework for the local information dynamics of distributed computation in complex systems ». In : *Guided self-organization : inception*. Springer, 2014, p. 115-158.

- [193] MT MARTIN, A PLASTINO et OA ROSSO. « Generalized statistical complexity measures : Geometrical and analytical properties ». In : *Physica A : Statistical Mechanics and its Applications* 369.2 (2006), p. 439-462.
- [194] OA ROSSO, HA LARRONDO, MT MARTIN, A PLASTINO et MA FUENTES. « Distinguishing noise from chaos ». In : *Physical review letters* 99.15 (2007), p. 154102.
- [195] Abraham LEMPEL et Jacob ZIV. « On the complexity of finite sequences ». In : *IEEE Transactions on information theory* 22.1 (1976), p. 75-81.
- [196] F KASPAR et HG SCHUSTER. « Easily calculable measure for the complexity of spatiotemporal patterns ». In : *Physical review A* 36.2 (1987), p. 842.
- [197] Steven M PINCUS. « Approximate entropy as a measure of system complexity ». In : *Proceedings of the National Academy of Sciences* 88.6 (1991), p. 2297-2301.
- [198] Joshua S RICHMAN et J Randall MOORMAN. « Physiological time-series analysis using approximate entropy and sample entropy ». In : *American journal of physiology-heart and circulatory physiology* 278.6 (2000), H2039-H2049.
- [199] Alfonso DELGADO-BONAL et Alexander MARSHAK. « Approximate entropy and sample entropy : A comprehensive tutorial ». In : *Entropy* 21.6 (2019), p. 541.
- [200] Peter GRASSBERGER et Itamar PROCACCIA. « Characterization of strange attractors ». In : *Physical review letters* 50.5 (1983), p. 346.
- [201] M BAUER, H HENG et W MARTIENSSEN. « Characterization of spatiotemporal chaos from time series ». In : *Physical review letters* 71.4 (1993), p. 521.
- [202] Wentian LI et Norman PACKARD. « The structure of the elementary cellular automata rule space ». In : *Complex systems* 4.3 (1990), p. 281-297.
- [203] Hector ZENIL. « Compression-based investigation of the dynamical properties of cellular automata and other systems ». In : *arXiv preprint arXiv :0910.4042* (2009).
- [204] Hector ZENIL. « On the dynamic qualitative behaviour of universal computation ». In : *arXiv preprint arXiv :1201.0824* (2012).
- [205] Wentian LI et Mats G NORDAHL. « Transient behavior of cellular automaton rule 110 ». In : *Physics Letters A* 166.5-6 (1992), p. 335-339.
- [206] Matthew COOK. « Universality in elementary cellular automata ». In : *Complex systems* 15.1 (2004), p. 1-40.
- [207] Shigeru NINAGAWA. « Power spectral analysis of elementary cellular automata ». In : *Complex Systems* 17.4 (2008), p. 399.
- [208] Marvin S KESHNER. « 1/f noise ». In : *Proceedings of the IEEE* 70.3 (1982), p. 212-218.
- [209] Yves POMEAU et Paul MANNEVILLE. « Intermittent transition to turbulence in dissipative dynamical systems ». In : *Communications in Mathematical Physics* 74.2 (1980), p. 189-197.

- [210] Shigeru NINAGAWA, Masaaki YONEDA et Sadaki HIROSE. « 1f fluctuation in the “Game of Life” ». In : *Physica D : Nonlinear Phenomena* 118.1-2 (1998), p. 49-52.
- [211] S NINAGAWA. « Dynamics of universal computation and 1/f noise in elementary cellular automata ». In : *Chaos, Solitons & Fractals* 70 (2015), p. 42-48.
- [212] Lazaros K GALLOS, Mariano SIGMAN et Hernán A MAKSE. « The conundrum of functional brain networks : small-world efficiency or fractal modularity ». In : *Frontiers in physiology* 3 (2012), p. 123.
- [213] Lazaros K GALLOS, Hernán A MAKSE et Mariano SIGMAN. « A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks ». In : *Proceedings of the National Academy of Sciences* 109.8 (2012), p. 2825-2830.
- [214] Zhixin LU, Jaideep PATHAK, Brian HUNT, Michelle GIRVAN, Roger BROCKETT et Edward OTT. « Reservoir observers : Model-free inference of unmeasured variables in chaotic systems ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 27.4 (2017), p. 041102.
- [215] Alexander HALUSZCZYNSKI et Christoph RÄTH. « Good and bad predictions : Assessing and improving the replication of chaotic attractors by means of reservoir computing ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.10 (2019), p. 103143.
- [216] Abd AlRahman R ALMOMANI, Jie SUN et Erik BOLLT. « How entropic regression beats the outliers problem in nonlinear system identification ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 30.1 (2020), p. 013107.
- [217] Tom BERTALAN, Felix DIETRICH, Igor MEZIĆ et Ioannis G KEVREKIDIS. « On learning Hamiltonian systems from data ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.12 (2019), p. 121107.
- [218] Duxin CHEN, Yuchen WANG, Ge WU, Mingyu KANG, Yongzheng SUN et Wenwu YU. « Inferring causal relationship in coordinated flight of pigeon flocks ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.11 (2019), p. 113118.
- [219] Daniel R GUREVICH, Patrick AK REINBOLD et Roman O GRIGORIEV. « Robust and optimal sparse regression for nonlinear PDE models ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.10 (2019), p. 103113.
- [220] Seungjoon LEE, Mahdi KOOSHKBAGHI, Konstantinos SPILOTIS, Constantinos I SIETTOS et Ioannis G KEVREKIDIS. « Coarse-scale PDEs from fine-scale observations via machine learning ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 30.1 (2020), p. 013141.
- [221] Cun MA, Qirui YANG, Xiaoqun WU et Jun-an LU. « Cluster synchronization : From single-layer to multi-layer networks ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.12 (2019), p. 123120.

- [222] Mark J PANAGGIO, Maria-Veronica CIOCANEL, Lauren LAZARUS, Chad M TOPAZ et Bin XU. « Model reconstruction from temporal data for coupled oscillator networks ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.10 (2019), p. 103116.
- [223] Dengfeng WU, Miaomiao FU et Jinqiao DUAN. « Discovering mean residence time and escape probability from data of stochastic dynamical systems ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.9 (2019), p. 093122.
- [224] Guorui SHEN, Jürgen KURTHS et Ye YUAN. « Sequence-to-sequence prediction of spatiotemporal systems ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 30.2 (2020), p. 023102.
- [225] Takayoshi HACHIJO, Shinga MASUDA, Takuya KUROSAKA et Hiroshi GOTODA. « Early detection of thermoacoustic combustion oscillations using a methodology combining statistical complexity and machine learning ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.10 (2019), p. 103123.
- [226] Pablo AMIL, Miguel C SORIANO et Cristina MASOLLER. « Machine learning algorithms for predicting the amplitude of chaotic laser pulses ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.11 (2019), p. 113111.
- [227] Martin LELLEP, Jonathan PREXL, Moritz LINKMANN et Bruno ECKHARDT. « Using machine learning to predict extreme events in the Hénon map ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 30.1 (2020), p. 013113.
- [228] Rosangela FOLLMANN et Epaminondas ROSA JR. « Predicting slow and fast neuronal dynamics with machine learning ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.11 (2019), p. 113119.
- [229] Zhong-Ke GAO, Yan-Li LI, Yu-Xuan YANG et Chao MA. « A recurrence network-based convolutional neural network for fatigue driving detection from EEG ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.11 (2019), p. 113126.
- [230] Christoph JANSEN, Thomas PENZEL, Stephan HODEL, Stefanie BREUER, Martin SPOTT et Dagmar KREFTING. « Network physiology in insomnia patients : Assessment of relevant changes in network topology with interpretable machine learning models ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.12 (2019), p. 123129.
- [231] Aleksei SELEZNEV, Dmitry MUKHIN, Andrey GAVRILOV, Evgeny LOSKUTOV et Alexander FEIGIN. « Bayesian framework for simulation of dynamical systems from multidimensional data using recurrent neural network ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.12 (2019), p. 123115.
- [232] S HERZOG, F WÖRGÖTTER et Ulrich PARLITZ. « Convolutional autoencoder and conditional random fields hybrid for predicting spatial-temporal chaos ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.12 (2019), p. 123116.

- [233] MA GANAIE, Saptarshi GHOSH, Naveen MENDOLA, Muhammad TANVEER et Sarika JALAN. « Identification of chimera using machine learning ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 30.6 (2020), p. 063128.
- [234] Sandip V GEORGE, Ranjeev MISRA et G AMBIKA. « Classification of close binary stars using recurrence networks ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.11 (2019), p. 113112.
- [235] Harikrishnan Nellippallil BALAKRISHNAN, Aditi KATHPALIA, Snehanshu SAHA et Nithin NAGARAJ. « ChaosNet : A chaos based artificial neural network architecture for classification ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 29.11 (2019), p. 113125.
- [236] Kyoung-jae KIM et Won Boo LEE. « Stock market prediction using artificial neural networks with optimal feature transformation ». In : *Neural computing & applications* 13 (2004), p. 255-260.
- [237] Klaus-Robert MÜLLER, Matthias KRAUEDAT, Guido DORNHEGE, Gabriel CURIO et Benjamin BLANKERTZ. « Machine learning techniques for brain-computer interfaces ». In : *Biomed. Tech* 49.1 (2004), p. 11-22.
- [238] Aravind GANAPATHIRAJU, Jonathan E HAMAKER et Joseph PICONE. « Applications of support vector machines to speech recognition ». In : *IEEE transactions on signal processing* 52.8 (2004), p. 2348-2355.
- [239] Young Ho KIM et Frank L LEWIS. « Neural network output feedback control of robot manipulators ». In : *IEEE Transactions on robotics and automation* 15.2 (1999), p. 301-309.
- [240] Takehisa YAIRI, Yoshinobu KAWAHARA, Ryohei FUJIMAKI, Yuichi SATO et Kazuo MACHIDA. « Telemetry-mining : a machine learning approach to anomaly detection and fault diagnosis for space systems ». In : *2nd IEEE International Conference on Space Mission Challenges for Information Technology (SMC-IT'06)*. IEEE. 2006, 8-pp.
- [241] Jerome R BELLEGARDA. « Unsupervised document clustering using multi-resolution latent semantic density analysis ». In : *2010 IEEE International Workshop on Machine Learning for Signal Processing*. IEEE. 2010, p. 361-366.
- [242] Gordon M SHEPHERD. *The synaptic organization of the brain*. Oxford university press, 2004.
- [243] Henryk SKOLIMOWSKI. *The theatre of the mind : Evolution in the sensitive cosmos*. Quest Books, 1984.
- [244] Simon HAYKIN. *Neural networks : a comprehensive foundation*. Prentice Hall PTR, 1998.
- [245] MF BEAR, BW CONNORS et MA PARADISO. « Synaptic transmission ». In : *Neuroscience Exploring the Brain, 3rd edn*. Lippincott Williams & Wilkins : Baltimore, MD (2007), p. 102-166.

- [246] Wulfram GERSTNER et Werner M KISTLER. *Spiking neuron models : Single neurons, populations, plasticity*. Cambridge university press, 2002.
- [247] Jürgen SCHMIDHUBER. « Deep learning in neural networks : An overview ». In : *Neural networks* 61 (2015), p. 85-117.
- [248] Danilo MANDIC et Jonathon CHAMBERS. *Recurrent neural networks for prediction : learning algorithms, architectures and stability*. Wiley, 2001.
- [249] Robert CALLAN. *Essence of neural networks*. Prentice Hall PTR, 1998.
- [250] Yilei WU, Qing SONG et Xulei YANG. « Robust recurrent neural network control of biped robot ». In : *Journal of Intelligent and Robotic Systems* 49 (2007), p. 151-169.
- [251] Alex GRAVES, Marcus LIWICKI, Santiago FERNÁNDEZ, Roman BERTOLAMI, Horst BUNKE et Jürgen SCHMIDHUBER. « A novel connectionist system for unconstrained handwriting recognition ». In : *IEEE transactions on pattern analysis and machine intelligence* 31.5 (2008), p. 855-868.
- [252] Thanasis G BARBOUNIS, John B THEOCHARIS, Minas C ALEXIADIS et Petros S DOKOPOULOS. « Long-term wind speed and power forecasting using local recurrent neural network models ». In : *IEEE Transactions on Energy Conversion* 21.1 (2006), p. 273-284.
- [253] J VERMAAK et EC BOTHA. « Recurrent neural networks for short-term load forecasting ». In : *IEEE Transactions on Power Systems* 13.1 (1998), p. 126-132.
- [254] Abdul Manan AHMAD, Saliza ISMAIL et DF SAMAON. « Recurrent neural network with backpropagation through time for speech recognition ». In : *IEEE International Symposium on Communications and Information Technology, 2004. ISCIT 2004*. T. 1. IEEE. 2004, p. 98-102.
- [255] Ken-ichi FUNAHASHI et Yuichi NAKAMURA. « Approximation of dynamical systems by continuous time recurrent neural networks ». In : *Neural networks* 6.6 (1993), p. 801-806.
- [256] Mantas LUKOŠEVIČIUS et Herbert JAEGER. « Reservoir computing approaches to recurrent neural network training ». In : *Computer science review* 3.3 (2009), p. 127-149.
- [257] David VERSTRAETEN, Benjamin SCHRAUWEN, Michiel D’HAENE et Dirk STROOBANDT. « An experimental unification of reservoir computing methods ». In : *Neural networks* 20.3 (2007), p. 391-403.
- [258] Herbert JAEGER. « The “echo state” approach to analysing and training recurrent neural networks-with an erratum note ». In : *Bonn, Germany : German National Research Center for Information Technology GMD Technical Report 148.34* (2001), p. 13.
- [259] Kohei NAKAJIMA et Ingo FISCHER. *Reservoir computing*. Springer, 2021.

- [260] Lyudmila GRIGORYEVA et Juan-Pablo ORTEGA. « Universal discrete-time reservoir computers with stochastic inputs and linear readouts using non-homogeneous state-affine systems ». In : *Journal of Machine Learning Research* 19 (2018), p. 1-40.
- [261] Lyudmila GRIGORYEVA et Juan-Pablo ORTEGA. « Echo state networks are universal ». In : *Neural Networks* 108 (2018), p. 495-508.
- [262] Gouhei TANAKA, Toshiyuki YAMANE, Jean Benoit HÉROUX, Ryosho NAKANE, Naoki KANAZAWA, Seiji TAKEDA, Hidetoshi NUMATA, Daiju NAKANO et Akira HIROSE. « Recent advances in physical reservoir computing : A review ». In : *Neural Networks* 115 (2019), p. 100-123.
- [263] Matthew DALE, Susan STEPNEY, Julian F MILLER et Martin TREFZER. « Reservoir computing in materio : a computational framework for in materio computing ». In : *2017 international joint conference on neural networks (IJCNN)*. IEEE. 2017, p. 2178-2185.
- [264] Andrew ADAMATZKY. « Towards fungal computer ». In : *Interface focus* 8.6 (2018), p. 20180029.
- [265] Lars BÜSING, Benjamin SCHRAUWEN et Robert LEGENSTEIN. « Connectivity, dynamics, and memory in reservoir computing with binary and analog neurons ». In : *Neural computation* 22.5 (2010), p. 1272-1311.
- [266] A ESPOSITO, M MARINARO et G CHOLLET. « Nonlinear speech modeling and applications ». In : *Some Notes on Nonlinearities of Speech*. T. 3445. Springer, 2005, p. 1-14.
- [267] Herbert JAEGER, Mantas LUKOŠEVIČIUS, Dan POPOVICI et Udo SIEWERT. « Optimization and applications of echo state networks with leaky-integrator neurons ». In : *Neural networks* 20.3 (2007), p. 335-352.
- [268] Fabian TRIEFENBACH, Azarakhsh JALALVAND, Benjamin SCHRAUWEN et Jean-Pierre MARTENS. « Phoneme recognition with large hierarchical reservoirs ». In : *Advances in neural information processing systems* 23 (2010).
- [269] Sharan NARANG, Erich ELSÉN, Gregory DIAMOS et Shubho SENGUPTA. « Exploring sparsity in recurrent neural networks ». In : *arXiv preprint arXiv :1704.05119* (2017).
- [270] Guillaume BELLEC, David KAPPEL, Wolfgang MAASS et Robert LEGENSTEIN. « Deep rewiring : Training very sparse deep networks ». In : *arXiv preprint arXiv :1711.05136* (2017).
- [271] Ashok LITWIN-KUMAR, Kameron Decker HARRIS, Richard AXEL, Haim SOMPOLINSKY et LF ABBOTT. « Optimal degrees of synaptic connectivity ». In : *Neuron* 93.5 (2017), p. 1153-1164.
- [272] Claudio GALLICCHIO. « Sparsity in reservoir computing neural networks ». In : *2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*. IEEE. 2020, p. 1-7.

- [273] L da F COSTA, Francisco A RODRIGUES, Gonzalo TRAVIESO et Paulino Ribeiro VILLAS BOAS. « Characterization of complex networks : A survey of measurements ». In : *Advances in physics* 56.1 (2007), p. 167-242.
- [274] Ingrid BOAS. « Social networking in a digital and mobile world : the case of environmentally-related migration in Bangladesh ». In : *Journal of Ethnic and Migration Studies* 46.7 (2020), p. 1330-1347.
- [275] Gueorgi KOSSINETTS et Duncan J WATTS. « Empirical analysis of an evolving social network ». In : *science* 311.5757 (2006), p. 88-90.
- [276] Steven H STROGATZ. « Exploring complex networks ». In : *nature* 410.6825 (2001), p. 268-276.
- [277] Paul ERDŐS et Alfréd RÉNYI. « On the strength of connectedness of a random graph ». In : *Acta Mathematica Hungarica* 12.1 (1961), p. 261-267.
- [278] Albert-László BARABÁSI, Réka ALBERT et Hawoong JEONG. « Mean-field theory for scale-free random networks ». In : *Physica A : Statistical Mechanics and its Applications* 272.1-2 (1999), p. 173-187.
- [279] Duncan J WATTS et Steven H STROGATZ. « Collective dynamics of ‘small-world’ networks ». In : *nature* 393.6684 (1998), p. 440-442.
- [280] Michael STOBBS, Joshua M PETERSON, Borbala MAZZAG et Ethan GAHTAN. « Graph theoretical model of a sensorimotor connectome in zebrafish ». In : *PLoS One* 7.5 (2012), e37292.
- [281] Santo FORTUNATO. « Community detection in graphs ». In : *Physics reports* 486.3-5 (2010), p. 75-174.
- [282] Mark EJ NEWMAN et Michelle GIRVAN. « Finding and evaluating community structure in networks ». In : *Physical review E* 69.2 (2004), p. 026113.
- [283] Sofia GARCIA-SANJUAN, Manuel LILLO-CRESPO, Ángela SANJUÁN-QUILES, Diana GIL-GONZÁLEZ et Miguel RICHART-MARTINEZ. « Life experiences of people affected by Crohn’s disease and their support networks : scoping review ». In : *Clinical Nursing Research* 25.1 (2016), p. 79-99.
- [284] Thiago Christiano SILVA et Liang ZHAO. « Network-based high level data classification ». In : *IEEE Transactions on Neural Networks and Learning Systems* 23.6 (2012), p. 954-970.
- [285] Matthew DALE, Simon O’KEEFE, Angelika SEBALD, Susan STEPNEY et Martin A TREFZER. « Reservoir computing quality : connectivity and topology ». In : *Natural Computing* 20 (2021), p. 205-216.
- [286] Yuji KAWAI, Jihoon PARK et Minoru ASADA. « A small-world topology enhances the echo state property and signal propagation in reservoir computing ». In : *Neural Networks* 112 (2019), p. 15-23.

- [287] Laercio de OLIVEIRA JUNIOR, Florian STELZER et Liang ZHAO. « Clustered and deep echo state networks for signal noise reduction ». In : *Machine Learning* 111.8 (2022), p. 2885-2904.
- [288] Ensieh NAJIBI et Habib ROSTAMI. « SCESN, SPESN, SWESN : Three recurrent neural echo state networks with clustered reservoirs for prediction of nonlinear and chaotic time series ». In : *Applied Intelligence* 43 (2015), p. 460-472.
- [289] Laércio Oliveira JUNIOR, Florian STELZER et Liang ZHAO. « Clustered echo state networks for signal observation and frequency filtering ». In : *Anais do VIII symposium on knowledge discovery, mining and learning*. SBC. 2020, p. 25-32.
- [290] X DUTOIT. « Reservoir Computing for Intelligent Mobile Systems ». Thèse de doct. PhD thesis, Department of Mechanical Engineering, Katholieke Universiteit ..., 2009.
- [291] David VERSTRAETEN, Joni DAMBRE, Xavier DUTOIT et Benjamin SCHRAUWEN. « Memory versus non-linearity in reservoirs ». In : *The 2010 international joint conference on neural networks (IJCNN)*. IEEE. 2010, p. 1-8.
- [292] Benjamin SCHRAUWEN, Jeroen DEFOUR, David VERSTRAETEN et Jan VAN CAMPENHOUT. « The introduction of time-scales in reservoir computing, applied to isolated digits recognition ». In : *Artificial Neural Networks-ICANN 2007 : 17th International Conference, Porto, Portugal, September 9-13, 2007, Proceedings, Part I 17*. Springer. 2007, p. 471-479.
- [293] David VERSTRAETEN. « Reservoir Computing : computation with dynamical systems ». Thèse de doct. Ghent University, 2009.
- [294] Francis WYFFELS, Benjamin SCHRAUWEN et Dirk STROOBANDT. « Stable output feedback in reservoir computing using ridge regression ». In : *Artificial Neural Networks-ICANN 2008 : 18th International Conference, Prague, Czech Republic, September 3-6, 2008, Proceedings, Part I 18*. Springer. 2008, p. 808-817.
- [295] Benjamin SCHRAUWEN, Marion WARDERMANN, David VERSTRAETEN, Jochen J STEIL et Dirk STROOBANDT. « Improving reservoirs using intrinsic plasticity ». In : *Neurocomputing* 71.7-9 (2008), p. 1159-1171.
- [296] Francis WYFFELS et Benjamin SCHRAUWEN. « A comparative study of reservoir computing strategies for monthly time series prediction ». In : *Neurocomputing* 73.10-12 (2010), p. 1958-1964.
- [297] Robert LEGENSTEIN et Wolfgang MAASS. « What makes a dynamical system computationally powerful ». In : *New directions in statistical signal processing : From systems to brain* (2007), p. 127-154.
- [298] Herbert JAEGER. « Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the "echo state network" approach ». In : (2002).
- [299] Chrisantha FERNANDO et Sampsa SOJAKKA. « Pattern recognition in a bucket ». In : *Advances in Artificial Life : 7th European Conference, ECAL 2003, Dortmund, Germany, September 14-17, 2003. Proceedings 7*. Springer. 2003, p. 588-597.

- [300] Kristof VANDOORNE, Wouter DIERCKX, Benjamin SCHRAUWEN, David VERSTRAETEN, Roel BAETS, Peter BIENSTMAN et Jan VAN CAMPENHOUT. « Toward optical signal processing using photonic reservoir computing ». In : *Optics express* 16.15 (2008), p. 11182-11192.
- [301] Matteo CUCCHI, Christopher GRUENER, Lautaro PETRAUSKAS, Peter STEINER, Hsin TSENG, Axel FISCHER, Bogdan PENKOVSKY, Christian MATTHUS, Peter BIRKHOLZ, Hans KLEEMANN et al. « Reservoir computing with biocompatible organic electrochemical networks for brain-inspired biosignal classification ». In : *Science Advances* 7.34 (2021), eabh0693.
- [302] Ken CALUWAERTS, Jérémie DESPRAZ, Atıl IŞÇEN, Andrew P SABELHAUS, Jonathan BRUCE, Benjamin SCHRAUWEN et Vytas SUNSPIRAL. « Design and control of compliant tensegrity robots through simulation and hardware validation ». In : *Journal of the royal society interface* 11.98 (2014), p. 20140520.
- [303] Chao DU, Fuxi CAI, Mohammed A ZIDAN, Wen MA, Seung Hwan LEE et Wei D LU. « Reservoir computing using dynamic memristors for temporal information processing ». In : *Nature communications* 8.1 (2017), p. 2204.
- [304] Lennert APPELTANT, Miguel Cornelles SORIANO, Guy VAN DER SANDE, Jan DANCKAERT, Serge MASSAR, Joni DAMBRE, Benjamin SCHRAUWEN, Claudio R MIRASSO et Ingo FISCHER. « Information processing using a single dynamical node as complex system ». In : *Nature communications* 2.1 (2011), p. 468.
- [305] Ozgur YILMAZ. « Reservoir computing using cellular automata ». In : *arXiv preprint arXiv :1410.0162* (2014).
- [306] Jean C COULOMBE, Mark CA YORK et Julien SYLVESTRE. « Computing with networks of nonlinear mechanical oscillators ». In : *PloS one* 12.6 (2017), e0178663.
- [307] Matthew DALE, Julian F MILLER, Susan STEPNEY et Martin A TREFZER. « Evolving carbon nanotube reservoir computers ». In : *Unconventional Computation and Natural Computation : 15th International Conference, UCNC 2016, Manchester, UK, July 11-15, 2016, Proceedings 15*. Springer. 2016, p. 49-61.
- [308] Matthew DALE, Susan STEPNEY, Julian F MILLER et Martin TREFZER. « Reservoir computing in materio : An evaluation of configuration through evolution ». In : *2016 IEEE symposium series on computational intelligence (SSCI)*. IEEE. 2016, p. 1-8.
- [309] Keisuke FUJII et Kohei NAKAJIMA. « Harnessing disordered-ensemble quantum dynamics for machine learning ». In : *Physical Review Applied* 8.2 (2017), p. 024030.
- [310] Jacob TORREJON, Mathieu RIOU, Flavio Abreu ARAUJO, Sumito TSUNEGI, Guru KHALSA, Damien QUERLIOZ, Paolo BORTOLOTTI, Vincent CROS, Kay YAKUSHIJI, Akio FUKUSHIMA et al. « Neuromorphic computing with nanoscale spintronic oscillators ». In : *Nature* 547.7664 (2017), p. 428-431.

- [311] Kristof VANDOORNE, Pauline MECHEZ, Thomas VAN VAERENBERGH, Martin FIERS, Geert MORTHER, David VERSTRAETEN, Benjamin SCHRAUWEN, Joni DAMBRE et Peter BIENSTMAN. « Experimental demonstration of reservoir computing on a silicon photonics chip ». In : *Nature communications* 5.1 (2014), p. 3541.
- [312] Kristof VANDOORNE, Joni DAMBRE, David VERSTRAETEN, Benjamin SCHRAUWEN et Peter BIENSTMAN. « Parallel reservoir computing using optical amplifiers ». In : *IEEE transactions on neural networks* 22.9 (2011), p. 1469-1481.
- [313] Daniel BRUNNER et Ingo FISCHER. « Reconfigurable semiconductor laser networks based on diffractive coupling ». In : *Optics letters* 40.16 (2015), p. 3854-3857.
- [314] Sylvain BARBAY, Robert KUSZELEWICZ et Alejandro M YACOMOTTI. « Excitability in a semiconductor laser with saturable absorber ». In : *Optics letters* 36.23 (2011), p. 4476-4478.
- [315] Thomas VAN VAERENBERGH, Martin FIERS, Pauline MECHEZ, Thijs SPUESENS, Rajesh KUMAR, Geert MORTHER, Benjamin SCHRAUWEN, Joni DAMBRE et Peter BIENSTMAN. « Cascadable excitability in microrings ». In : *Optics express* 20.18 (2012), p. 20292-20308.
- [316] Bhavin J SHASTRI, Mitchell A NAHMIA, Alexander N TAIT, Alejandro W RODRIGUEZ, Ben WU et Paul R PRUCNAL. « Spike processing with a graphene excitable laser ». In : *Scientific reports* 6.1 (2016), p. 1-12.
- [317] Yvan PAQUOT, Joni DAMBRE, Benjamin SCHRAUWEN, Marc HAELTERMAN et Serge MASSAR. « Reservoir computing : a photonic neural network for information processing ». In : *Nonlinear optics and applications IV*. T. 7728. SPIE. 2010, p. 58-69.
- [318] Laurent LARGER, Miguel C SORIANO, Daniel BRUNNER, Lennert APPELTANT, Jose M GUTIÉRREZ, Luis PESQUERA, Claudio R MIRASSO et Ingo FISCHER. « Photonic information processing beyond Turing : an optoelectronic implementation of reservoir computing ». In : *Optics express* 20.3 (2012), p. 3241-3249.
- [319] Yvan PAQUOT, Francois DUPORT, Antoneo SMERIERI, Joni DAMBRE, Benjamin SCHRAUWEN, Marc HAELTERMAN et Serge MASSAR. « Optoelectronic reservoir computing ». In : *Scientific reports* 2.1 (2012), p. 287.
- [320] K IKEDA, H DAIDO et O AKIMOTO. « Optical turbulence : chaotic behavior of transmitted light from a ring cavity ». In : *Physical Review Letters* 45.9 (1980), p. 709.
- [321] Jean-Pierre GOEDGEBUER, Laurent LARGER, Henri PORTE et Franck DELORME. « Chaos in wavelength with a feedback tunable laser diode ». In : *Physical Review E* 57.3 (1998), p. 2795.
- [322] Thomas L CARROLL. « Using reservoir computers to distinguish chaotic signals ». In : *Physical Review E* 98.5 (2018), p. 052209.
- [323] Sanjukta KRISHNAGOPAL, Michelle GIRVAN, Edward OTT et Brian R HUNT. « Separation of chaotic signals by reservoir computing ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 30.2 (2020).

- [324] Kengo NAKAI et Yoshitaka SAIKI. « Machine-learning inference of fluid variables from data using reservoir computing ». In : *Physical Review E* 98.2 (2018), p. 023111.
- [325] Roland S ZIMMERMANN et Ulrich PARLITZ. « Observing spatio-temporal dynamics of excitable media using reservoir computing ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 28.4 (2018).
- [326] Shahrokh SHAHI, Flavio H FENTON et Elizabeth M CHERRY. « Prediction of chaotic time series using recurrent neural networks and reservoir computing techniques : A comparative study ». In : *Machine learning with applications* 8 (2022), p. 100300.
- [327] Katsuhiko OGATA. *Modern control engineering, prectice-hall*. 1990.
- [328] Robert HERMANN et Arthur KRENER. « Nonlinear controllability and observability ». In : *IEEE Transactions on automatic control* 22.5 (1977), p. 728-740.
- [329] Martijn P VAN DEN HEUVEL, Edward T BULLMORE et Olaf SPORNS. « Comparative connectomics ». In : *Trends in cognitive sciences* 20.5 (2016), p. 345-361.
- [330] Jonathan J CROFTS, Michael FORRESTER, Stephen COOMBES et Reuben D O'DEA. « Structure-function clustering in weighted brain networks ». In : *Scientific Reports* 12.1 (2022), p. 16793.
- [331] Danielle Smith BASSETT et ED BULLMORE. « Small-world brain networks ». In : *The neuroscientist* 12.6 (2006), p. 512-523.
- [332] Marcus KAISER et Claus C HILGETAG. « Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems ». In : *PLoS computational biology* 2.7 (2006), e95.
- [333] Olaf SPORNS et Jonathan D ZWI. « The small world of the cerebral cortex ». In : *Neuroinformatics* 2 (2004), p. 145-162.
- [334] Maxwell A BERTOLERO, BT Thomas YEO et Mark D'ESPOSITO. « The modular and integrative functional architecture of the human brain ». In : *Proceedings of the National Academy of Sciences* 112.49 (2015), E6798-E6807.
- [335] Richard F BETZEL, John D MEDAGLIA, Lia PAPADOPOULOS, Graham L BAUM, Ruben GUR, Raquel GUR, David ROALF, Theodore D SATTERTHWAITTE et Danielle S BASSETT. « The modular organization of human anatomical brain networks : Accounting for the cost of wiring ». In : *Network Neuroscience* 1.1 (2017), p. 42-68.
- [336] Zhang J CHEN, Yong HE, Pedro ROSA-NETO, Jurgen GERMANN et Alan C EVANS. « Revealing modular architecture of human brain structural networks by using cortical thickness from MRI ». In : *Cerebral cortex* 18.10 (2008), p. 2374-2381.
- [337] Emma K TOWLSON, Petra E VÉRTES, Sebastian E AHNERT, William R SCHAFER et Edward T BULLMORE. « The rich club of the C. elegans neuronal connectome ». In : *Journal of Neuroscience* 33.15 (2013), p. 6380-6387.
- [338] Martijn P VAN DEN HEUVEL, René S KAHN, Joaquín GOÑI et Olaf SPORNS. « High-cost, high-capacity backbone for global brain communication ». In : *Proceedings of the National Academy of Sciences* 109.28 (2012), p. 11372-11377.

- [339] Sarah JARVIS, Stefan ROTTER et Ulrich EGERT. « Extending stability through hierarchical clusters in echo state networks ». In : *Frontiers in neuroinformatics* 4 (2010), p. 11.
- [340] Zhidong DENG et Yi ZHANG. « Collective behavior of a small-world recurrent neural system with scale-free distribution ». In : *IEEE Transactions on neural networks* 18.5 (2007), p. 1364-1375.
- [341] Fabrizio DAMICELLI, Claus C HILGETAG et Alexandros GOULAS. « Brain connectivity meets reservoir computing ». In : *PLoS Computational Biology* 18.11 (2022), e1010639.
- [342] Christopher W LYNN et Danielle S BASSETT. « The physics of brain network structure, function and control ». In : *Nature Reviews Physics* 1.5 (2019), p. 318-332.
- [343] Thomas SCHREIBER. « Measuring information transfer ». In : *Physical review letters* 85.2 (2000), p. 461.
- [344] Terry BOSSOMAIER, Lionel BARNETT, Michael HARRÉ, Joseph T LIZIER, Terry BOSSOMAIER, Lionel BARNETT, Michael HARRÉ et Joseph T LIZIER. *Transfer entropy*. Springer, 2016.
- [345] Raul VICENTE, Michael WIBRAL, Michael LINDNER et Gordon PIPA. « Transfer entropy—a model-free measure of effective connectivity for the neurosciences ». In : *Journal of computational neuroscience* 30.1 (2011), p. 45-67.
- [346] Shinya ITO, Michael E HANSEN, Randy HEILAND, Andrew LUMSDAINE, Alan M LITKE et John M BEGGS. « Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model ». In : *PloS one* 6.11 (2011), e27431.
- [347] Oliver OBST, Joschka BOEDECKER et Minoru ASADA. « Improving recurrent neural network performance using transfer entropy ». In : *Neural Information Processing. Models and Applications : 17th International Conference, ICONIP 2010, Sydney, Australia, November 22-25, 2010, Proceedings, Part II 17*. Springer. 2010, p. 193-200.
- [348] Joseph T LIZIER, Siddharth PRITAM et Mikhail PROKOPENKO. « Information dynamics in small-world Boolean networks ». In : *Artificial life* 17.4 (2011), p. 293-314.
- [349] Joseph T LIZIER, Fatihcan M ATAY et Jürgen JOST. « Information storage, loop motifs, and clustered structure in complex networks ». In : *Physical Review E* 86.2 (2012), p. 026110.
- [350] Daniel J AMIT et Daniel J AMIT. *Modeling brain function : The world of attractor neural networks*. Cambridge university press, 1989.
- [351] Philip W ANDERSON. « More Is Different : Broken symmetry and the nature of the hierarchical structure of science. » In : *Science* 177.4047 (1972), p. 393-396.
- [352] Hermann HAKEN, JA Scott KELSO et Heinz BUNZ. « A theoretical model of phase transitions in human hand movements ». In : *Biological cybernetics* 51.5 (1985), p. 347-356.

- [353] John J HOPFIELD. « Neural networks and physical systems with emergent collective computational abilities. » In : *Proceedings of the national academy of sciences* 79.8 (1982), p. 2554-2558.
- [354] Kim SNEPPEN. *Models of life*. Cambridge University Press, 2014.
- [355] Paul M CHAIKIN, Tom C LUBENSKY et Thomas A WITTEN. *Principles of condensed matter physics*. T. 10. Cambridge university press Cambridge, 1995.
- [356] H Eugene STANLEY. *Phase transitions and critical phenomena*. T. 7. Clarendon Press, Oxford, 1971.
- [357] Per BAK. *How nature works : the science of self-organized criticality*. Springer Science & Business Media, 2013.
- [358] John M BEGGS. « The criticality hypothesis : how local cortical networks might optimize information processing ». In : *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences* 366.1864 (2008), p. 329-343.
- [359] Dante R CHIALVO. « Emergent complex neural dynamics ». In : *Nature physics* 6.10 (2010), p. 744-750.
- [360] Dietmar PLENZ. « Criticality in cortex : neuronal avalanches and coherence potentials ». In : *Criticality in neural systems* (2014), p. 5-42.
- [361] Alan Mathison TURING. « Mind ». In : *Mind* 59.236 (1950), p. 433-460.
- [362] Luca COCCHI, Leonardo L GOLLO, Andrew ZALESKY et Michael BREAKSPEAR. « Criticality in the brain : A synthesis of neurobiology, models and cognition ». In : *Progress in neurobiology* 158 (2017), p. 132-152.
- [363] Gustavo DECO, Giulio TONONI, Melanie BOLY et Morten L KRINGELBACH. « Rethinking segregation and integration : contributions of whole-brain modelling ». In : *Nature Reviews Neuroscience* 16.7 (2015), p. 430-439.
- [364] Mike LI, Yinuo HAN, Matthew J ABURN, Michael BREAKSPEAR, Russell A POLDRACK, James M SHINE et Joseph T LIZIER. « Transitions in information processing dynamics at the whole-brain network level are driven by alterations in neural gain ». In : *PLoS computational biology* 15.10 (2019), e1006957.
- [365] Lawrence K SAUL, Yair WEISS et Léon BOTTOU. *Advances in neural information processing systems 17 : proceedings of the 2004 conference*. T. 17. MIT press, 2005.
- [366] Marouane AYYAD et Saliya COULIBALY. « The cellular automata inside optical chimera states ». In : *Chaos, Solitons & Fractals* 153 (2021), p. 111524.
- [367] Annick LESNE et Michel LAGUËS. *Scale invariance : From phase transitions to turbulence*. Springer Science & Business Media, 2011.
- [368] T GISIGER. « Scale invariance in biology : coincidence or footprint of a universal mechanism? » In : *Biological Reviews* 76.2 (2001), p. 161-209.
- [369] HE STANLEY, LAN AMARAL, P GOPIKRISHNAN et V PLEROU. « Scale invariance and universality of economic fluctuations ». In : *Physica A : Statistical Mechanics and its Applications* 283.1-2 (2000), p. 31-41.

- [370] S NINAGAWA. « Dynamics of universal computation and $1/f$ noise in elementary cellular automata ». In : *Chaos, Solitons & Fractals* 70 (2015), p. 42-48.
- [371] Benoit MANDELBROT. « How long is the coast of Britain? Statistical self-similarity and fractional dimension ». In : *science* 156.3775 (1967), p. 636-638.
- [372] Ben POOLE, Subhaneil LAHIRI, Maithra RAGHU, Jascha SOHL-DICKSTEIN et Surya GANGULI. « Exponential expressivity in deep neural networks through transient chaos ». In : *Advances in neural information processing systems* 29 (2016).
- [373] Christophe ELOY. « Leonardo's rule, self-similarity, and wind-induced stresses in trees ». In : *Physical review letters* 107.25 (2011), p. 258101.
- [374] Vijay SHARMA. « Deterministic chaos and fractal complexity in the dynamics of cardiovascular behavior : perspectives on a new frontier ». In : *The Open Cardiovascular Medicine Journal* 3 (2009), p. 110.
- [375] Michael F SHLESINGER et Bruce J WEST. « Complex fractal dimension of the bronchial tree ». In : *Physical review letters* 67.15 (1991), p. 2106.
- [376] Elaine Nicpon MARIEB et Katja HOEHN. *Human anatomy & physiology*. Pearson education, 2007.
- [377] Ewald R WEIBEL. « Fractal geometry : a design principle for living organisms ». In : *American Journal of Physiology-Lung Cellular and Molecular Physiology* 261.6 (1991), p. L361-L369.
- [378] Julian H SMITH, Conor ROWLAND, B HARLAND, S MOSLEHI, RD MONTGOMERY, K SCHOBERT, WJ WATTERSON, J DALRYMPLE-ALFORD et RP TAYLOR. « How neurons exploit fractal geometry to optimize their network connectivity ». In : *Scientific reports* 11.1 (2021), p. 2332.
- [379] Thomas F VARLEY, Michael CRAIG, Ram ADAPA, Paola FINOIA, Guy WILLIAMS, Judith ALLANSON, John PICKARD, David K MENON et Emmanuel A STAMATAKIS. « Fractal dimension of cortical functional connectivity networks & severity of disorders of consciousness ». In : *PloS One* 15.2 (2020), e0223812.
- [380] Conor ROWLAND, Julian H SMITH, Saba MOSLEHI, Bruce HARLAND, John DALRYMPLE-ALFORD et Richard P TAYLOR. « Neuron arbor geometry is sensitive to the limited-range fractal properties of their dendrites ». In : *Frontiers in Network Physiology* 3 (2023), p. 1072815.
- [381] Andrea I LUPPI, Michael M CRAIG, Peter COPPOLA, Alexander RD PEATTIE, Paola FINOIA, Guy B WILLIAMS, Judith ALLANSON, John D PICKARD, David K MENON et Emmanuel A STAMATAKIS. « Preserved fractal character of structural brain networks is associated with covert consciousness after severe brain injury ». In : *NeuroImage : Clinical* 30 (2021), p. 102682.
- [382] Stephen WOLFRAM. « Statistical mechanics of cellular automata ». In : *Reviews of modern physics* 55.3 (1983), p. 601.

- [383] Alan WOLF, Jack B SWIFT, Harry L SWINNEY et John A VASTANO. « Determining Lyapunov exponents from a time series ». In : *Physica D : nonlinear phenomena* 16.3 (1985), p. 285-317.
- [384] Jaideep PATHAK, Zhixin LU, Brian R HUNT, Michelle GIRVAN et Edward OTT. « Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 27.12 (2017).
- [385] Otto E RÖSSLER. « An equation for continuous chaos ». In : *Physics Letters A* 57.5 (1976), p. 397-398.
- [386] Venkata Anirudh PAMMI et Sylvain BARBAY. « Micro-lasers for neuromorphic computing ». In : *Photoniques* 104 (2020), p. 26-29.
- [387] Leandro M ALONSO. « Complex behavior in chains of nonlinear oscillators ». In : *Chaos : An Interdisciplinary Journal of Nonlinear Science* 27.6 (2017).
- [388] Lazaros K GALLOS, Chaoming SONG et Hernán A MAKSE. « A review of fractality and self-similarity in complex networks ». In : *Physica A : Statistical Mechanics and its Applications* 386.2 (2007), p. 686-691.
- [389] Jian LI, Qian DU et Caixin SUN. « An improved box-counting method for image fractal dimension estimation ». In : *Pattern recognition* 42.11 (2009), p. 2460-2469.

Mesures statistiques

Ici, nous présentons brièvement les mesures statistiques utilisées dans notre étude :

A.1 Local Active information storage

Pour quantifier le stockage d'information présent dans les séries temporelles, LIZIER a introduit une mesure appelée stockage actif local d'informations (**AIS**) [115]. En bref, cette mesure quantifie la quantité d'informations dans un échantillon d'une série temporelle neuronale qui est prévisible à partir de son passé et qui a donc été stockée dans ce passé. Cela se fait en calculant simplement les informations mutuelles locales entre le passé d'un signal neuronal et son prochain échantillon à chaque instant, et pour chaque canal d'un enregistrement.

Nous désignerons les variables aléatoires par X, Y, \dots , et laisserons x_i, y_i, \dots représenter le i -ième pas de temps d'une série temporelle tirée d'une variable aléatoire. De nombreuses mesures prennent en compte les k -historiques (c'est-à-dire les k -blocs) de la série chronologique, par exemple : $x_i^k = (x_{i-k+1}, x_{i-k+2}, \dots, x_i)$.

Généralement, on écrira $p(x_i)$, $p(x_i^k)$ et $p(x_i^k, x_{i+1})$ pour dénoter la probabilité empirique d'observer l'état x_i , les x_i^k , k -historique, et la probabilité conjointe d'observer (x_i^k, x_{i+1}) . AIS d'une variable temporelle est donnée par :

$$a_{X,i}(k) = \log_2 \frac{p(x_i^k, x_{i+1})}{p(x_i^k)p(x_{i+1})} \tag{A.1}$$

où les probabilités sont construites empiriquement à partir de l'ensemble de la série chronologique. De la variante locale, les informations actives temporellement globales telles que :

$$A_X(k) = \langle a_{X,i}(k) \rangle_i = \sum_{x_i^k, x_{i+1}} p(x_i^k, x_{i+1}) \log_2 \frac{p(x_i^k, x_{i+1})}{p(x_i^k)p(x_{i+1})} \tag{A.2}$$

Au sens strict, les informations actives locales et moyennes sont définies comme :

$$a_{X,i} = \lim_{k \rightarrow \infty} a_{X,i}(k) \quad \text{et} \quad A_X = \lim_{k \rightarrow \infty} A_X(k) \quad (\text{A.3})$$

A.2 Transfer entropy

L'entropie de transfert (**TE**) a été introduite par SCHREIBER [343] pour quantifier le transfert d'informations entre une source d'informations et une destination, en conditionnant les effets d'histoire partagée. TE a été formulée à l'origine en ne considérant que la source et la destination ; cependant, de nombreux systèmes d'intérêt ont plus que ces deux composants. En tant que tel, il peut être nécessaire de conditionner les probabilités sur les états de tous les composants "d'arrière-plan" du système. Ces deux formes sont parfois appelées respectivement entropie de transfert apparente et complète [188].

Cette implémentation de TE permet à l'utilisateur de conditionner les probabilités sur n'importe quel nombre de processus d'arrière-plan, dans les limites matérielles bien sûr. Pour la description suivante, prenez X comme source, Y comme cible et $W = (W_1, \dots, W_l)$ comme processus d'arrière-plan par rapport auxquels nous aimerions conditionner. Par exemple, nous pouvons prendre l'état de deux nœuds dans un réseau dynamique comme source et cible, tandis que tous les autres nœuds du réseau sont traités comme arrière-plan. L'entropie de transfert est alors définie en terme d'une variante temporelle locale :

$$t_{X \rightarrow Y, W, i}(k) = \log_2 \frac{p(y_{i+1}, x_i | y_i^k, W_{(1,i)}, \dots, W_{(l,i)})}{p(y_{i+1} | y_i^k, W_{(1,i)}, \dots, W_{(l,i)}) p(x_i | y_i^k, W_{(1,i)}, \dots, W_{(l,i)})} \quad (\text{A.4})$$

En faisant la moyenne dans le temps, nous avons :

$$T_{X \rightarrow Y, W}(k) = \langle t_{X \rightarrow Y, W, i}(k) \rangle_i \quad (\text{A.5})$$

Comme dans le cas des informations actives (AIS), l'entropie de transfert est formellement définie comme la limite de l'entropie de transfert de k -historique lorsque $k \rightarrow \infty$:

$$t_{X \rightarrow Y, W, i} = \lim_{k \rightarrow \infty} t_{X \rightarrow Y, W, i}(k) \quad \text{et} \quad T_{X \rightarrow Y, W} = \lim_{k \rightarrow \infty} T_{X \rightarrow Y, W}(k). \quad (\text{A.6})$$

A.3 Approximate Entropy

Approximate Entropy (**ApEn**) a été proposée, pour discerner des niveaux de régularité dans des données réelles sans aucune connaissance du système source [197]. Elle est robuste au bruit, ce qui était une limitation centrale des outils précédents développés par la théorie de l'information [199]. Depuis lors, ApEn a été utilisée pour examiner de courts signaux expérimentaux, en particulier des ensembles de données physiologiques et biologiques [199].

Afin de trouver l'entropie approximative d'une série temporelle de données : $u = \{u(1), u(2), \dots, u(N)\}$ de longueur N . Tout d'abord, nous fixons m (dimension plongeante),

un entier non négatif, de sorte que des blocs de vecteurs $(N - m + 1)$ soient formés $x(i) = \{u(i), u(i + 1), \dots, u(i + m - 1)\}$ et $x(j) = \{u(j), u(j + 1), \dots, u(j + m - 1)\}$, et on calcule la distance qui les sépare, donnée par :

$$d[x(i), x(j)] = \max_{k=1,2,\dots,m} (|u(i + k - 1) - u(j + k - 1)|) \quad (\text{A.7})$$

Dans un second temps, on calcule la valeur de

$$C_i^m(r) = \text{nombre de } j \leq N - m + 1 \text{ tel que } d[x(i), x(j)] \leq r / (N - m + 1) \quad (\text{A.8})$$

où r spécifie le niveau de filtrage. Ensuite, on calcule

$$\Phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \log(C_i^m(r)). \quad (\text{A.9})$$

Enfin, cette statistique est définie comme

$$\text{ApEn}(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r) \quad (\text{A.10})$$

A.4 Sample Entropy

ApEn présente une sensibilité majeure aux paramètres d'entrée : m (dimension d'intégration), r (niveau de filtrage) et L (Longueur des données). Pour éviter cet inconvénient principal, l'entropie d'échantillon (**SampEn**) [198] a été introduite pour surmonter la dépendance à la longueur de la série temporelle et maintenir la cohérence relative [198, 199]. Théoriquement, les principales différences entre ApEn et SampEn concernent le calcul des probabilités. En fait, SampEn évite le problème d'auto-comptage et adopte une méthode de sommation différente des correspondances entre les vecteurs modèles.

Pour calculer SampEn d'un ensemble de données de séries chronologiques : Tout d'abord, nous déterminons le nombre total de vecteurs *possibles* pour chaque vecteur modèle et les additionnons en calculant la formule :

$$B^m(r) = \frac{1}{(N - m - 1)(N - m)} \sum_{i=1}^{N-m} \sum_{j=1, j \neq i}^{N-m} \text{nombre de fois que } d(|x_m(j) - x_m(i)|) < r \quad (\text{A.11})$$

De même, nous déterminons le nombre total de *matches* pour chaque vecteur modèle et les additionnons en calculant la formule :

$$A^m(r) = \frac{1}{(N - m - 1)(N - m)} \sum_{i=1}^{N-m} \sum_{j=1, j \neq i}^{N-m} \text{nombre de fois que } d(|x_{m+1}(j) - x_{m+1}(i)|) < r \quad (\text{A.12})$$

Enfin, la valeur de SampEn est estimée par :

$$\text{SampEn}(m, r, N) = -\log(A^m(r)/B^m(r)) \quad (\text{A.13})$$

La combinaison de paramètres typiques ($r = 0,2 * std$, $m = 2$) [197, 198] a été ajustée pour effectuer notre analyse pour ApEn ou SampEn.

A.5 Lempel-Ziv complexity

Un autre outil puissant qui a prouvé sa capacité à mesurer et à caractériser le caractère aléatoire des modèles dynamiques, est la compression des données. En gros, les algorithmes de compression fournissent la longueur de la forme la plus courte qui peut réexprimer les informations "essentiels" dans un message. Le LEMPEL-ZIV complexity (**LZ**), l'un des algorithmes de compression optimale, est utilisé pour estimer le taux d'entropie d'une source ergodique, qui est lié à la valeur asymptotique du taux de croissance du LZ atteint dans la limite de grande longueur de chaîne ($L \gg 1$) [195, 196]. LZ a été utilisé pour la classification de CA [195, 196], et des systèmes dynamiques non linéaires [387].

Soit s une suite binaire de longueur N . La complexité LZ, $C_{LZ}(s)$, de s est définie comme le nombre de facteurs dans son historique exhaustif [195, 196]. Comme nos chaînes binaires ont une grande longueur, nous considérerons la mesure normalisée de complexité donnée par :

$$LZ_{\text{complexité}} = C_{LZ}(s)/(N/\log(N)) \quad (\text{A.14})$$

A.6 Dimension fractale : Méthode de comptage des boîtes

Il existe de nombreuses façons de calculer la propriété de mise à l'échelle des objets naturels et de nombreuses méthodes et algorithmes empiriques ont été suggérés dans la littérature [388]. Le plus courant est la méthode *Box Counting*, qui s'approche de la dimension fractale de HAUSDORFF [389]. L'algorithme du Box Counting donne des indications fiables sur l'existence d'auto-similarité, et peut être facilement mis en œuvre dans des données bidimensionnelles binaires.

La méthode du Box Counting couvre l'image avec une grille, puis nous comptons combien de cases de la grille couvrent une partie de l'image. Ensuite on fait la même chose mais en utilisant une grille plus fine avec des cases plus petites. En réduisant la taille de la grille à plusieurs reprises, nous finissons par capturer avec plus de précision la structure du motif.

En utilisant la méthode du Box Counting, la dimension fractale est définie comme la pente de la ligne lorsque nous traçons la valeur de $\log(N)$ sur l'axe Y par rapport à la valeur de $\log(r)$ sur l'axe X . La même équation est utilisée pour définir la dimension fractale D . Cette fois, N est le nombre de cases qui couvrent le motif et r est le grossissement, ou l'inverse de la taille de la case :

$$D = \frac{\log(N)}{\log(r)} \quad (\text{A.15})$$

La pente de la ligne est égale à la dimension fractale et elle est définie comme la quantité

de changement le long de l'axe Y , divisée par la quantité de changement le long de l'axe X . Les pentes (et les dimensions) varient entre 1 et 2 pour ce type d'analyse, ce qui correspond à la plage entre une ligne droite (dimension = 1) et une ligne si ondulée qu'elle remplit complètement un plan bidimensionnel. Une pente plus raide signifie que l'objet est plus 'fractal', c'est-à-dire qu'il gagne en complexité à mesure que la taille de la boîte diminue. Une pente plus plate et de valeur inférieure signifie que l'objet est plus proche d'une ligne droite, moins 'fractalement' et que la quantité de détails n'augmente pas aussi rapidement avec l'augmentation du grossissement.

Le traitement d'information avec des états chimères optiques

Résumé

Dans la mythologie grecque, une chimère est une créature fantastique dont certaines parties du corps appartiennent à des animaux différents. Par analogie à cette mythologie, en physique et plus particulièrement dans l'étude des systèmes complexes discrets spatialement étendus, ces états chimères correspondent à la coexistence de deux comportements dynamiques spatio-temporels opposés. La coexistence de deux domaines l'un cohérent et l'autre incohérent dans une chaîne d'oscillateurs non-linéaires couplés en est l'exemple historique, à l'image des différentes parties du corps d'une chimère. Ces auto-organisations spatio-temporelles ont été largement étudiées théoriquement et expérimentalement. Cependant, rares sont les études menées pour explorer les liens entre ce type de dynamique et les automates cellulaires. Ces automates, malgré leur simplicité, possèdent des propriétés dynamiques remarquables et, par conséquent, représentent un des socles de la théorie d'information.

Pour répondre à cette problématique, nous avons considéré des états chimères obtenus dans une chaîne de résonateurs optiques identiques couplés. Ces structures ont alors fait l'objet d'analyses quantitatives et qualitatives par les mêmes outils que ceux utilisés pour caractériser les automates cellulaires. Cela nous a permis de mettre en évidence une dynamique de type automate cellulaire élémentaire cachée dans l'évolution de nos états chimères. Nous avons alors été en mesure de déduire, un ensemble de propriétés en terme de calculabilité, ouvrant des perspectives vers des potentielles applications pour le traitement de l'information.

Par la suite, nous avons utilisé nos états chimères optiques dans le cadre des réseaux de neurones récurrents. Il s'agit d'un nouveau paradigme, qui se distingue par sa grande simplicité, sa rapidité ainsi que son efficacité incontournable dans le traitement de l'information. Cependant, les performances de cette technique d'apprentissage automatique, dépendent notamment du design du réservoir. Nos résultats montrent que l'implémentation de nos états chimères optiques au lieu des réservoirs 'classiques', peut fournir une alternative architecturale prometteuse permettant d'améliorer davantage la vitesse du traitement d'information.

Mots clés : chimera states, spatiotemporal complexity, cellular automata, neural network, reservoir computing

Information processing with optical chimera states

Abstract

In Greek mythology, a chimera is a fantastic creature whose body parts belong to different animals. By analogy with this mythology, in physics and more particularly in the study of spatially extended discrete complex systems, these chimera states correspond to the coexistence of two opposing spatio-temporal dynamic behaviors. The coexistence of two domains, one coherent and the other incoherent in a chain of coupled non-linear oscillators is the historical example, like the different parts of the body of a chimera. These spatio-temporal self-organizations have been widely studied theoretically and experimentally. However, few studies have been carried out to explore the links between this type of dynamics and cellular automata. These automata, despite their simplicity, have remarkable dynamic properties and, consequently, represent one of the foundations of information theory.

To answer this problem, we considered chimera states obtained in a chain of identical coupled optical resonators. These structures were then the subject of quantitative and qualitative analyzes using the same tools as those used to characterize cellular automata. This allowed us to highlight an elementary cellular automaton type dynamic hidden in the evolution of our chimera states. We were then able to deduce a set of properties in terms of computability, opening perspectives towards potential applications for information processing.

Subsequently, we used our optical chimera states in the context of recurrent neural networks. This is a new paradigm, which stands out for its great simplicity, speed and essential efficiency in the processing of information. However, the performance of this machine learning technique depends in particular on the design of the reservoir. Our results show that the implementation of our optical chimeric states instead of 'classic' reservoirs can provide a promising architectural alternative to further improve the speed of information processing.

Keywords: états chimères, complexité spatiotemporelle, automates cellulaires, réseau de neurones, calcul par réservoir

Laboratoire de Physique des Lasers, Atomes et Molécules

2 Avenue Jean Perrin – BP 70478 – 59658 Villeneuve d'Ascq Cedex – France