Technische Universität Ilmenau
Fakultät für Mathematik und Naturwissenschaften
Fachgebiet Theoretische Physik 2

Investigating Quantum Evolution Using Krylov Complexity and Its Applications in Quantum Reservoir Computing

Dissertation
zur Erlangung des akademischen Grades
Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt der
Fakultät für Mathematik und Naturwissenschaften
der Technischen Universität Ilmenau

von Herrn
Saud Čindrak
geboren am 23.04.1990 in Bijelo Polje, Montenegro

Betreuerin: Prof. Dr. Kathy Lüdge

Gutachter: Prof. Dr. Jörg Schumacher
Gutachter: Prof. Dr. Christopher Gies

URN: urn:nbn:de:gbv:ilm1-202500026
DOI: 10.22032/dbt.67384
Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung 4.0 International Lizenz.

Anlage 1

Erklärung

Ich versichere, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise entgeltlich/unentgeltlich[1] geholfen:

1.
2.
3.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder anderer Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer Prüfungsbehörde vorgelegt.

Ich bin darauf hingewiesen worden, dass die Unrichtigkeit der vorstehenden Erklärung als Täuschungsversuch bewertet wird und gemäß § 7 Abs. 10 der Promotionsordnung den Abbruch des Promotionsverfahrens zur Folge hat.

(Ort, Datum) *(Unterschrift)*

---

1) Unzutreffendes bitte streichen.

Deutsche Zusammenfassung

Diese Dissertation untersucht das Zusammenspiel von Quantendynamik, Quanteninformationstheorie und Quantum Reservoir Computing (QRC). Im Zentrum steht die Frage, wie sich Quantensysteme zeitlich entwickeln und inwieweit das Verständnis dieser Zeitentwicklung Einblicke in QRC geben kann.

Die Arbeit beginnt mit einem Theoriekapitel, in dem dynamische Systeme, die Postulate der Quantenmechanik sowie Konzepte des Reservoir Computing behandelt werden (**Kapitel** II). Aufbauend darauf richtet sich der Fokus auf das Feld der *Krylov Complexity* in **Kapitel** III. Hier wird untersucht, wie sich zeitentwickelte Zustände oder Observablen im zugehörigen Krylov-Raum verteilen. Ein wesentlicher Beitrag dieser Arbeit besteht darin zu zeigen, dass sich Krylov-Räume durch zeitentwickelte Operatoren oder Zustände erzeugen lassen. Das ist besonders wichtig für QRC, wo der zugrunde liegende Hamiltonian oft nicht bekannt ist. Es wurde außerdem gezeigt, dass solche Krylov-Räume eine geringere Krylov Complexity aufweisen können. Diese Beobachtung stellt die verbreitete Annahme infrage, dass Krylov-Basen aus Potenzen des Hamiltonians minimale Komplexität erzeugen. Um die effektive Dimension der zeitabhängigen Krylov-Räume zu erfassen, werden *Krylov Expressivity* und *Krylov Observability* eingeführt.

In **Kapitel** IV werden diese Maße im Rahmen des Quantum Reservoir Computing untersucht. Dabei zeigt sich, dass *Krylov Observability* die *data expressivity* deutlich besser beschreibt als die *Krylov Spread Complexity*. Dies stützt die Interpretation, dass QRC Daten auf den zugehörigen Krylov-Raum abbildet. Des Weiteren befasst sich dieses Kapitel mit dem Messprozess und dessen Implikationen für QRC. Da jede Messung den Zustand kollabieren lässt, ist eine vollständige Rücksetzung des Systems erforderlich. Um diesem Problem zu begegnen, wird ein modellhafter Ansatz mit beschränktem Gedächtnis vorgestellt, der die natürliche *fading memory* Eigenschaft von Reservoir-Computern ausnutzt. Diese Methode reduziert den Rechenaufwand erheblich und verbessert zugleich die Leistung.

Das abschließende **Kapitel** V widmet sich der Untersuchung von Phasenübergängen in ungeordneten Quantensystemen mithilfe quanteninformativer Maße. *Krylov Complexity* erweist sich hierbei nicht nur als geeignet, um den Übergang von lokalisiertem zu ergodischem Verhalten zu identifizieren, sondern auch, um zwischen unterschiedlichen lokalisierten Phasen zu unterscheiden. Eine charakteristische Zeitgröße, die aus der *Krylov Complexity* abgeleitet wird, korreliert am besten mit der *data expressivity* im QRC.

Zusammenfassend vereint diese Arbeit Ansätze aus der Quanteninformation, der Quantenstatistik, der *Krylov Complexity*, dem maschinellen Lernen, der Theorie der Phasenübergänge und der dynamischen Systeme. Die entwickelten Methoden liefern neue Erkenntnisse über das Verhalten und die Steuerbarkeit quantenmechanischer Systeme und eröffnen vielfältige Perspektiven für weiterführende Forschung in der Quanteninformation.

Abstract

This thesis examines the relationship between quantum dynamics, quantum information theory, and quantum reservoir computing. The work is driven by the goal of understanding how quantum systems evolve and how their behavior over time influences their value for computational tasks.

The study begins with the necessary theoretical background, covering dynamical systems, quantum mechanics, and reservoir computing in Chapter II. The focus then moves to Krylov complexity, a more recent method used to describe the spreading behavior of quantum states and operators within Krylov spaces, detailed in Chapter III. One of the main contributions of this thesis is the reinterpretation of Krylov complexity through the use of time-evolved states and operators. It is demonstrated that Krylov spaces can be built from measurements alone, without prior knowledge of the Hamiltonian. These time-evolved spaces can lead to lower complexity values, challenging previous assumptions that Krylov bases constructed from Hamiltonian powers always result in minimal spreading. From these developments, two new quantities are introduced: Krylov expressivity and Krylov observability. These serve to describe the effective dimensional growth of the time-dependent Krylov spaces.

In Chapter IV, the field of Krylov complexity is applied to the framework of quantum reservoir computing. It is shown that Krylov observability more accurately reflects data expressivity than conventional Krylov complexity. This leads to the interpretation that quantum reservoirs transform data into Krylov space, where observability can act as a useful guide for system design. This chapter also addresses practical challenges that arise during quantum measurement, where the collapse of the state often demands a reset of the system after every step. To mitigate this, a model based on limited memory is proposed. This approach takes advantage of the fading memory of reservoir computers, resulting in both faster computation and improved performance.

The final chapter, Chapter V, explores the use of quantum information measures to study phase transitions in disordered systems. Krylov complexity is found to detect transitions between localized and ergodic phases and also to distinguish between different types of localization, such as spin glass and many-body localized states. A characteristic time scale derived from Krylov complexity is shown to coincide with data expressivity in quantum reservoir computing, pointing to its potential for uncovering structural features in quantum dynamics.

In summary, this thesis brings together ideas from quantum information, quantum statistics, Krylov complexity, machine learning, phase transitions, and dynamical systems. The techniques developed here contribute to a deeper understanding of quantum reservoir computing and provide versatile tools for analyzing the evolution of quantum systems. These results open new possibilities for future research in quantum information science and computational theory.

Contents

List of symbols and abbreviations e

I Chapter: Introduction 3
- I.1 Aspects of Quantum Dynamics . . . 3
- I.2 Quantum Reservoir Computing . . . 6
- I.3 Outline of This Thesis . . . 8

II Chapter: Theory 11
- II.1 Dynamical Systems . . . 13
  - II.1.1 Mathematical Definition of a Dynamical System . . . 13
  - II.1.2 Physical Dynamical Systems and Fixed-Point Analysis . . . 14
  - II.1.3 Linear Systems . . . 16
  - II.1.4 Non-Linear and Chaotic Systems . . . 18
- II.2 Quantum Mechanics . . . 23
  - II.2.1 Postulates of Quantum Mechanics and the Schrödinger Equation 23
  - II.2.2 Evolution of Mixed States and Operators . . . 24
- II.3 Reservoir Computing . . . 27
  - II.3.1 Basics of Neural Networks . . . 27
  - II.3.2 Reservoir Computing . . . 31
  - II.3.3 Information Processing Capacity and prediction tasks . . . 33
- II.4 Quantum reservoir computing . . . 40
  - II.4.1 Quantum Reservoir Computing Scheme . . . 40
  - II.4.2 Motivation of Quantum Reservoir Computing . . . 43

III Chapter: Krylov Complexity in Unitary Time Evolution 47
- III.1 Krylov Spaces for Pure States . . . 50
  - III.1.1 Spread Complexity through the Krylov spaces of the Hamiltonian 52
  - III.1.2 Spread Complexity via Measurable Krylov Spaces . . . 57
  - III.1.3 Krylov Expressivity: Phase Space Measure on the Krylov Space 65
- III.2 Krylov Spaces for Operators . . . 68
  - III.2.1 Operator Complexity through the Powers of the Liouvillian . . . 69
  - III.2.2 Operator Complexity through Time-Evolved Operators . . . 74
  - III.2.3 Capturing the Phase Space Dimension of Operators with Krylov Observability . . . 80

IV Chapter: Quantum Reservoir Computing **83**
IV.1 Summary of Krylov Spaces and Quantum Reservoir Computing 84
IV.2 Quantum Reservoir Computing Maps Data onto the Krylov Space . . . 90
  IV.2.1 Understanding Saturation in Quantum Reservoir Computing . . 92
  IV.2.2 Krylov Observability in Quantum Reservoir Computing 93
  IV.2.3 Quantum Reservoir Computing Maps Data onto the Krylov Space 98
IV.3 Engineering Quantum Reservoir Computers 103
IV.4 Memory-Restricted Quantum Reservoir Computing 112

V Chapter: Probing Quantum Phase Transitions **125**
V.1 Level spacing statistics, Loschmidt Echo and Correlators 127
  V.1.1 Transverse-Field Ising Model 127
  V.1.2 Level Spacing Statistics 128
  V.1.3 Loschmidt Echo 131
  V.1.4 Out-of-Time-Ordered Correlators (OTOCs) 133
V.2 Understanding Phase Transitions using Krylov Complexity 135
  V.2.1 Revisiting Krylov spread Complexity 135
  V.2.2 Spread Complexity at the Spin Glass to Ergodic Transition . . . 137
  V.2.3 Spread Complexity at the MBL to Ergodic Transition 140
  V.2.4 Krylov saturation time 142
V.3 Quantum Reservoir Computing and Quantum Phase Transitions 145
  V.3.1 Conclusion 148

VI Chapter: Summary and Outlook **151**

Anhang

A Loschmidt Echo and Out of Order Time Correlators **155**

B Engineering Quantum Reservoir Computing **161**

List of Figures **CLXV**

List of Tables **CLXVII**

Bibliography **CLXIX**

List of symbols and abbreviations

$G(t)$	Loschmidt Echo
P_C	Pearson Correlation
$\mathbf{S}$	State matrix
$\mathcal{E}_K$	Krylov Expressivity
$\mathcal{K}_O$	Krylov Operator Complexity
$\mathcal{K}_S$	Krylov Spread Complexity
O_K	Krylov observability
K_m	Krylov Space
$\odot$	pointwise multiplication
ρ	Density matrix
ETH	Eigenenergy Thermalization Hypothesis
FFNN	Feed Forward Neural Network
IPC	Information Processing Capacity
LCQA	Linear Complexity Quantum Algorithm
MBL	Many-Body Localization
NRMSE	Normalized Root Mean Squarred Error
OTOC	Out-of-Time Ordered Correlator
OTOC	Out-of-time-ordered correlator
QCQA	Quadratic Complexity Quantum Algorithm
QML	Quantum Machine Learninng
QRC	Quantum Reservoir Computing
RC	Reservoir Computing
RNN	Recurrent Neural Network
SYK	Sachdev-Ye-Kitaev
TFIM	Transverse field Ising model

Acknowledgement

At this point, I would like to thank Professor Kathy Lüdge and Lina Jaurigue for their supervision and support throughout this work. I am also grateful to all members of the research group for creating such a pleasant working and scientific environment. In particular, I thank my office colleagues Jonnel Jaurigue and Alireza Molla Ali Hosseini for lively discussions, entertaining conversations, and a congenial office atmosphere. I would also like to thank Maximilian Zier and our neighbors Luci Fumagalli, Hamid Shahraki, and Niloufar Mashhadi-Gholamali for their friendly collegiality and engaging discussions, as well as the neighbors on the opposite side, Malte Grunert, Max Großmann, Jannis Waldmann, and Marc Leon Thieme, for contributing to a welcoming work climate. Last but not least, I thank Oliver Voigt and Jonas Naujoks for their insightful discussions, unwavering support, and for proofreading this thesis.

I Chapter: Introduction

I.1 Aspects of Quantum Dynamics

The question of how physical systems develop over time is a central theme in physics. From the predictable motions of celestial bodies in classical mechanics to the evolution of quantum states, the pursuit of a unified description of dynamics has guided scientific progress for centuries. In quantum theory, however, the situation is more subtle. Although the time evolution of isolated systems is governed by unitary dynamics, phenomena such as thermalization, chaos, and the behavior of observables remain complex and not yet fully understood.

This dissertation examines quantum dynamics through the lens of complexity and information. At its core, it addresses three intertwined questions: What constitutes complexity in the evolution of a quantum system? How does this complexity relate to the computational power of such systems, especially in the setting of quantum reservoir computing (QRC)? And to what extent can these insights be used to improve the design and understanding of QRC architectures?

Although chaos is a dynamical phenomenon, it has mostly been described using spectral and eigenvector statistics [1, 2, 3]. Recently, it was shown that isolated quantum systems can exhibit thermalization in local observables, as explained by the eigenstate thermalization hypothesis (ETH) [2, 3]. The ETH fails in describing systems exhibiting many body localization (MBL), where strong disorder results in an increase in the number of local integrals of motion, preventing the system from thermalization [4]. Transitions between MBL and thermalization are called dynamical phase transitions [5, 6] and have gained a lot of interest.

While effective in some respects, spectral and eigenvector statistics fall short of capturing time dependent features that are crucial for understanding quantum dynamics.

More recently, quantities such as the Loschmidt Echo [7, 8, 9], out of time ordered correlators (OTOCs) [10, 11, 12, 13], and Krylov complexity [14, 15, 16] have emerged as alternatives that offer a more dynamical perspective. Despite their potential, the application of these measures to quantum information processing tasks remains largely unexplored.

A major contribution of this work lies in showing that Krylov complexity can be formulated using time evolved states and operators. In the case of states, the constructed Krylov space is fully measurable and does not require prior knowledge of the system's Hamiltonian. The proposed Krylov spaces challenge claims that Krylov bases constructed from powers of the Hamiltonian always minimize spread complexity [16]. Based on numerical results, this thesis demonstrates that alternative constructions can result in slower growth of complexity. Similar behavior is observed for Krylov operator complexity [14], which motivates the definition of time dependent Krylov spaces. From these spaces, two new measures are introduced: Krylov expressivity and Krylov observability. These quantities capture the system's accessible state space as it evolves over time.

Building on this, the expressiveness of quantum reservoirs is examined through Krylov based metrics. Results show that Krylov complexity does not adequately explain task performance. On the other hand, the proposed notions of Krylov expressivity and Krylov observability offer a more accurate reflection of a system's ability to map information. In particular, the nearly identical behavior of Krylov observability and data expressivity in quantum reservoir computing supports the interpretation that QRC maps input data into the Krylov basis. These ideas are then applied to engineer custom reservoirs capable of learning and predicting complex time dependent patterns.

Another practical difficulty in quantum reservoir computing arises from the nature of quantum measurement. Each readout disturbs the system and requires complete reinitialization of past inputs, which leads to quadratic time complexity. This thesis proposes a solution that makes use of the fading memory property of reservoirs. By restricting inputs to a finite temporal window, the computational cost is reduced while task performance is preserved or even improved.

Lastly, the thesis explores several measures aimed at detecting phase transitions in disordered quantum systems. In doing so, it shows that Krylov complexity is not only sensitive to the localized to ergodic crossover but can also differentiate between distinct localized regimes, such as spin glass and many body localization. Interestingly, the best performance of higher order processing capacities in QRC coincides with a critical time scale defined on Krylov complexity. These observations, while preliminary, suggest that

Krylov based diagnostics may serve as practical tools for probing structure in quantum dynamical systems.

By linking the concepts of quantum complexity, machine learning, and dynamical phase transitions, this thesis contributes to a broader understanding of quantum systems not just as physical entities but as platforms for information processing. The methods developed and the results obtained open new avenues for fields such as quantum phase transitions, quantum chaos, and quantum reservoir computing. On a more fundamental level, the defined measures also apply to homogeneous differential equations.

I.2 Quantum Reservoir Computing

The interplay between physics and machine learning has proven to be highly impactful, underscored by the awarding of the Nobel Prize to John Hopfield and Geoffrey Hinton for their foundational contributions. In parallel, quantum computation has emerged as a rapidly advancing field of research [17]. Quantum computing uses quantum gates instead of classical logic operations, and represents information using qubits that evolve in the system's Hilbert space, rather than conventional bits. This can sometimes result in lower time-complexity to solve tasks. One of the most prominent demonstrations of quantum advantage is Shor's algorithm [18].

Given the significance of both machine learning and quantum computing, a natural question arises: to what extent can machine learning algorithms be implemented on quantum systems? What form might these hybrid models take, and what advantages could they offer? These questions have given rise to the field of quantum machine learning (QML), which aims to integrate the computational strengths of both domains.

QML models are designed to exploit the exponential dimensionality of Hilbert space to achieve powerful computational capabilities. Within QML, one particularly active area of research is quantum reservoir computing (QRC). Inspired by classical reservoir computing, where dynamical physical systems are harnessed to perform machine learning tasks [19, 20, 21, 22, 23, 24, 25], QRC adapts this concept to the quantum realm. In this framework, input data is encoded into the quantum state of a system, which then evolves over time. Observables are measured at discrete intervals, and the resulting readouts are multiplied with trained weights for tasks such as time series prediction.

QRC has been actively studied, particularly in the context of the Ising model [26, 27, 28, 29, 30, 31, 32, 33], and has seen early experimental implementations on IBM's quantum processors [34, 35, 36, 37]. Notably, QRC has demonstrated the ability to treat noise and dissipation not as obstacles, but as computational resources [38, 39, 40]. To address the time complexity challenges associated with time-series analysis, several techniques have been proposed, including weak measurement protocols [41, 42], reinitialization strategies [43], feedback mechanisms through output reinjection [44, 45], and continuous measurement schemes [46]. Additionally, [47] employs methods from parameterized quantum circuits to investigate the expressive power limits of QRC architectures. A deeper understanding of quantum machine learning, and quantum reservoir computing in particular, is essential for identifying effective and practical applications. In classical machine learning, insights are often obtained through analysis of expressivity and interpretability. In quantum mechanics, however, quantum information measures provide a more suitable framework for exploring the structure and behavior of these systems. This invites the question of how recent developments in quantum information theory might inform our understanding of quantum machine learning, and to what extent new insights in quantum information can be gained by posing questions aimed at explaining quantum machine learning task performance. Th previous section discussed several such tools—namely, the Loschmidt echo, Krylov expressivity, and Krylov observability—were discussed in this regard.

I.3 Outline of This Thesis

This thesis investigates the time evolution of quantum states and operators through the lens of Krylov complexity. Building on established measures, it introduces novel quantum information quantities aimed at capturing the effective phase space dimension in quantum dynamics. These measures are then applied to quantum reservoir computing, providing insights into the design of more efficient quantum reservoirs. This thesis is organized into five sections, beginning with the Introduction, of which this outline is a part of.

Chapter II presents the necessary theoretical background for the discussions that follow. Section II.1 introduces dynamical systems and classifies them into linear, non-linear, and chaotic categories, alongside methods for their quantification. Next, Section II.2 discusses the postulates and time evolution in quantum mechanics. Section II.3 provides a brief overview of machine learning and reservoir computing, followed by Section II.4, which covers the quantum reservoir computing framework and the potential benefits of quantum machine learning.

The focus then shifts to Krylov complexity in Chapter III, a recently emerging field for studying time evolution in quantum systems. This chapter includes the main theoretical contributions of the thesis and aims to offer a visual and intuitive understanding of how quantum states and operators evolve in their respective Krylov space. Section III.1 discusses how pure quantum states evolve in the Krylov basis and introduces Krylov spread complexity as a measure of this evolution. It is demonstrated that the Krylov space can be constructed from time-evolved states, and that the common assumption that the Krylov basis minimizes state spread is incorrect. Instead, Krylov spaces built from appropriately chosen time-evolved states can exhibit lower spread. Building on this, a new measure called Krylov expressivity is introduced. Rather than tracking a state's evolution within a fixed Krylov space, this measure defines a time-dependent Krylov space and computes its effective phase space dimension at each point in time.

Following the discussion of states, the thesis turns to how operators evolve in the corresponding Krylov space of the Liouvillian in Section III.2. It is shown that time-evolved operators can be used to define Krylov operator complexity, which captures how operators spread in Krylov space. Like the case with states, it is demonstrated that Krylov operator complexity can be reduced using time-evolved Krylov spaces. A further measure, Krylov observability, is introduced to characterize the phase space dimension based on multiple operators and the frequency with which they are measured.

In Chapter IV, Krylov spread complexity, Krylov expressivity, and Krylov observability are computed and applied as quantifiers in the context of quantum reservoir computing (QRC). In Section IV.1, a short summary of quantum reservoir computing and Krylov complexity is given, which is necessary for understanding the this chapter.

Section IV.2 analyzes how these measures correlate with QRC task performance. It is shown that Krylov observability effectively captures data expressivity, suggesting that QRC maps input data onto Krylov space. Subsequently, Section IV.3 explores how these measures can guide the optimization and interpretation of quantum reservoir architectures.

The thesis then switches to discussing the time-complexity problem in QRC in Section IV.4. In QRC, measurement collapses the quantum wavefunction, requiring full reinitialization of the time series. This challenge is addressed through memory-restricted quantum reservoir computing, which enhances system performance by avoiding costly reinitializations.

The last chapter, Chapter V, discusses how quantum phase transitions can be probed using selected quantum information measures. It emphasizes that quantum evolution is unitary and therefore not chaotic in the classical sense. Since the eigenvalues remain constant over time, defining a Lyapunov exponent is not practical. Nevertheless, quantum systems often exhibit chaos-like behavior either in the classical limit or in the evolution of observables. For this reason, the chapter includes a discussion of these aspects.

Originally, this chapter was intended to be part of the theoretical background. However, during the course of the work, it became clear that the study of quantum chaos is an active area of research with no consensus on its precise definition or how best to probe it. As a result, the chapter evolved into a research-focused section, presenting new insights and connections between phase transitions, quantum information measures, and task performance in quantum reservoir computing. In simulations, it can be shown that Krylov complexity is a powerful tool for probing phase transitions, especially when time is rescaled to the Heisenberg time. This approach makes it possible to distinguish between different localized regions, but an in-depth study is required for a final conclusion. Therefore, these results should be viewed as an outlook for future work.

List of Publications

- S. Čindrak, B. Donvil, K. Lüdge and L. C. Jaurigue, “Enhancing the performance of quantum reservoir computing and solving the time-complexity problem by artificial memory restriction,” *Phys. Rev. Res.* **6**, 013051 (2024).
- S. Čindrak, A. Paschke, L. C. Jaurigue and K. Lüdge, “Measurable Krylov Spaces and Eigenenergy Count in Quantum State Dynamics,” *J. High Energ. Phys.* **2024**, 83 (2024).
- S. Čindrak, K. Lüdge and L. C. Jaurigue, “Quantum reservoir computing maps data onto the Krylov space,” arXiv:2502.12157 [submitted].
- S. Čindrak, L. C. Jaurigue and K. Lüdge, “Engineering Quantum Reservoirs through Krylov Complexity, Expressivity and Observability,” arXiv:2409.12079 [submitted].

II Chapter: Theory

Describing and predicting how physical systems change over time has always been, if not the most, one of the most fundamental questions in physics. Whether it concerns understanding how the planets orbit the sun, how a ball falls from the sky, how the intensity of a laser changes, how a quantum mechanical state evolves, or how biological systems develop, the underlying interest is the same. Whenever a physical effect is observed, it is the result of some quantity changing over time. Interestingly, some systems appear deterministic, and the value of interest approaches a final state. An example of such a system is a capacitor that becomes charged. Others seem to remain in continuous motion, such as the moon orbiting the Earth. Some systems, such as weather patterns, cannot be modeled with complete precision, and the best available predictions are probabilistic, indicating the likelihood of rain, sun, or snow. In chaotic systems, a small change in initial conditions can lead to dramatically different outcomes. This phenomenon is often referred to as the butterfly effect or chaos.

Of particular importance in this work are quantum systems. These systems are best described as a superposition of deterministic and probabilistic behavior.Quantum systems can be interpreted in the context of dynamical systems as being a superposition of both deterministic and probabilistic dynamics. When no measurement is performed, quantum systems may be viewed as dynamical systems evolving under the Schrödinger equation. However, the deterministic behavior of the quantum state disappears once a measurement is made, and the system is then best interpreted by a probability distribution. This is by no means a physical interpretation of quantum mechanics, but serves more to give a feeling of how quantum systems can be thought of in regards to dynamical systems.

This chapter begins with an overview of the basics of dynamical systems, describing when they are linear, non-linear, dynamical, or chaotic. While this may seem very basic to the reader, it becomes important when quantum mechanics are discussed. The unitary time evolution is linear with respect to the quantum state, yet behavior can be observed that mimics non-linear or chaotic dynamics. A quantum-mechanical description of quantum chaos under unitary time evolution remains largely unresolved and has led to the introduction of quantum-based information measures that probe such systems. Here, unresolved implies that there is no definite consensus like it is in classical chaos theory and new ways to probe quantum chaos pop up every now and then.

Following the discussion of dynamical systems and quantum mechanics, a brief overview of reservoir computing will be provided, followed by a preliminary discussion of quantum reservoir computing.

II.1 Dynamical Systems

This section provides a recap of the important aspects of dynamical systems that are necessary for understanding this thesis. After discussing dynamical, linear, and non-linear systems, chaotic systems will be introduced. A detailed description of most concepts in dynamical systems can be found in standard textbooks such as [48].

To this end, the Lorenz system is examined as an example of chaos [49], while the Lotka-Volterra system is discussed as an example of a non-linear and non-chaotic system [50]. The time evolution shows how both of these systems change with a small initial perturbation and should give the reader an intuitive sense of chaos.

II.1.1 Mathematical Definition of a Dynamical System

A **dynamical system** is one in which the current state completely determines both its future and past evolution. The **phase space** X represents the set of all possible states of the system, where each state is denoted by a point $x \in X$. This space can be a Euclidean space, such as $\mathbb{R}^N$ or $\mathbb{C}^N$, or a more general manifold. The evolution can be either continuous ($t \in \mathbb{R}$) or discrete ($t \in \mathbb{Z}$). I¡n physics, the time evolution is typically governed by equations of motion, such as the Euler-Lagrange equations in classical mechanics or the Schrödinger equation in quantum mechanics.

The evolution is described by a function $x(t, x_0)$, where x_0 is the initial state at $t = 0$, and the system evolves according to:

$$x(0) = x_0 \in X, \tag{II.1}$$

$$x(t, x_0) \in X. \tag{II.2}$$

This can be expressed via a family of maps ϕ, where

$$\phi(t_0, t_1, x_0) = x(t_1, x_0). \tag{II.3}$$

By setting $t_0 = 0$ and $t_1 = t$, this simplifies to:

$$\phi(x_0, t) = \phi_t(x_0) = x(t, x_0). \tag{II.4}$$

To qualify as a **dynamical system**, the evolution must satisfy two key properties:

$$x(0, x_0) = x_0, \tag{II.5}$$

$$x(t + s, x_0) = x(t, x(s, x_0)) = x(s, x(t, x_0)). \tag{II.6}$$

This means the system's evolution is associative over time. Setting $s = -t$ leads to the reversibility condition:

$$\phi(-t) = (\phi(t))^{-1}, \tag{II.7}$$

implying that the system is reversible, where applicable. The following example helps to verify whether a given system qualifies as a dynamical system:

$$x(t, x_0) = \frac{x_0}{1 - tx_0}. \tag{II.8}$$

Checking the dynamical system properties:

$$\begin{aligned} x(0, x_0) &= \frac{x_0}{1 - 0} = x_0, \\ x(t, x(s, x_0)) &= \frac{\frac{x_0}{1 - sx_0}}{1 - t\frac{x_0}{1 - sx_0}} = \frac{x_0}{1 - (t + s)x_0} = x(t + s, x_0). \end{aligned} \tag{II.9}$$

This confirms the dynamical properties. For $t = \frac{1}{x_0}$ the function becomes undefined. Hence, the system is a valid dynamical system only for $t < \frac{1}{x_0}$. Such systems are classified as **nonlinear dynamical systems**. Other classes include **chaotic dynamical systems**, **stochastic systems**, and **time-delay systems**, which are often collectively referred to as **dynamical systems**.

II.1.2 Physical Dynamical Systems and Fixed-Point Analysis

Physical systems are frequently modeled by differential equations, which are often solvable only numerically. These are typically of the form:

$$\frac{d}{dt} x(t) = F(x(t), t), \tag{II.10}$$

where $x(t) \in \mathbb{R}^n$ or $\mathbb{C}^n$ is the **dynamical variable**, and $F : \mathbb{R}^n \times \mathbb{R}_t \to \mathbb{R}^n$ is the **vector field**. Note that higher order differential equations can be brought in this form via variable substitution.

A commonly used method for analyzing such systems is **fixed-point analysis**, which involves finding points where the time derivative vanishes, i.e., solving for x^* such that:

$$F(x^*, t) = 0. \tag{II.11}$$

These are known as **fixed points**, and they may be either **stable** or **unstable**. A stable fixed point is one where small perturbations decay over time, returning the system to equilibrium—similar to a ball resting in a valley. An unstable fixed point behaves like a ball balanced on a hilltop, where any small disturbance drives it away from equilibrium.

To determine the stability of a fixed point, the system is linearized around that point by considering a small perturbation δx:

$$\delta x = x - x^*. \tag{II.12}$$

Applying a Taylor expansion gives:

$$\delta \dot{x}_i = \sum_{k=1}^{n} \left(\frac{\partial F}{\partial x_k} \right)_{x^*} \delta x_k. \tag{II.13}$$

The **Jacobian matrix** is defined as:

$$J = \begin{pmatrix} \partial_1 F_1 & \partial_2 F_1 & \cdots & \partial_n F_1 \\ \partial_1 F_2 & \partial_2 F_2 & \cdots & \partial_n F_2 \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 F_n & \partial_2 F_n & \cdots & \partial_n F_n \end{pmatrix}. \tag{II.14}$$

At the fixed point x^*, equation (II.13) becomes:

$$\delta \dot{x} = J(x^*) \delta x. \tag{II.15}$$

This reveals that the system's response to perturbations is governed by a linear differential equation involving the Jacobian.

To evaluate the stability, one computes the **eigenvalues** of the Jacobian by solving:

$$J(x^*, t) v = \lambda v, \tag{II.16}$$

where v is an eigenvector and λ is the corresponding eigenvalue.

The nature of the fixed point is then classified as:

- **Stable** if all eigenvalues have negative real parts.
- **Unstable** if any eigenvalue has a positive real part.
- **Neutrally stable** if the eigenvalues are purely imaginary, leading to oscillatory behavior.

Previous sections have introduced key concepts but not yet classified dynamical systems based on behavior. Such classifications assist in understanding their evolution over time. The systems discussed in this work primarily fall into three categories: **linear systems**, **nonlinear systems**, and **chaotic systems**[48].

II.1.3 Linear Systems

Definition 1. *A linear dynamical system, in the context of this work, is described by*

$$\begin{aligned} x_0 &= x(0), \\ \dot{x}(t) &= Ax(t), \end{aligned} \tag{II.17}$$

where $x_0 \in \mathbb{R}^N$ or $x_0 \in \mathbb{C}^N$ is the initial condition, and $A \in \mathbb{R}^{N\times N}$ or $A \in \mathbb{C}^{N\times N}$ is a matrix with ***constant*** *coefficients. If the system is described by a higher-order differential equation, it can be rewritten into the form of Eq. (II.17).*

A system is called linear if for any two solutions $x_1(t)$ and $x_2(t)$, the linear combination $x(t) = \alpha x_1(t) + \beta x_2(t)$ is also a solution, where $\alpha, \beta \in \mathbb{R}$ or $\mathbb{C}$ are scalar coefficients. The solution to the differential equation is given by the matrix exponential:

$$x(t) = e^{At} x_0. \tag{II.18}$$

The Jacobian at any time is simply:

$$J(t) = A. \tag{II.19}$$

Since A is constant, the Jacobian J is also constant across time and space, implying that the eigenvalues remain constant for all points $x(t)$ in the phase space. Solving the eigenvalue equation yields:

$$Av_k = \lambda_k v_k. \tag{II.20}$$

The system is stable if all $\mathrm{Re}(\lambda_k) < 0$ and unstable if at least one $\mathrm{Re}(\lambda_k) > 0$.

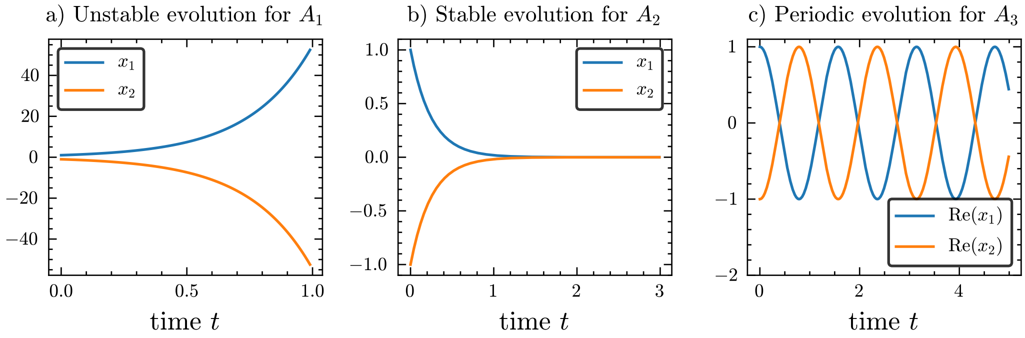


Figure II.1: Simulation of three linear systems. The panels a), b), and c) show the behavior for A_1, A_2, and A_3, respectively.

To develop an intuition for linear systems and their properties, three example matrices A_1, A_2, and A_3 are simulated and analyzed:

$$A_1 = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}, \quad A_2 = -A_1, \quad A_3 = iA_1, \quad x_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{II.21}$$

The fixed point for each system is $x^* = (0, 0)^T$, and the stability is determined as follows:

1. For A_1: $\lambda_1 = 4,\ \lambda_2 = 2 \Rightarrow$ The fixed point is unstable.

2. For A_2: $\lambda_1 = -4,\ \lambda_2 = -2 \Rightarrow$ The fixed point is stable.

3. For A_3: $\lambda_1 = 4i,\ \lambda_2 = 2i \Rightarrow$ The fixed point exhibits oscillatory behavior and is a focus.

Figure II.1 shows the time evolution of $x(t) = (x_1(t), x_2(t))$ for matrices A_1, A_2, and A_3 in panels a, b, and c, respectively. The dynamical variable becomes complex in the case of A_3, and only the real part is visualized. The matrix A_1 induces instability around the fixed point $x^* = 0$, as evidenced by the positive eigenvalues. For large times, $x_1(t) \to \infty$, while $x_2(t) \to -\infty$. For A_2, with negative eigenvalues, the system converges to the fixed point, as seen in panel b. In the case of A_3, the evolution is periodic due to the purely imaginary eigenvalues, shown in panel c.

The time evolution for A_3 is expressed as:

$$x(t) = U(t)x_0, \quad \text{where} \quad U(t) = e^{iA_1 t}. \tag{II.22}$$

Since A_1 is Hermitian ($A_1 = A_1^\dagger$), it possesses a complete set of orthonormal eigenvectors. Using Eq. (II.20), the evolution can be written as:

$$x(t) = \sum_{k=1}^{2} e^{i\lambda_k t} v_k v_k^\dagger x_0. \tag{II.23}$$

This evolution satisfies the property $U^\dagger(t) = U^{-1}(t)$, as shown:

$$\begin{aligned} U^\dagger(t)U(t) &= \sum_{k=1}^{2} e^{-i\lambda_k t} v_k v_k^\dagger \sum_{j=1}^{2} e^{i\lambda_j t} v_j v_j^\dagger \\ &= \sum_{k=1}^{2}\sum_{j=1}^{2} e^{-i\lambda_k t} e^{i\lambda_j t} v_k (v_k^\dagger v_j) v_j^\dagger. \end{aligned} \tag{II.24}$$

Because A is Hermitian, the eigenvectors v_k form an orthonormal basis, i.e., $v_k^\dagger v_j = \delta_{kj}$. Thus,

$$\begin{aligned} U^\dagger(t)U(t) &= \sum_{k=1}^{2} e^{-i\lambda_k t} e^{i\lambda_k t} v_k v_k^\dagger = \sum_{k=1}^{2} v_k v_k^\dagger = I, \\ \Rightarrow \quad U^\dagger(t) &= U^{-1}(t). \end{aligned} \tag{II.25}$$

This in-depth discussion of unitary time evolution is motivated by its fundamental role in quantum mechanics. Time evolution governed by the Schrödinger or von Neumann equation is always unitary.

II.1.4 Non-Linear and Chaotic Systems

While many physical systems can be described by linear dynamical systems, others require more complexity. Some systems exhibit non-linear behavior, and among these, some also display chaotic dynamics, where small differences in initial conditions lead to vastly different evolutions over time.

This subsection introduces non-linear and chaotic systems. Two examples are discussed: the Lotka-Volterra system as a non-linear but non-chaotic system, and the Lorenz system as a non-linear and chaotic system. The Lorenz attractor is a widely used benchmark in time-series prediction due to its complexity and is applied in this thesis for chaotic time-series prediction tasks.

Definition 2 (Non-Linear System)**.** *A non-linear dynamical system, in the context of this thesis, is described by*

$$\begin{aligned} x_0 &= x(0), \\ \dot{x}(t) &= A(x(t), t)x(t), \end{aligned} \tag{II.26}$$

where $x_0 \in \mathbb{R}^N$ *or* $x_0 \in \mathbb{C}^N$ *is the initial condition and* $A(x,t) \in \mathbb{R}^{N\times N}$ *or* $\mathbb{C}^{N\times N}$ *is a matrix that typically depends on the state* x*. The system is non-linear if it does* ***not*** *satisfy the superposition principle, i.e., if* $x_1(t)$ *and* $x_2(t)$ *are solutions, then*

$$x(t) = \alpha x_1(t) + \beta x_2(t) \tag{II.27}$$

is not necessarily a solution for arbitrary $\alpha, \beta \in \mathbb{R}$ *or* $\mathbb{C}$*.*

A simple example is:

$$\dot{x}(t) = x(t) + x(t)^2 = f(x(t)). \tag{II.28}$$

Assuming $x_1(t)$ and $x_2(t)$ are solutions, a superposition yields:

$$\begin{aligned} \frac{d}{dt}(x_1 + x_2) &= x_1 + x_2 + (x_1 + x_2)^2 \\ &= x_1 + x_2 + x_1^2 + x_2^2 + 2x_1x_2 \\ &= f(x_1) + f(x_2) + 2x_1x_2 \\ \Rightarrow \frac{d}{dt}(x_1 + x_2) &\neq f(x_1) + f(x_2). \end{aligned} \tag{II.29}$$

This demonstrates a key distinction between linear and non-linear systems. Among non-linear systems, a particularly important subclass is chaotic systems, where small perturbations can lead to entirely different long-term behavior. The Lyapunov exponent is a standard tool to assess whether a system is chaotic.

Definition 3 (Lyapunov Exponent and Chaos)**.** *Let* $x(t, x_0)$ *be the trajectory of an* n*-dimensional system with initial condition* x_0*, and let* $x(t, x_0 + \delta_k)$ *be the trajectory with a small perturbation in the* k*-th direction. The* k*-th Lyapunov exponent is defined as:*

$$\lambda_k = \lim_{t\to\infty} \frac{1}{t} \ln \left| x_k(t, x_0) - x_k(t, x_0 + \delta_k) \right|. \tag{II.30}$$

The largest Lyapunov exponent is given by:

$$\lambda = \sup\{\lambda_1, \lambda_2, \ldots, \lambda_n\}. \tag{II.31}$$

A system is classified as follows:

1. *If $\lambda > 0$, the system is **chaotic**.*

2. *If $\lambda < 0$, the system is **non-chaotic**.*

The 3D Lotka-Volterra model describes predator-prey interactions[50]. Here, x is the prey population, y is a primary predator feeding on x, and z is a secondary predator feeding on y. In contrast, the Lorenz model represents atmospheric convection, where x denotes convection and y, z represent horizontal and vertical temperature variations. Both systems are non-linear. However, the parameters of the Lorenz system are chosen so that it exhibits a chaotic attractor—meaning the system's state evolves within a bounded region of phase space in a non-repeating, sensitive manner. In contrast, the Lotka-Volterra model exhibits a periodic orbit and remains non-chaotic. The Lorenz system is a classic benchmark in chaos theory and reservoir computing[49].

Lotka-Volterra (non-chaotic)

$$\begin{aligned} \alpha =&0.1, \quad \beta = 0.02, \quad \delta = 0.01, \quad \gamma = 0.1 \\ &\epsilon = 0.015, \quad \zeta = 0.008, \quad \eta = 0.1 \\ &\frac{dx}{dt} = \alpha x - \beta xy - \epsilon x^2 \\ &\frac{dy}{dt} = \delta xy - \gamma y \\ &\frac{dz}{dt} = \zeta yz - \eta z \end{aligned}$$

Lorenz (chaotic)

$$\begin{aligned} &\sigma = 10.0, \quad \rho = 28.0, \quad \beta = \frac{8}{3} \\ &\frac{dx}{dt} = \sigma(y - x) \\ &\frac{dy}{dt} = x(\rho - z) - y \\ &\frac{dz}{dt} = xy - \beta z \end{aligned}$$

Figure II.2 shows the evolution of the Lotka-Volterra system under two initial conditions: the original (blue) and a 1% perturbed version (orange). The three variables exhibit nearly identical behavior, and their trajectories remain close in phase space (a–c). This demonstrates that the system is non-linear but not chaotic, which is also evident in the trajectory evolutions shown in Figure II.2(d).

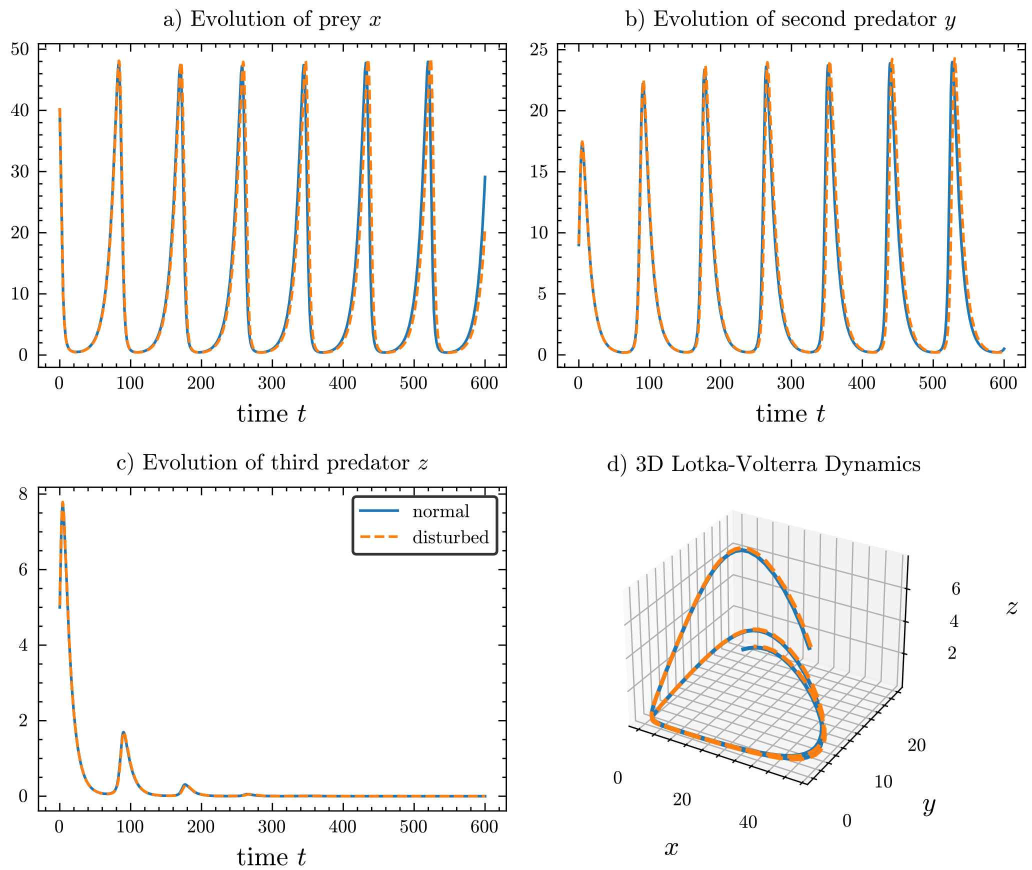


Figure II.2: Time evolution of the Lotka-Volterra system with original initial condition (blue) and a 1% perturbed initial condition (orange). The trajectories are nearly identical, indicating non-chaotic behavior.

Figure II.3 presents the Lorenz system under similar conditions. Initially, the trajectories are nearly indistinguishable, but around $t \approx 20$, the x and y variables begin to diverge (Figure II.3.a-b). The z variable (Figure II.3.c) remains close initially but diverges at later times. The phase space visualization confirms that the two trajectories deviate significantly, demonstrating chaotic behavior. The Lyapunov exponents for the Lorenz system are:

$$\lambda_1 = 0.905, \quad \lambda_2 = 0.0, \quad \lambda_3 = -14.6. \tag{II.32}$$

Here, λ_3 corresponds to contraction toward the attractor, λ_2 represents evolution along the attractor, and λ_1 quantifies sensitivity to perturbations orthogonal to the attractor.

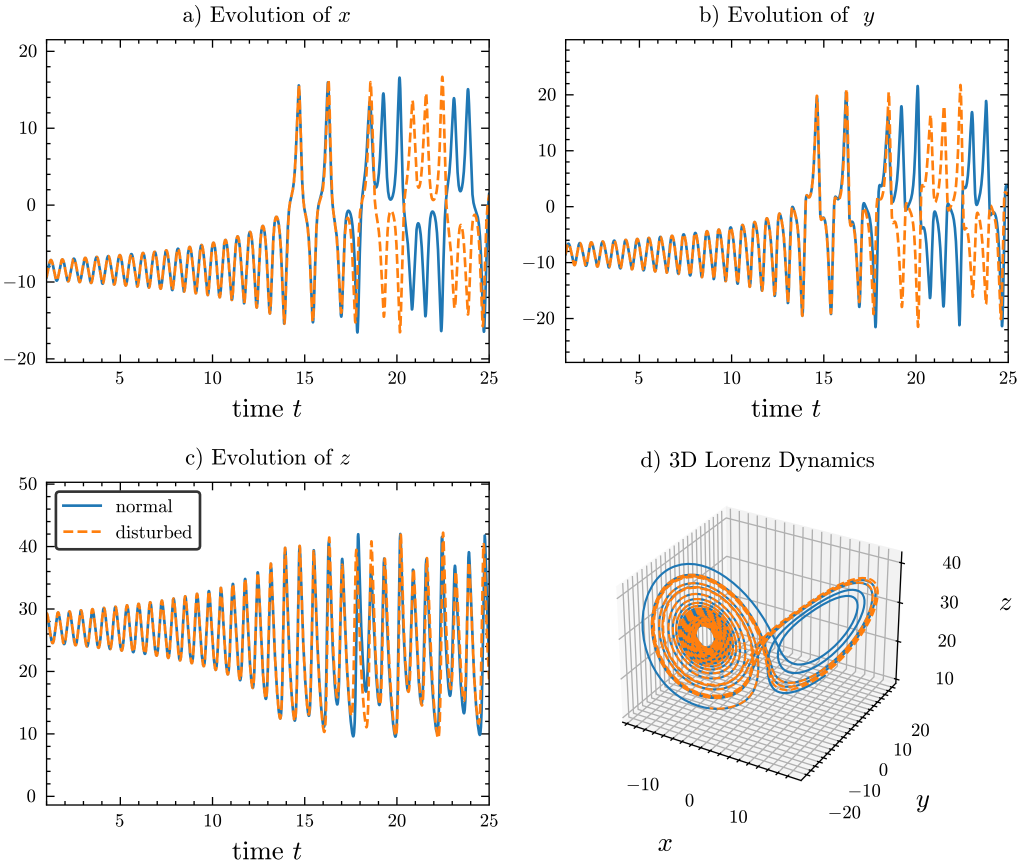


Figure II.3: Time evolution of the Lorenz system with original (blue) and 1% perturbed (orange) initial conditions. Divergence in phase space and positive Lyapunov exponent indicate chaotic behavior.

II.2 Quantum Mechanics

After discussing classical dynamics in detail, the focus now shifts to quantum mechanics. In the Copenhagen interpretation, quantum mechanics is built on five fundamental postulates that govern the time evolution of a system. These postulates are as follows:

II.2.1 Postulates of Quantum Mechanics and the Schrödinger Equation

1. **State of a Quantum System:** The state of a quantum system is described by a state vector $|\psi\rangle$ in a complex Hilbert space $\mathcal{H}$, i.e., $|\psi\rangle \in \mathcal{H}$. Additionally, the state is normalized such that $\| \, |\psi\rangle \, \| = 1$.

2. **Observables and Operators:** Every observable quantity in quantum mechanics corresponds to a Hermitian operator A acting on the state $|\psi\rangle$.

3. **Measurement and Eigenvalues:** Upon measuring an observable A, the only possible outcomes are the eigenvalues a_n of the corresponding operator, satisfying

   $$A \, |a_n\rangle = a_n \, |a_n\rangle \, .$$

   If the eigenvalue a_n is measured, the quantum state collapses to the corresponding eigenstate $|a_n\rangle$, i.e.,

   $$\text{measuring } a_n \text{ results in } \; |\psi\rangle \xrightarrow{\text{collapse}} |a_n\rangle \, . \tag{II.33}$$

4. **Probability Interpretation (Born Rule):** The probability of measuring the eigenvalue a_n is given by

   $$P(a_n) = |\langle a_n|\psi\rangle|^2.$$

5. **Time Evolution (Schrödinger Equation):** The time evolution of a quantum system is governed by the Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle ,$$

where H is the Hamiltonian operator representing the total energy of the system.

Consider now only the Schrödinger equation. If the Hamiltonian H is time-independent, the system evolves linearly in time. The solution to the Schrödinger equation in this case is given by the matrix exponential, as discussed in Section II.1.3:

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}Ht} |\psi_0\rangle , \tag{II.34}$$

where $|\psi_0\rangle = |\psi(0)\rangle$ is the initial state. The Hamiltonian H is always Hermitian, i.e., $H = H^\dagger$, resulting in unitary time evolution. This ensures that the normalization condition $\| \, |\psi(t)\rangle \, \| = 1$ holds for all times t. Unitary time evolution, discussed in Section II.1.3, also implies that such systems are not chaotic in the classical sense.

Nonetheless, quantum chaos remains an active field of research. Although quantum systems evolve unitarily, they can exhibit features reminiscent of classical chaos. One broad interpretation of quantum chaos lies in the exponential scaling of the Hilbert space dimension $\mathcal{H}$. Even a small number of quantum particles can result in a Hilbert space of such high dimensionality that it exceeds the capabilities of modern supercomputers. This has led to the development of approximation methods such as Hartree-Fock and density functional theory.

In this context, quantum chaos is understood as the emergence of chaotic behavior in the expectation values of certain observables. Although this thesis does not directly study quantum chaos, many of the proposed methods contribute novel ways of characterizing it.

As introduced earlier in this chapter, quantum systems can be interpreted as exhibiting both deterministic and probabilistic behavior. The probalistic nature arises from the third and fourth postulates, where a quantum measurement causes the system to collapse randomly to one of the eigenstates $|a_n\rangle$, with probability $P(a_n) = |\langle a_n|\psi\rangle|^2$.

II.2.2 Evolution of Mixed States and Operators

The previous chapter discussed how the quantum state collapses after a measurement and how the measurement outcome is random, governed by the probability distribution

$$P(\alpha_n) = |\langle a_n|\psi\rangle|^2. \tag{II.35}$$

The expectation value of this distribution is given by

$$\begin{aligned}\langle A\rangle &= \sum_n a_n P(\alpha_n) = \sum_n a_n |\langle a_n|\psi\rangle|^2 \\ &= \sum_n \langle\psi|\, a_n\, |a_n\rangle \langle a_n|\psi\rangle = \langle\psi|\, A\, |\psi\rangle\,. \end{aligned} \tag{II.36}$$

Another way to compute the expectation value is by defining the density matrix $\rho = |\psi\rangle\langle\psi|$, which leads to the expression

$$\langle A\rangle = \mathrm{Tr}(\rho A) = \mathrm{Tr}(A\rho). \tag{II.37}$$

Now, one important aspect of quantum mechanics remains to be addressed: mixed states. Suppose an intern is blamed for a mistake during a measurement in which the outcome was not recorded. According to the postulates, the state collapses $|\psi\rangle \to |a_n\rangle$ with probability $P(a_n) = P_n$. A seemingly clever (but incorrect) approach would be to represent the system as a superposition of all possible measurement outcomes:

$$|\psi_a\rangle = \sum_n \sqrt{P_n} a_n\, |a_n\rangle\,. \tag{II.38}$$

While this may seem reasonable, it is **incorrect** because it implies a non-linear relationship between the state and the measurement probabilities. This issue becomes evident when computing the expectation value:

$$\langle A\rangle = \mathrm{Tr}(A\rho) = \mathrm{Tr}\left(A \sum_n \sqrt{P_n}\, |a_n\rangle \sum_m \sqrt{P_m}^{*}\, \langle a_m|\right), \tag{II.39}$$

which leads to off-diagonal terms of the form $|a_n\rangle\langle a_m|$, introducing unwanted correlations between different outcomes.

The correct way to model this *uncertainty* is through a mixed state represented by the density matrix:

$$\rho = \sum_n P_n\, |a_n\rangle\langle a_n|\,. \tag{II.40}$$

This construction preserves linearity in the measurement process and the computation of expectation values.

One might argue that the mixed-state formalism is more general than the pure-state formalism. While that is true in a mathematical sense, most interpretations of quantum mechanics consider mixed states to arise from incomplete knowledge about a fundamentally pure state, i.e., one described by $\rho = |\psi\rangle\langle\psi|$.

Nevertheless, the **von Neumann equation**

$$\dot{\rho}(t) = -\frac{i}{\hbar}[H, \rho(t)] \tag{II.41}$$

with initial condition $\rho(0) = \rho_0$, governs the time evolution of all (pure or mixed) quantum states. Its solution is given by

$$\rho(t) = e^{-\frac{i}{\hbar}Ht}\rho_0 e^{\frac{i}{\hbar}Ht}. \tag{II.42}$$

Another useful perspective in quantum mechanics is to consider the time evolution of expectation values:

$$\langle A(t) \rangle = \langle \psi(t) | A | \psi(t) \rangle . \tag{II.43}$$

Substituting the unitary evolution $|\psi(t)\rangle = e^{-\frac{i}{\hbar}Ht}|\psi_0\rangle$, we get

$$\langle A(t) \rangle = \langle \psi_0 | e^{\frac{i}{\hbar}Ht} A e^{-\frac{i}{\hbar}Ht} | \psi_0 \rangle . \tag{II.44}$$

Instead of interpreting the state $|\psi(t)\rangle$ as evolving, one can equivalently consider the observable as evolving in time:

$$A_H(t) = e^{\frac{i}{\hbar}Ht} A e^{-\frac{i}{\hbar}Ht}, \tag{II.45}$$

with the state remaining fixed. The equation of motion corresponding to this picture is the **Heisenberg equation**:

$$\frac{d}{dt}A_H(t) = \frac{i}{\hbar}[H, A_H(t)] + \left(\frac{\partial A}{\partial t}\right). \tag{II.46}$$

For time-independent observables, as considered in this work, the partial derivative vanishes, and the equation of motion in the Heisenberg picture simplifies to:

$$\frac{d}{dt}A(t) = \frac{i}{\hbar}[H, A]. \tag{II.47}$$

II.3 Reservoir Computing

The last few decades have brought about a revolution in computer science through the introduction of machine learning. The demonstration of machine learning algorithms that outperform humans in image recognition for the first time in 2012 has led to the rapid development seen today. This rapid progress spans time series forecasting, image recognition, classification, and now large language models, which enable human-like conversations and can implement algorithms and code. Models are currently trained to create images and videos from text. The development of modern machine learning methods is based on the backpropagation algorithm and optimization with respect to a loss function. The significance of machine learning in physics, and vice versa, is evident from the awarding of the Nobel Prize in Physics to John J. Hopfield and Geoffrey Hinton for their contributions to machine learning.

II.3.1 Basics of Neural Networks

Before introducing reservoir computing, it is necessary to establish the foundation and gain at least a basic understanding of neural networks. To achieve this, the concept of a neuron must first be introduced.

Neuron

A neuron takes numbers $(x_1, x_2, \dots, x_n) = \mathbf{x} \in \mathbb{R}^n$ as inputs. It is a mapping $g : \mathbb{R}^n \to \mathbb{R}$ from the input space, such that $y = g(\mathbf{X}) \in \mathbb{R}$ represents the output of the neuron. One widely used mapping is the sigmoid neuron σ, which is defined by Eq. (II.48):

$$y = g(\mathbf{x}) = \sigma(\mathbf{x}) = \frac{1}{1 + \exp\Big(-\sum_j w_j x_j - b\Big)} \tag{II.48}$$

Here, $(w_1, w_2, \dots, w_n) = \mathbf{W} \in \mathbb{R}^n$ are called weights, and $b \in \mathbb{R}$ is called the bias.

A neuron can be understood as a function with multiple input variables x_i, defined by the analytical function $g(\cdot) = 1/(1+\exp(\cdot))$, and parameterized by the constants $\mathbf{W}$ and b.

Having a multiple number of neurons results in a neuron layer and the definition is given below.

Neuron Layer Consider a set of k neurons $g_j^{(1)}$, indexed by $j \in \{1, 2, \ldots, k\}$. The superscript (1) indicates the output of the first neuron layer, which will become important in the following definition. Each neuron $g_j^{(1)}$ receives the same input vector $\mathbf{X}$. The output of these k neurons forms a new vector:

$$\mathbf{X}^{(1)} = (g_1^{(1)}(\mathbf{x}), g_2^{(1)}(\mathbf{x}), \ldots, g_k^{(1)}(\mathbf{x})) = \mathbf{G}^{(1)}(\mathbf{x}). \tag{II.49}$$

A collection of neurons that share the same input vector and produce an output vector is referred to as a **neuron layer**.

Feed-Forward Neural Network

A Feed-Forward Neural Network (FFNN) consists of multiple layers of neurons, each applied sequentially to the output of the previous layer. The second neuron layer takes the output of the first neuron layer, $\mathbf{X}^{(1)}$, as input and returns:

$$\mathbf{X}^{(2)} = (g_1^{(2)}(\mathbf{X}^{(1)}), g_2^{(2)}(\mathbf{X}^{(1)}), \ldots, g_l^{(2)}(\mathbf{X}^{(1)})) = \mathbf{G}^{(2)}(\mathbf{X}^{(1)}) = \mathbf{G}^{(2)}(\mathbf{G}^{(1)}(\mathbf{x})). \tag{II.50}$$

The update equation for an arbitrary layer n can be written as:

$$\mathbf{X}^{(n)} = \mathbf{G}^{(n)}(\mathbf{W}^{(n)}\mathbf{X}^{(n-1)} + \mathbf{b}^{(n)}), \tag{II.51}$$

where $\mathbf{W}^{(n)}$ and $\mathbf{b}^{(n)}$ represent the weights and biases of the n-th layer, respectively, and $\mathbf{G}^{(n)}$ is the non-linear activation function. For $1 \leq m \leq L$, the output of the m-th neuron layer is given by the chain rule:

$$\mathbf{X}^{(m)} = \mathbf{G}^{(m)}(\mathbf{G}^{(m-1)}(\ldots \mathbf{G}^{(1)}(\mathbf{x}))). \tag{II.52}$$

This recursive structure enables deeper networks, where each successive layer extracts higher-level features from the input. In this notation, the input of the first layer is set to the input data as $\mathbf{X}^{(0)} = \mathbf{x}$. The output of the final layer is typically denoted as

$\mathbf{y} = \mathbf{X}^{(L)}$, and the weights $\mathbf{W}^{(m)}$ and biases $\mathbf{b}^{(m)}$ are optimized such that $\mathbf{y}$ corresponds to a desired target output.

At this point, it has not been discussed how the weights and biases should be trained so that, for a given input $\mathbf{x}$, the network returns the desired target output $\hat{\mathbf{y}}$. The algorithm that enables the training of neural networks—and that revolutionized their development—is the backpropagation algorithm.

Backpropagation Algorithm

The update equation for a neural network layer is given by:

$$\mathbf{X}^{(n)} = \mathbf{G}^{(n)}(\mathbf{Z}^{(n)}), \tag{II.53}$$

where

$$\mathbf{Z}^{(n)} = \mathbf{W}^{(n)}\mathbf{X}^{(n-1)} + \mathbf{b}^{(n)}. \tag{II.54}$$

To optimize the network parameters, a **loss function** is computed, which quantifies the difference between the actual output $\mathbf{y}_i$ and the target output $\hat{\mathbf{y}}_i$ for a given input $\mathbf{x}_i$. A commonly used loss function is the **mean-squared loss**, where for N input samples $\mathbf{x}_1, \ldots, \mathbf{x}_N$, the loss is given by:

$$L = \frac{1}{2}\sum_i (\mathbf{y}_i - \hat{\mathbf{y}}_i)^2. \tag{II.55}$$

Again, the notation where the input data is referred to as the input of the first layer, $\mathbf{X}^{(0)} = \mathbf{x}_j$, is used. The **backpropagation algorithm** uses the **chain rule** and **gradient descent** to adjust the weights and biases in order to minimize the loss L. The following steps are performed:

1. **Forward Pass:** Compute the network output according to Eq. (II.53).
2. **Compute the Loss:** Evaluate the loss function L, such as the one in Eq. (II.55).
3. **Backward Pass:** Compute the output layer error as

   $$\delta^{(L)} = \frac{\partial L}{\partial \mathbf{X}^{(L)}} \odot G'^{(L)}(\mathbf{Z}^{(L)}), \tag{II.56}$$

   where $\odot$ represents element-wise multiplication, and $G'^{(L)}$ is the derivative of the activation function. The error at each previous layer is computed recursively:

   $$\delta^{(n)} = (\mathbf{W}^{(n+1)})^T \delta^{(n+1)} \odot G'^{(n)}(\mathbf{Z}^{(n)}), \tag{II.57}$$

for all $n = L-1, L-2, \ldots, 1$. This step determines how errors propagate backward through the network. Finally, compute the gradients of the weights and biases using $\delta^{(n)}$:

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{W}^{(n)}} &= \delta^{(n)} (\mathbf{X}^{(n-1)})^T \\ \frac{\partial L}{\partial \mathbf{b}^{(n)}} &= \delta^{(n)}. \end{aligned} \tag{II.58}$$

4. **Update Weights and Biases:** After computing the gradients, update the weights and biases using **gradient descent**:

$$\begin{aligned} \mathbf{W}^{(n)} &\leftarrow \mathbf{W}^{(n)} - \eta \frac{\partial L}{\partial \mathbf{W}^{(n)}}, \\ \mathbf{b}^{(n)} &\leftarrow \mathbf{b}^{(n)} - \eta \frac{\partial L}{\partial \mathbf{b}^{(n)}}, \end{aligned} \tag{II.59}$$

   where η is the **learning rate**, a hyperparameter that controls the step size of updates.

5. **Iteration Until Convergence:** The entire process (forward pass, loss computation, backward pass, and parameter update) is repeated iteratively until the loss function converges, meaning the neural network learns to approximate the target function.

This algorithm enabled efficient training of the weights and sparked the machine learning revolution we witness today. However, revisiting the function of a feedforward neural network (FFNN), it becomes evident that its state does not depend on previous inputs. For sequential data—such as weather prediction, stock market analysis, or real-time monitoring—the temporal correlation between data points is crucial. Therefore, a more suitable architecture involves designing neural networks that incorporate information from previous inputs. One of the earliest proposals in this direction is the recurrent neural network (RNN), which is defined below.

Recurrent Neural Network (RNN)

A Recurrent Neural Network (RNN) is a neural network designed to process sequential data by maintaining a state over discrete time steps. Consider a discrete time sequence $\{t_l\}$ where $t_l \in \mathbb{R}$ for $l = 1, 2, \ldots, n_T$. Let $\mathbf{f} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ be a non-linear, differentiable function defining the state update, where:

- $\mathbf{u}(t_l) \in \mathbb{R}^{n_u}$ is the input at time t_l
- and where $\mathbf{x}(t_l) \in \mathbb{R}^{n_x}$ represents the state of the system.

At time t_{l+1}, the function $\mathbf{f}$ maps the inputs

$$\mathbf{u}(t_{l+1}) = (u_1(t_{l+1}), u_2(t_{l+1}), \ldots, u_{n_u}(t_{l+1})) \in \mathbb{R}^{n_u}$$

and the current state

$$\mathbf{x}(t_l) = (x_1(t_l), x_2(t_l), \ldots, x_{n_x}(t_l)) \in \mathbb{R}^{n_x}$$

to the updated state

$$\mathbf{x}(t_{l+1}) = (x_1(t_{l+1}), x_2(t_{l+1}), \ldots, x_{n_x}(t_{l+1})).$$

The update equation is given by:

$$x_i(t_{l+1}) = f\left(\sum_{j=1}^{n_u} \mathbf{W}^{\text{in}}_{i,j} u_j(t_{l+1}) + \sum_{k=1}^{n_x} \mathbf{W}^{\text{int}}_{i,k} x_k(t_l)\right), \tag{II.60}$$

where:

- $\mathbf{W}^{\text{in}}_{i,j}$ represents the input weights and
- $\mathbf{W}^{\text{int}}_{i,k}$ represents the internal recurrent weights.

For an RNN with n_{out} readout nodes $\{s_i\} \subset \{x_i\}$, the output at time t_{l+1} is computed as:

$$y_i(t_{l+1}) = \sum_{j=1}^{n_{\text{out}}} \mathbf{W}^{\text{out}}_{i,j} s_j(t_{l+1}), \tag{II.61}$$

where $\mathbf{W}^{\text{out}}_{i,j}$ refers to the readout weights.

II.3.2 Reservoir Computing

One major challenge with Feedforward Neural Networks (FFNNs) and Recurrent Neural Networks (RNNs) is that they require extensive training. This is especially true in the case of RNNs, where the backpropagation algorithm must consider all previous time steps, resulting in a costly training procedure. Therefore, one might ask whether this problem can be mitigated and whether training speed can be improved. One possible approach is to randomly initialize the internal weights and train only a linear readout layer, resulting in a convex optimization problem.

The hidden layer is interpreted as a mapping onto a high-dimensional space, allowing this interpretation to be leveraged to replace the hidden layer with any mapping or system that transforms an initial state into a high-dimensional representation. This is where reservoir computing (RC) comes into play. Instead of using a conventional hidden layer, a physical system can be utilized. These physical systems can vary widely, including mechanical systems, electrical circuits, lasers, or even quantum systems. Most of these systems are described by differential equations, and measuring their physical quantities is used to construct an output. The measured variables might include currents or voltages in circuits, pressure or deflection in mechanical oscillators, electric fields or light intensities in laser systems, or expectation values of local observables in quantum systems.

Compared to classical machine learning, reservoir computing requires a thorough understanding of the underlying physics, including phase transitions and bifurcations in these systems. Analyzing a system as a reservoir can often yield novel insights into its underlying dynamics. One interesting feature is that reservoirs often perform best at the edge of chaos. This implies that the dynamics become complex yet remain predictable, ensuring that the reservoir computer can still be trained effectively within that parameter regime.

Reservoir Computer

In the most general case, a reservoir is a physical system described by a differential equation:

$$\frac{d}{dt}\mathbf{x}(t) = f(\mathbf{x}(t)) \tag{II.62}$$

where $\mathbf{x}(t) \in \mathbb{R}^{N_x}$ represents the state of the system.

The input $\mathbf{u}(t_l) \in \mathbb{R}^{n_u}$ is introduced into the system at discrete time steps, making the system's evolution input-dependent:

$$\frac{d}{dt}\mathbf{x}(t_l) = f(\mathbf{x}(t_l), \mathbf{u}(t_l)). \tag{II.63}$$

The time between two inputs, denoted as $T := t_{l+1} - t_l$, is called the clock cycle. The readout is constructed based on the state of the system $\mathrm{s}_l = \mathbf{x}(t_l)$ or a subset of the state $\mathrm{s}_l \subset \mathbf{x}(t_l)$, with $\mathbf{s}_i \in \mathbb{R}^{N_S}$.

Given an input series $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M)$, where $\mathbf{u}_i = \mathbf{u}(t_i)$, and the corresponding reservoir states $(\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M)$, the output is given by:

$$\mathbf{Y} = \mathbf{S}\mathbf{W}^{\text{out}} \tag{II.64}$$

where $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_M)^T$ represents the outputs, and $\mathbf{S} = (\mathbf{s}_1, \ldots, \mathbf{s}_M)^T$ is called the state matrix. The weights $\mathbf{W} \in \mathbb{R}^{N_S}$ are trained using linear regression:

$$\mathbf{W}^{\text{out}} = (\mathbf{S}^T\mathbf{S} + \lambda\mathbf{I})^{-1}\mathbf{S}^T\mathbf{Y}^{\text{targ}} \tag{II.65}$$

where $\mathbf{Y}^{\text{targ}}$ represents the vector of all target outputs. λ is a regularization parameter.

II.3.3 Information Processing Capacity and prediction tasks

Understanding typically emerges through a well-defined quantity that captures key aspects of a system. By interpreting this quantity, one can infer the system's behavior, enabling optimization or new discoveries. Physical systems are often characterized by their stability, but alternative insights can arise through measures of expressivity, offering a different lens for understanding. These measures aim to quantify how well a system can map data onto a set of outputs and have found various applications. In reservoir computing, the most established data expressivity measure is the **information processing capacity** (IPC). This measure quantifies how well a system can retain and non-linearly transform input data. To compute the IPC, Legendre polynomials l_i are used as a set of orthogonal functions (Definition IPC.1). The computation of IPC involves two combinatorial challenges (Definition IPC.2). First, all possible combinations of the Legendre polynomials of the same polynomial order must be calculated, as described in Definition IPC.2. Second, all valid combinations of past inputs must be determined, as outlined in Definition IPC.3. Summing over all orders of the information processing capacity, results in the total IPC or just IPC (see Def. IPC.4)

Definition IPC.1: Legendre Polynomials The Legendre polynomials $l_n : [-1, 1] \to [-1, 1]$ are defined as follows:

$$\begin{aligned}
l_0(x) &= 1 \\
l_1(x) &= x \\
l_2(x) &= \frac{1}{2}(3x^2 - 1) \\
l_{n+1}(x) &= \frac{2n+1}{n+1} x l_n(x) - \frac{n}{n+1} l_{n-1}(x), \text{ for } n \geq 2
\end{aligned} \tag{II.66}$$

They are orthogonal under the scalar product:

$$\int_{-1}^{1} l_i(x) l_j(x) dx = \delta_{i,j} \tag{II.67}$$

Definition IPC.2: Combinations of Legendre Polynomials Let k be the order of a polynomial, and let $\mathcal{D}^k = [d_1, d_2, .., d_q]$ be a tuple where $0 < d_j \leq k$ and $d_a \leq d_b$ for $a < b$, such that $\sum_{j=1}^{q} d_j = k$. For each order k, a finite number of tuples N_k satisfying these conditions exist. The o-th tuple is denoted as $\mathcal{D}_o^k = [d_{o,1}, d_{o,2}, .., d_{o,q}]$, where $o \in \{1, 2, .., N_k\}$. Let M_o^k be the length of $\mathcal{D}_o^k$. The o-th polynomial of order k is given by:

$$p_o^k = \prod_{j=1}^{M_o^k} l_{d_{o,j}}, \tag{II.68}$$

where $l_{d_{o,j}}$ is the $d_{o,j}$-th Legendre polynomial.

We will provide an example for a polynomial of order $k = 3$. There exist three valid combinations satisfying the above conditions: $\mathcal{D}_1^3 = [d_{1,1}] = [3]$, $\mathcal{D}_2^3 = [d_{2,1}, d_{2,2}] = [1, 2]$, and $\mathcal{D}_3^3 = [d_{3,1}, d_{3,2}, d_{3,3}] = [1, 1, 1]$. From these tuples, the lengths $M_1^3 = 1$, $M_2^3 = 2$, and $M_3^3 = 3$ can be calculated. The polynomials of the third order are given by:

$$\begin{aligned} p_1^3 &= \prod_{j=1}^{M_1^3} l_{d_{1,j}} = l_3 \\ p_2^3 &= \prod_{j=1}^{M_2^3} l_{d_{2,j}} = l_1 l_2 \\ p_3^3 &= \prod_{j=1}^{M_3^3} l_{d_{3,j}} = l_1 l_1 l_1. \end{aligned} \tag{II.69}$$

Assume the input vector $\mathbf{u}_{all} = (u_{-b+1}, u_{-b+2}, .., u_0, u_1, u_2.., u_n)$. This vector can be split into the vector for the state matrix $\mathbf{u} = (u_1, u_2.., u_n)$ and the prior vector $\mathbf{u}_{prior} = (u_{-b+1}, u_{-b+2}, .., u_0)$. The values $u(m)$ and $u(m - i_1)$ denote the m-th input and the input shifted by i_1 time steps into the past, where $-b \leq m \leq n$ and $i_1 \leq b$. The system vector $\mathbf{u}$ shifted by i_1 time steps into the past is denoted as $\mathbf{u}(i_1) = (u_{1-i_1}, u_{2-i_1}.., u_{n-i_1})$. The a-th Legendre polynomial of the input vector $\mathbf{u}$ is given by:

$$\mathbf{l}_a(\mathbf{u}) = (l_a(u_1), l_a(u_2), .., l_a(u_n)). \tag{II.70}$$

The shifted a-th Legendre polynomial is defined as:

$$\mathbf{l}_a(i_1) = \mathbf{l}_a(\mathbf{u}(i_1)) = (l_a(u_{1-i_1}), l_a(u_{2-i_1}), .., l_a(u_{n-i_1})). \tag{II.71}$$

The element-wise multiplication between two vectors $\mathbf{a} = (a_1, a_2, .., a_n)$ and $\mathbf{b} = (b_1, b_2, .., b_n)$ results in a new vector:

$$\mathbf{c} = \mathbf{a} \odot \mathbf{b} = (a_1, a_2, .., a_n) \odot (b_1, .., b_n) = (a_1 b_1, a_2 b_2, .., a_n b_n). \tag{II.72}$$

The polynomials for different time steps are calculated by

$$\mathbf{p}_o^k(i_1, \dots, i_q) = \bigodot_{j=1}^{M_o^k} \mathbf{l}_{d_{o,j}}(i_j). \tag{II.73}$$

As an example, the polynomials of third order with respect to the past inputs i_j are given by:

$$\begin{aligned}
\mathbf{p}_1^3(i_1) &= \bigodot_{j=1}^{M_1^3} \mathbf{l}_{d_{1,j}}(i_j) = \mathbf{l}_3(i_1), \\
\mathbf{p}_2^3(i_1, i_2) &= \bigodot_{j=1}^{M_2^3} \mathbf{l}_{d_{2,j}}(i_j) = \mathbf{l}_1(i_1)\mathbf{l}_2(i_2), \\
\mathbf{p}_3^3(i_1, i_2, i_3) &= \bigodot_{j=1}^{M_3^3} \mathbf{l}_{d_{3,j}}(i_j) = \mathbf{l}_1(i_1)\mathbf{l}_1(i_2)\mathbf{l}_1(i_3).
\end{aligned} \tag{II.74}$$

Definition IPC.3: Legendre Memory Accuracy
Assume that the target function is of the form

$$\mathbf{y}^{\text{targ}}(i_1, i_2, \dots) = \mathbf{p}_l^k(i_1, i_2, \dots)$$

for one tuple $(i_1, i_2, \dots)$. The capacity of such a target is defined as

$$\begin{aligned}
C_{k,l}^{\text{comb}}(i_1, i_2, \dots) &= \frac{\operatorname{cov}(\mathbf{y}, \mathbf{y}^{\text{targ}}(i_1, i_2, \dots))}{\sigma^2(\mathbf{y})\,\sigma^2(\mathbf{y}^{\text{targ}}(i_1, i_2, \dots))} \\
&= \frac{\operatorname{cov}(\mathbf{y}, \mathbf{p}_l^k(i_1, i_2, \dots))}{\sigma^2(\mathbf{y})\,\sigma^2(\mathbf{p}_l^k(i_1, i_2, \dots))}.
\end{aligned} \tag{II.75}$$

Definition IPC.3: Combinations of Past Inputs
Valid combinations of $(i_1, i_2, \dots)$ are those that comply with the following two conditions:

1. Only combinations are allowed where all i_1, i_2, $\dots$ are distinct (i.e., $i_a \neq i_b$ if $a \neq b$).

2. The second condition is best understood with an example.
   Let
   $$\mathbf{p}_1^7(i_1, \dots, i_4) = \mathbf{l}_1(i_1)\mathbf{l}_1(i_2)\mathbf{l}_1(i_3)\mathbf{l}_2(i_4)\mathbf{l}_2(i_5).$$
   There are three Legendre polynomials of first order with indices (i_1, i_2, i_3) and two Legendre polynomials of second order with indices (i_4, i_5). The second condition applies to indices corresponding to Legendre polynomials of the same order. In these subsets, only decreasing indices are allowed, meaning $i_1 > i_2 > i_3$ and $i_4 > i_5$. Note that i_4 can be greater than i_3. One valid combination is
   $$(i_1, i_2, i_3, i_4, i_5) = (5, 4, 2, 3, 1).$$

Example IPC

The following example illustrates how these combinations should be chosen. Let $(i_1, i_2, \dots, i_n) \in \text{comb}$, where the set comb consists of all valid combinations. The combinations for $1 \le i_j \le 4$ for the first three orders of polynomials and one realization of fourth order are given by:

Order	Polynomial	Index Set
1	$p_{1,1}(i_1) = l_1(i_1)$	$i_1 \in \{1,2,3,4\}$
2	$p_{2,1}(i_1) = l_2(i_1)$	$i_1 \in \{1,2,3,4\}$
	$p_{2,2}(i_1, i_2) = l_1(i_1)l_1(i_2)$	$(i_1, i_2) \in \{(2,1),(3,1),(4,1),(3,2),(4,2),(4,3)\}$
3	$p_{3,1}(i_1) = l_3(i_1)$	$i_1 \in \{1,2,3,4\}$
	$p_{3,2}(i_1, i_2) = l_1(i_1)l_2(i_2)$	$(i_1, i_2) \in \{(2,1),(3,1),(4,1),(1,2),(3,2),(4,2),(1,3),(2,3),(4,3),(1,4),(2,4),(3,4)\}$
	$p_{3,3}(i_1, i_2, i_3) = l_1(i_1)l_1(i_2)l_1(i_3)$	$(i_1, i_2, i_3) \in \{(3,2,1),(4,2,1),(4,3,1),(4,3,2)\}$
4	$p_{4,4}(i_1, i_2, i_3) = l_1(i_1)l_1(i_2)l_2(i_3)$	$(i_1, i_2, i_3) \in \{(3,2,1),(4,2,1),(4,3,1),(3,1,2),(4,1,2),(4,3,2),(2,1,3),(4,1,3),(4,2,3),(2,1,4),(3,1,4),(3,2,4)\}$

Table II.1: Selecting targets for computing the information processing capacity.

This example should help the reader better understand how the i_j are selected for each polynomial. In the next step, a sum over all valid combinations $\text{comb}_{k,l}$ of $C_{k,l}^{\text{comb}}$ is taken to obtain the capacity $C_{k,l}$ with respect to the target polynomial $\mathbf{p}_l^k$:

$$C_{k,l} = \sum_{\text{comb}_{k,l}} C_{k,l}^{\text{comb}}(i_1, i_2, \dots). \tag{II.76}$$

Definition IPC.4: Information Processing Capacity (IPC)

As described in Def. 3.1, N_k different polynomial combinations of degree k exist, where $o \in \{1, 2, \dots, N_k\}$ is the index for these polynomials. The k-th order capacity C_k is given as the sum over all polynomial combinations of degree k:

$$\mathrm{IPC}_k = \sum_l C_{k,l}, \tag{II.77}$$

and the overall **information processing capacity (IPC)** is given by

$$\mathrm{IPC} = \sum_k \mathrm{IPC}_k. \tag{II.78}$$

The information processing capacity (IPC) is often used as a measure to describe how the nonlinear behavior of the past inputs can be characterized. It should be noted that further constraints on the combinations need to be imposed, as the system exhibits the fading memory property. This means that

$$\lim_{i_1 \to \infty} \to 0,$$

and a termination condition needs to be added. Additionally, the IPC is upper bounded by

$$C \leq N_R,$$

where N_R is the number of readout nodes [51].

Time Series Prediction Tasks

While information processing capacity (IPC) aims to quantify how well a system can retain and nonlinearly map past inputs, it often fails to capture a system's ability to predict future states in time series prediction tasks. Therefore, reservoir computing is frequently evaluated using benchmark prediction tasks to assess its predictive performance.

Typically, one samples a real-valued time series $u(t)$, representing real-world data at discrete time steps $u_i = u(t_i)$, resulting in a sequence $\mathbf{u} = (u_1, \dots, u_M)$, where $u_i \in \mathbb{R}^m$. The sampling interval $\Delta t = t_{i+1} - t_i$ is generally assumed to be fixed.

The goal of time series prediction tasks is to predict future values of the time series or other quantities that are assumed to depend on it. These tasks can be categorized into three types:

1. **Direct prediction:** Given u_n, predict u_{n+p}. This is a standard form of time series forecasting, such as stock market prediction.

2. **Cross prediction (function of present):** Given u_n, predict $f(u_n)$. This involves predicting a function of the current state, often used when the full state of the system is unknown or when exploring correlations between system variables.
3. **Cross prediction (function of future):** Given u_n, predict $f(u_{n+p})$. This corresponds to predicting the future state of an unknown or indirectly observed variable. Applications include earthquake forecasting—where local sensor data is used to predict seismic activity—or weather prediction based on local meteorological measurements.

Lorenz Prediction Task

The Lorenz attractor is a widely used benchmark system for evaluating time series prediction tasks in reservoir computing. Although several variations of the Lorenz system exist, this work employs the following standard setup.

The continuous Lorenz time series is discretized to produce three sequences: $\mathbf{x} = (x_1, \ldots, x_N)$, $\mathbf{y} = (y_1, \ldots, y_N)$, and $\mathbf{z} = (z_1, \ldots, z_N)$.

This study investigates the following three prediction tasks:

- **Future prediction:** Use x_n as input to predict x_{n+p}.
- **Cross prediction:** Use x_n as input to predict z_n.
- **Future cross prediction:** Use x_n as input to predict z_{n+p}.

II.4 Quantum reservoir computing

Exploiting the exponential scaling of the Hilbert space in quantum mechanics for computation has resulted in the current interest in quantum computation. However, current hardware exhibits dissipation and noise, resulting in shallow circuits and few qubit operations, rendering them impossible for a possible quantum algorithms. Quantum machine learning on the other hand is known to be more robust to noise and has therefore garnered a lot of attention recently. Within QML, quantum reservoir computing (QRC) has gained a lot of attention recently. Quantum reservoir computing is inspired by classical reservoir computing, where physical systems are used to solve machine learning tasks [19, 20, 21, 22, 23, 24, 25]. Significant progress in quantum reservoir computing (QRC) has been made through investigations using the Ising model [26, 27, 28, 29, 30, 31, 32, 33], with early experimental realizations conducted on IBM's quantum hardware [34, 35, 36, 37]. Interestingly, studies have demonstrated that inherent quantum features such as noise and dissipation can be leveraged as beneficial computational elements within QRC frameworks [38, 39, 40]. To tackle the challenges of time-series tasks, particularly with respect to computational overhead, methods such as weak measurement protocols [41, 42] and system reinitialization strategies [43] have been introduced, both contributing to more efficient and accurate learning. Additional advancements include techniques where output data is fed back into the reservoir [44, 45], as well as the use of continuous measurement dynamics [46].

II.4.1 Quantum Reservoir Computing Scheme

Figure II.4 illustrates a quantum reservoir scheme, which is explained in the following.

1. **Initialization:** The reservoir is first initialized with an input time series of length N_{In}.

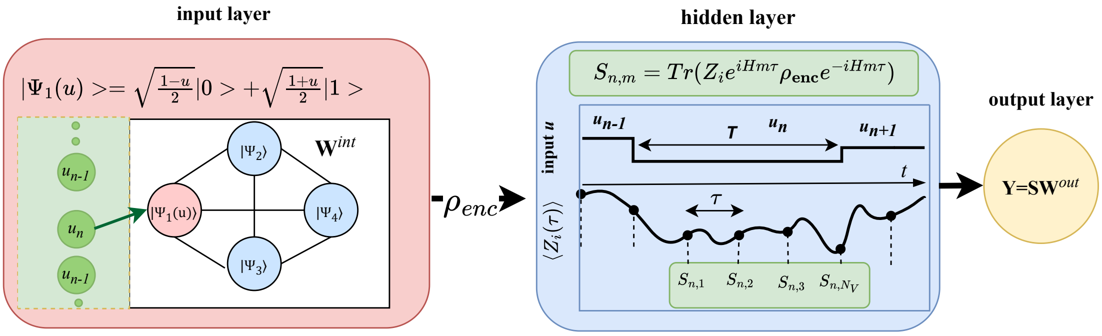

input layer
$|\Psi_1(u)\rangle = \sqrt{\frac{1-u}{2}}|0\rangle + \sqrt{\frac{1+u}{2}}|1\rangle$
u_{n-1}
u_n
u_{n-1}
$|\Psi_1(u)\rangle$
$|\Psi_2\rangle$
$|\Psi_3\rangle$
$|\Psi_4\rangle$
W^{int}
ρ_{enc}
hidden layer
$S_{n,m} = Tr(Z_i e^{iHm\tau} \rho_{enc} e^{-iHm\tau})$
input u
u_{n-1}
T
u_n
u_{n+1}
t
τ
$\langle Z_i(\tau)\rangle$
$S_{n,1}$
$S_{n,2}$
$S_{n,3}$
S_{n,N_V}
output layer
Y=SW^{out}

Figure II.4: A time series u_n is encoded into a qubit of a quantum system, described by a Hamiltonian H. The system state, represented by the density matrix ρ_{enc}, evolves over time t. Measurements of selected observables Z_i at times $m\tau$ are used to construct the state matrix $\mathbf{S}$. The total number of measurements per observable is V and the clock cycle is defined as $V\tau = T$. In the output layer, the state matrix $\mathbf{S}$ is multiplied by the readout weights $\mathbf{W}^{out}$ to produce the output vector $\mathbf{Y} = \mathbf{S}\mathbf{W}^{out}$.

2. **Input Layer (green) and Encoding (red):** The n-th input $u_n \in \{u_1, \ldots, u_{N_u}\}$ is encoded into the first qubit, resulting in the state

$$|\Psi_n\rangle = \sqrt{\frac{1-u_n}{2}}\,|0\rangle + \sqrt{\frac{1+u_n}{2}}\,|1\rangle\,. \tag{II.79}$$

After encoding, the system state is described by

$$\rho_{enc} = |\Psi_n\rangle\langle\Psi_n| \otimes \mathrm{Tr}_1(\rho_{n-1}). \tag{II.80}$$

Overwriting the first qubit at each time step leads to a loss of information from previous steps. This fading memory property is a critical requirement for reservoir computing, as reservoirs inherently exhibit a memory–nonlinearity trade-off [19, 20, 52, 53, 54]. In [43], this trade-off is demonstrated for quantum reservoir computing.

3. **Hidden Layer (blue):** The reservoir undergoes unitary evolution for a clock cycle T via

$$U_R = e^{-iHT}.$$

A discrete set of observables $\{Z_i\}_{i=1}^{K}$ is measured to obtain their expectation values. The expectation value of observable Z_i at time t after the n-th input is given by

$$\langle Z_i(t)\rangle = \mathrm{Tr}\Big(Z_i\, e^{-iHt}\,\rho_{enc}\, e^{iHt}\Big). \tag{II.81}$$

To increase the readout dimension, each observable is sampled V times within each clock cycle T, i.e., at times $t_{n,j} = nT + (j+1)\tau$ with $\tau = T/V$. Figure II.4 shows a quantum reservoir with four virtual nodes ($V = 4$). This results in $N_R = VR$ readout nodes for each input u_n, where R is the number of measured observables. Writing the readouts for each input as a row produces the state matrix $\mathbf{S} \in \mathbb{R}^{N_u \times N_R}$.

$$\mathbf{S} = \begin{pmatrix} \langle Z_1(t_{1,1})\rangle & \dots & \langle Z_R(t_{1,1})\rangle & \dots & \langle Z_1(t_{1,V})\rangle & \dots & \langle Z_R(t_{1,V})\rangle \\ \langle Z_1(t_{2,1})\rangle & \dots & \langle Z_R(t_{2,1})\rangle & \dots & \langle Z_1(t_{2,V})\rangle & \dots & \langle Z_R(t_{2,V})\rangle \\ \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ \langle Z_1(t_{N_u,1})\rangle & \dots & \langle Z_R(t_{N_u,1})\rangle & \dots & \langle Z_1(t_{N_u,V})\rangle & \dots & \langle Z_R(t_{N_u,V})\rangle \end{pmatrix} \tag{II.82}$$

4. **Output Layer (yellow):** The state matrix is multiplied by the readout weights $\mathbf{W}^{out}$ to yield the output

$$\mathbf{Y} = \mathbf{S}\mathbf{W}^{\mathrm{out}}. \tag{II.83}$$

5. **Training:** The data is split into an initialization set u_{In}, a training set u_{Tr}, and a testing set u_{Te} with lengths N_{In}, N_{Tr}, and N_{Te}, respectively. A buffer of length N_b is inserted between the sets.

   First, the system is initialized with N_{In} steps. Then, the weights are optimized using the training set u_{Tr}, resulting in the trained weights $\mathbf{W}^{\mathrm{out}}$. The testing set is used to evaluate whether the reservoir can generalize to new data. The weights are adjusted to minimize the loss

$$L := (\mathbf{Y} - \mathbf{Y}^{\mathrm{targ}})^2, \tag{II.84}$$

where $\mathbf{Y}^{\text{targ}}$ is the target vector. To account for measurement shot noise, Gaussian-distributed noise $\mathcal{N} \in \mathbb{R}^{N_u \times N_R}$ is added to the state matrix:

$$\mathbf{S} \leftarrow \mathbf{S} + \eta\mathcal{N}. \tag{II.85}$$

The noise term $\eta\mathcal{N}$ can serve as a regularization parameter when computing the state matrix, a technique often referred to as regularization by noise. The readout weights $\mathbf{W}^{out}$ are then determined by

$$\mathbf{W}^{\text{out}} = (\mathbf{S}_{Tr}^{T}\mathbf{S}_{Tr})^{-1}\mathbf{S}_{Tr}^{T}\mathbf{Y}^{\text{targ}}. \tag{II.86}$$

II.4.2 Motivation of Quantum Reservoir Computing

As a quantum reservoir, this work uses the Transverse Field Ising Model discussed in Section V.1.1 and given by

$$H = \sum_{i>j} J_{ij}\sigma_x^{(i)}\sigma_x^{(j)} + \frac{1}{2}\sum_{i=1}^{N}(h + D_i)\sigma_z^{(i)}. \tag{II.87}$$

The on-site disorder is turned off with $D_i = 0$ and $h = 0.5$, while the inter-spin couplings are sampled from a uniform distribution $J_{ij} \in \mathcal{U}([0.25, 0.25])$, unless stated otherwise. This chapter will explain some features of quantum reservoir computing and discuss some of its unique traits, which will be analyzed in the later parts of this work.

To give a brief overview of quantum reservoir computing, the IPC and the Lorenz tasks will be computed for varying parameters in the quantum reservoir computing scheme. For the Lorenz task, the x variable will be used as an input and discretized with $\Delta t = 0.02$, resulting in a discrete time series $\mathbf{x} = (x_1, x_2, \ldots, x_N)$. The two tasks will be to predict the x variable of the Lorenz tasks five steps into the future, i.e., given x_n, predict x_{n+5}; and another task will consist of cross-predicting the z variable of the Lorenz task (z_n).

The following results are based on a training set of $N_{\text{Tr}} = 10{,}000$ steps and a test set of $N_{\text{Te}} = 5{,}000$ steps. One of the main motivations behind using quantum reservoir computing is its potential for exponential scaling with respect to the Hilbert space dimension. Figure II.5 presents both the information processing capacity up to third order (a) and the Lorenz task performance (b), as a function of the number of spin sites N_s. The parameters are kept fixed with a clock cycle of $T = 50$ and $V = 50$ measurements, where all spins are measured along the σ_z-axis.

Panel (a) illustrates that a five-site system reaches an information processing capacity of IPC $= 120$, which is notably high. Initially, the IPC appears to increase exponentially before transitioning into a linear regime. This behavior may arise from several factors. First, while the readout dimension increases linearly, the Hilbert space dimension grows exponentially. Second, the IPC calculation is limited to third-order correlations, which may restrict its ability to fully capture the reservoir's representational capacity. Despite these limitations, the exponential growth observed from $N_s = 2$ to $N_s = 4$ indicates that the system's data expressivity increases rapidly with size—even in this unoptimized setup.

The IPC is a well established data expressivity, but more insights might be gained from the behavior of task performance. Therefore, panel (b) of Figure II.5 shows the Lorenz prediction task results for $x \to x$ and the cross-prediction task $x \to z$. In the logarithmic plot, the prediction error for the direct task (orange) decreases linearly, which indicates exponential improvement as the number of spin sites increases. On the other hand, the cross-prediction task (green) flattens out, showing diminishing returns with higher N_s.

Understanding whether a quantum reservoir truly leverages its exponentially growing Hilbert space for learning remains an open research question. This work does not attempt to resolve these issues, but instead aims to offer a clearer picture of what quantum reservoir computing could potentially achieve. The core idea is compelling: achieving a nonlinear boost in performance through a modest increase in system size.

Ultimately, the objective is to design systems with minimal readout dimensions and few qubits that still offer exponential gains in learning performance.

Another important aspect of quantum reservoir computing concerns the readout dimension of the system, i.e., the number of measurements V. Figure II.6 shows the IPC and Lorenz task performance versus the number of measurements for a system with four spin sites $N_s = 4$ and a clock cycle of $T = 50$.

The IPC (Figure II.6.a) increases and then saturates around IPC ≈ 63 for approximately $V \approx 25$ measurements. This is somewhat atypical, since in classical reservoir computing there is a region with optimal readout dimension and no saturation. The Lorenz task performance shows a steep initial improvement: up to $V = 5$ for the prediction task (orange), and up to $V = 15$ for the cross-prediction task (green). Beyond these points, the performance gradually transitions into a saturation phase around $V \approx 35$ for both tasks.

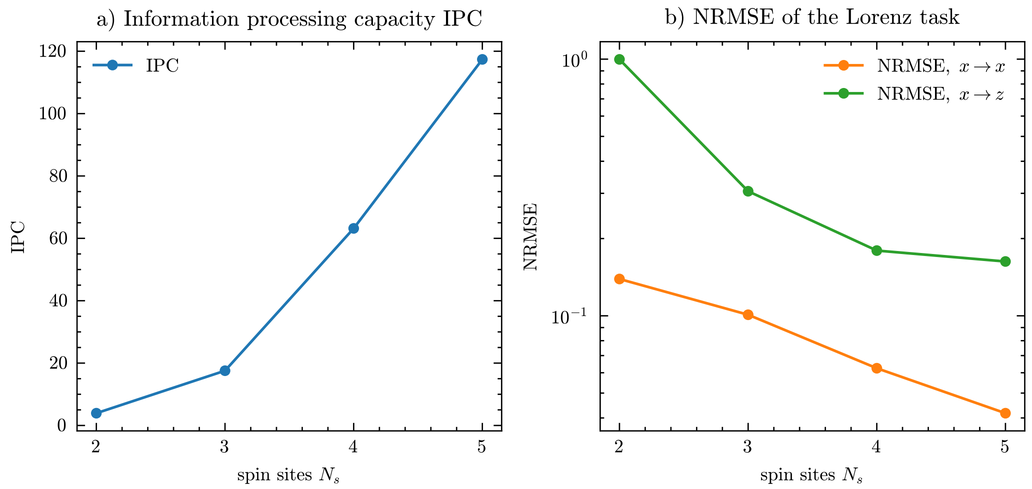


Figure II.5: Quantum reservoir performance as a function of the number of spin sites N_s. (a) Information processing capacity (IPC) up to third order. (b) Prediction error for the Lorenz task using direct prediction $x \to x$ (orange) and cross-prediction $x \to z$ (green). The plots demonstrate strong scaling in task performance for small system sizes.

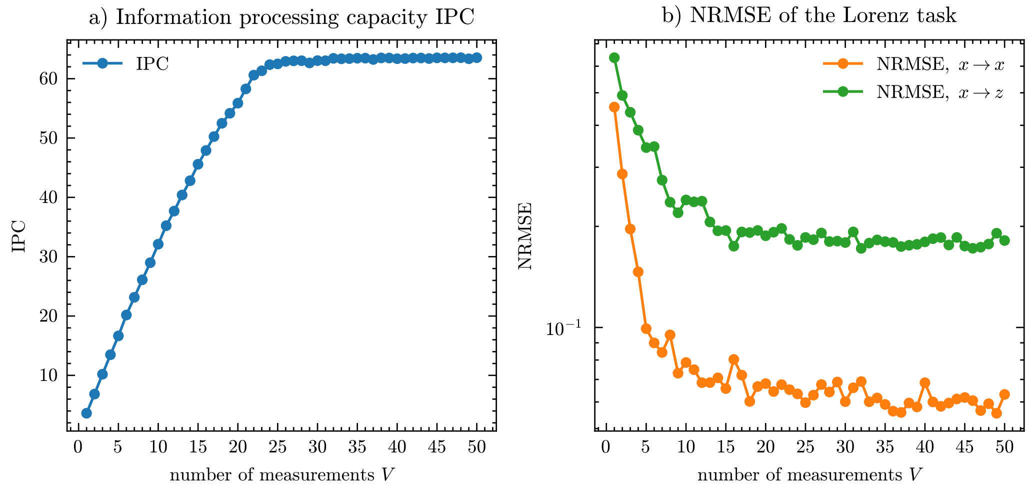


Figure II.6: Effect of the number of measurements V on quantum reservoir computing performance for a system with $N_s = 4$ spin sites. (a) Information Processing Capacity (IPC) shows rapid growth followed by saturation. (b) Lorenz task prediction error for direct prediction $x \to x$ (orange) and cross-prediction $x \to z$ (green), both showing early improvement and eventual saturation.

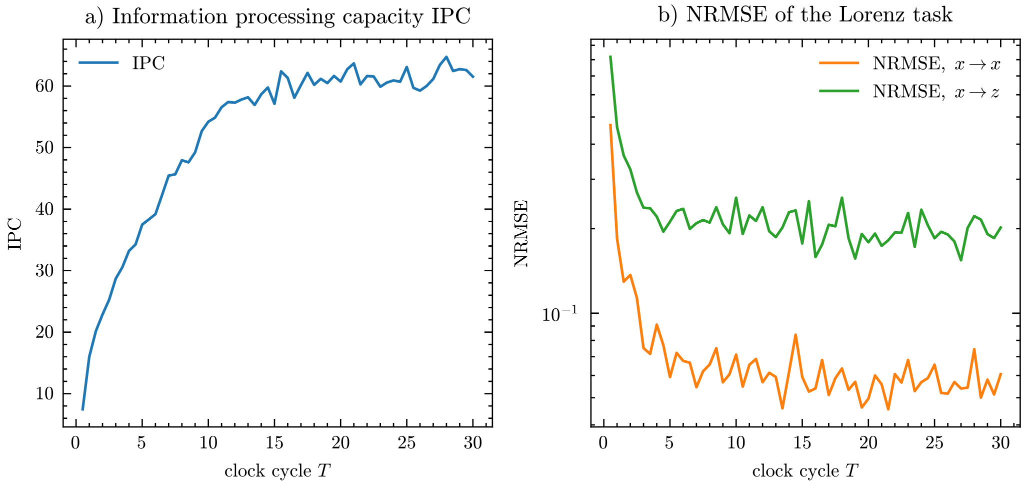


Figure II.7: Dependence of information processing capacity (a) and Lorenz task performance (b) on the clock cycle T for a quantum reservoir with $N_s = 4$ spin sites and $V = 30$ measurements. Both metrics show saturation behavior around $T \approx 20$, which differs from typical classical reservoir behavior.

Lastly, the behavior of the IPC and task performance is analyzed as a function of the time interval between two inputs, i.e., the clock cycle T, for a four-qubit system $(N_s = 4)$ with $V = 30$ measurements per observable. A saturation effect is observed once again. This behavior is atypical, as most classical reservoir computers exhibit a distinct region of optimal performance with respect to the clock cycle T.

Figure II.7.a displays the IPC, where saturation occurs around $T \approx 20$. The Lorenz task performance also increases up to $T \approx 20$, after which it saturates, with minor oscillatory fluctuations observed.

III Chapter: Krylov Complexity in Unitary Time Evolution

As discussed in Section II.2, quantum systems are described by a Hamiltonian H acting on a Hilbert space $\mathcal{H}$. This description will be briefly revisited in the context of this chapter. In the case of pure states $|\Psi\rangle$ with initial condition $|\Psi_0\rangle$, the dynamics are governed by the Schrödinger equation

$$\partial_t |\Psi(t)\rangle = -iH |\Psi(t)\rangle \,, \tag{III.1}$$

$$|\Psi(0)\rangle = |\Psi_0\rangle \,. \tag{III.2}$$

This results in the unitary time evolution of the quantum state, described by

$$|\Psi(t)\rangle = e^{-iHt} |\Psi_0\rangle \,. \tag{III.3}$$

The density matrix

$$\rho(t) = |\Psi(t)\rangle \langle\Psi(t)| \tag{III.4}$$

is introduced and is an important concept in quantum mechanics.

Observing a quantum system typically requires the application of a measurement process, usually described by Hermitian observables O, with eigenvalue equation $\mathrm{O} |o_i\rangle = o_i |o_i\rangle$. The probability of measuring the eigenvalue o_i at time t is given by

$$P(m = o_i,\ t) = |\langle o_i|\Psi(t)\rangle|^2 = \langle o_i|\Psi(t)\rangle \langle\Psi(t)|o_i\rangle = \langle o_i| \rho(t) |o_i\rangle \,. \tag{III.5}$$

The probabilistic nature allows the definition of an expectation value of the observable O, given by

$$\langle \mathrm{O} \rangle = \langle \Psi | \, \mathrm{O} \, | \Psi \rangle = \mathrm{Tr}(\rho \mathrm{O}). \tag{III.6}$$

If the quantum state is described by a density matrix as

$$\rho_0 = \sum_i p_i \, |\Phi_i\rangle \, \langle \Phi_i| \, . \tag{III.7}$$

, then this state is called mixed, where p_i represents the probability of the state being in $|\Phi_i\rangle \, \langle\Phi_i|$. The equation of motion for such systems is

$$\partial_t \rho(t) = -i[H, \rho], \quad \rho(0) := \rho_0. \tag{III.8}$$

The solution to this equation is the unitary time evolution for observables:

$$\rho(t) = e^{-iHt} \rho_0 e^{iHt}. \tag{III.9}$$

Similarly, instead of describing the evolution of states, one can describe the evolution of an operator, which results in the Heisenberg equation

$$\partial_t \mathrm{O}(t) = i[H, \mathrm{O}], \quad \mathrm{O}(0) := \mathrm{O}_0 \tag{III.10}$$

with the solution

$$\mathrm{O}(t) = e^{iHt} \mathrm{O}_0 e^{-iHt}. \tag{III.11}$$

In this chapter, Krylov-based information measures are discussed mathematically, particularly with respect to how well they capture the behavior of time evolution in quantum mechanics. These measures are defined based on how effectively the state or operator evolves over time within the Krylov basis. Throughout this work, the term ”Krylov complexity” will be used to refer to all measures that are defined in relation to the Krylov space. Special focus is given to Krylov operator complexity introduced in [14] and Krylov spread complexity introduced in [16]. These measures have recently garnered significant attention and were initially proposed to quantify quantum chaos in systems ranging from Ising to SYK models [14, 55, 56, 57, 58, 59, 60, 61, 62, 63]. The authors demonstrated that Krylov complexity can successfully differentiate between integrable and chaotic systems. Several relationships involving Krylov complexity have been explored. It has also been discussed in connection with Nielsen complexity [64], operator entropy [65], and its geometrical aspects have been analyzed in [66]. Krylov complexity has been studied in the context of phase transitions [67, 68], as well as in open and driven quantum systems [69, 70, 71, 72]. Another field that has seen widespread use is quantum field theory [73, 74, 75, 76, 77], while [78] analyzes quantum scars using Krylov complexity. An extension to multiple operators was proposed in [79], and a detailed explanation of these methods is provided in [80].

An interesting interpretation of Krylov complexity is that a Krylov operator can be defined such that its expectation value corresponds to Krylov spread and operator complexity. A major challenge in computing spread and operator complexity lies in the requirement of classical simulation to construct Krylov spaces. This work addresses that limitation by introducing quantum-mechanically measurable Krylov spaces for computing spread complexity. This is achieved by demonstrating that time-evolved states and operators can be used in constructing the Krylov space, based on results from [81, 82]. Additionally, Krylov expressivity and Krylov observability are introduced as indicators of the effective space dimension of Krylov spaces.

Finally, this work disproves the widely held assumption that Krylov spaces constructed from Hamiltonian powers always minimize spread complexity, showing through multiple simulations that time-evolved states and operators can exhibit slower growth in Krylov spread and operator complexity.

III.1 Krylov Spaces for Pure States

This section discusses Krylov spaces for pure states. First, a definition of a Krylov space is provided in Def. 4. Theo. 1 demonstrates that the space in which all future states lie consists of different powers of the Hamiltonian H^i applied to the initial state $|\Psi_0\rangle$, i.e.,

$$|\Psi(t)\rangle \in \mathrm{Span}\{|\Psi_0\rangle, H|\Psi_0\rangle, \ldots, H^{m-1}|\Psi_0\rangle\} := \mathrm{K}_m \tag{III.12}$$

for some $m \in \mathbb{N}$ [14, 16]. Based on this, spread complexity (Def. 6) is introduced as a measure to quantify how the time-evolved states $|\Psi(t)\rangle$ spread in K_m.

While this approach has yielded valuable insights into various phenomena in quantum mechanics, it encounters significant limitations in the context of quantum reservoir computing and quantum machine learning. A primary challenge is the requirement of complete knowledge of the Hamiltonian H. In practical settings, especially in experimental implementations, this information is often unavailable or only partially accessible due to noise, decoherence, or limited control over the system. As a result, the applicability of spread complexity remains constrained in real-world scenarios. To overcome this limitation and enhance the interpretability of quantum systems, it is essential to develop a formulation of spread complexity that is experimentally measurable.

To address these challenges, this work demonstrates that instead of constructing the space through powers of the Hamiltonian H^n, an equivalent space can be constructed by measuring the state at discrete times $t_0 < t_1 < t_2 < \ldots < t_{m-1} < T_P$, where t_{m-1} is assumed to be smaller than the period T_P:

$$|\Psi(t)\rangle \in \mathrm{Span}\{e^{-iHt_0}|\Psi_0\rangle, e^{-iHt_1}|\Psi_0\rangle, \ldots, e^{-iHt_{m-1}}|\Psi_0\rangle\} := \mathrm{G}_m. \tag{III.13}$$

Furthermore, it is demonstrated that G_m is independent of the global phase and that the number m equals the number of pairwise distinct eigenvalues of the quantum system. Since G_m is independent of a global phase, it can be measured and used to construct the Krylov space.

Definition 4 (Krylov Space)**.** *Consider a linear function $f : \mathbb{C}^N \to \mathbb{C}^N$ satisfying the following properties for all $\mathbf{v}, \mathbf{w} \in \mathbb{C}^N$ and $\alpha \in \mathbb{C}$:*

$$f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w}), \quad f(\alpha\mathbf{w}) = \alpha f(\mathbf{w}). \tag{III.14}$$

Let $\mathbf{v}_0 \in \mathbb{C}^N \setminus \{\mathbf{0}\}$ *be a non-zero vector. There exists a uniquely defined smallest* $m \leq N \in \mathbb{N}$ *such that the vectors*

$$\mathbf{v}_0, f(\mathbf{v}_0), f^2(\mathbf{v}_0), \ldots, f^{m-1}(\mathbf{v}_0)$$

are linearly independent, while the vectors

$$\mathbf{v}_0, f(\mathbf{v}_0), f^2(\mathbf{v}_0), \ldots, f^{m}(\mathbf{v}_0) \tag{III.15}$$

are linearly dependent.

The integer m *is referred to as the grade of the vector* $\mathbf{v}_0$ *with respect to* f*. The* j*-th Krylov space of* f *with respect to* $\mathbf{v}_0$ *is defined as*

$$\mathrm{K}_j(f, \mathbf{v}_0) = \mathrm{Span}\{\mathbf{v}_0, f(\mathbf{v}_0), \ldots, f^{j-1}(\mathbf{v}_0)\} \subseteq \mathbb{C}^N. \tag{III.16}$$

It follows that $\mathrm{K}_M(f, \mathbf{v}_0) = \mathrm{K}_m(f, \mathbf{v}_0)$ *for all* $M \geq m$*. When the initial state* $\mathbf{v}_0$ *and the function* f *are clear, the Krylov space with grade* m *is denoted by* K_m*.*

Definition 5 (State Evolution)**.** *Consider a quantum system described by the Hamiltonian* $H \in \mathbb{C}^{N\times N}$*, where* $H = H^\dagger$*. Let* $|\Psi_0\rangle, |\Psi(t)\rangle \in \mathbb{C}^N$ *represent the initial state and the state at time* $t \in \mathbb{R}_{\geq 0}$*, respectively. Given the initial condition* $|\Psi(0)\rangle = |\Psi_0\rangle$*, the state at time* t *is determined by*

$$|\Psi(t)\rangle = e^{-iHt} |\Psi_0\rangle . \tag{III.17}$$

To gain insight into the time evolution of quantum states, the results will be discussed for a transverse field Ising model with six sites. The inter-spin couplings are sampled from a uniform distribution $J_{i,j} \in \mathcal{U}([0.25, 0.5])$, resulting in

$$H_\mathrm{S} = \sum_{i=1,j>i}^{6} J_{i,j} X_{i-1} X_i + \frac{1}{2} \sum_{i=0}^{6} Z_i. \tag{III.18}$$

III.1.1 Spread Complexity through the Krylov spaces of the Hamiltonian

Theorem 1 (Time Evolution as a Map onto a Krylov Space)**.** *Using the notation introduced in Def. 5, let* $t \in \mathbb{R}$*, and let* $|\Psi(t)\rangle = \exp(-iHt)\,|\Psi_0\rangle$ *represent the time-evolved state at time* t*, where* $|\Psi_0\rangle \neq 0$ *is the initial state. For all* $t \in \mathbb{R}$*, the following holds:*

$$\begin{aligned} |\Psi(t)\rangle \in \mathrm{K}_m(-iH, |\Psi_0\rangle) :&= \mathrm{Span}\{f^0(|\Psi_0\rangle), \dots, f^{m-1}(|\Psi_0\rangle)\} \\ &= \mathrm{Span}\{H^0\,|\Psi_0\rangle, \dots, H^{m-1}\,|\Psi_0\rangle\}, \end{aligned} \tag{III.19}$$

where $m \in \mathbb{N}$ *is the grade of the initial state* $|\Psi_0\rangle$ *with respect to the linear function* $f := -iH$ *[14, 16].*

Proof. The time-evolved state is defined as $|\Psi(t)\rangle = \exp(-iHt)\,|\Psi_0\rangle$. Introducing the linear function $f : \mathbb{C}^N \to \mathbb{C}^N$ as $f(|\Psi\rangle) := -iH\,|\Psi\rangle$, the Taylor expansion of the time evolution can be expressed as

$$\begin{aligned} |\Psi(t)\rangle &= e^{-iHt}\,|\Psi_0\rangle \\ &= \sum_{k=0}^{\infty}(-iH)^k \frac{t^k}{k!}\,|\Psi_0\rangle = \sum_{k=0}^{\infty} f^k(|\Psi_0\rangle)\frac{t^k}{k!}. \end{aligned} \tag{III.20}$$

This implies that $|\Psi(t)\rangle$ is a superposition of vectors $f^k(|\Psi_0\rangle)$ with coefficients $\frac{t^k}{k!}$. Therefore, the state $|\Psi(t)\rangle$ satisfies:

$$|\Psi(t)\rangle \in \mathrm{Span}\{f^0(|\Psi_0\rangle), f^1(|\Psi_0\rangle)t, f^2(|\Psi_0\rangle)\frac{t^2}{2!}, \dots\}. \tag{III.21}$$

For $t = 0$, $|\Psi(t)\rangle = |\Psi_0\rangle$. For $t > 0$, the coefficients $\frac{t^k}{k!} > 0$, so the span becomes:

$$\mathrm{Span}(f, |\Psi_0\rangle) = \mathrm{Span}\{f^0(|\Psi_0\rangle), f^1(|\Psi_0\rangle), f^2(|\Psi_0\rangle), \dots\}. \tag{III.22}$$

By the linearity of f and the properties of the Krylov space in Def. 4, there exists an $m \leq N$ such that the following vectors are linearly independent:

$$|\Psi_0\rangle\,, f(|\Psi_0\rangle), f^2(|\Psi_0\rangle), \dots, f^{m-1}(|\Psi_0\rangle), \tag{III.23}$$

while the set

$$|\Psi_0\rangle\,, f(|\Psi_0\rangle), f^2(|\Psi_0\rangle), \dots, f^{m}(|\Psi_0\rangle), \tag{III.24}$$

is linearly dependent. Thus, it follows that:

$$\begin{aligned}\mathrm{Span}(f, |\Psi_0\rangle) &= \mathrm{Span}\{f^0(|\Psi_0\rangle), f^1(|\Psi_0\rangle)t, f^2(|\Psi_0\rangle)\frac{t^2}{2!}, \dots\} \\ &= \mathrm{Span}\{f^0(|\Psi_0\rangle), f^1(|\Psi_0\rangle), f^2(|\Psi_0\rangle), \dots\} \\ &= \mathrm{Span}\{f^0(|\Psi_0\rangle), f^1(|\Psi_0\rangle), \dots, f^{m-1}(|\Psi_0\rangle)\} \\ &= \mathrm{K}_m(-iH, |\Psi_0\rangle), \end{aligned} \tag{III.25}$$

which confirms that $|\Psi(t)\rangle \in \mathrm{K}_m(-iH, |\Psi_0\rangle)$, where m is the grade of $|\Psi_0\rangle$ with respect to $f := -iH$[14, 16]. □

Definition 6 (Spread Complexity)**.** *Assume the quantum system is described by the Hamiltonian H with initial state $|\Psi_0\rangle$ and the corresponding Krylov space*

$$\mathrm{K}_m = \mathrm{Span}\{f^0(|\Psi_0\rangle), f^1(|\Psi_0\rangle), \dots, f^{m-1}(|\Psi_0\rangle)\}.$$

The space K_m is then orthonormalized, resulting in

$$\mathrm{K}_m = \mathrm{Span}\{|k_0\rangle, |k_1\rangle, \dots, |k_{m-1}\rangle\}.$$

Any time-evolved state can be represented in this basis via

$$|\Psi(t)\rangle = \sum_{n=0}^{m-1} |k_n\rangle \langle k_n|\Psi(t)\rangle . \tag{III.26}$$

With $\alpha_n := \langle k_n|\Psi(t)\rangle$, the spread complexity $\mathcal{K}_S$ is defined as

$$\mathcal{K}_S(t) = \sum_{n=0}^{m-1} (n+1)|\alpha_n(t)|^2. \tag{III.27}$$

One interpretation of spread complexity arises from defining the **Krylov operator**

$$\mathrm{K}_S := \sum_{n=0}^{m-1} (n+1)\, |k_n\rangle \langle k_n| , \tag{III.28}$$

which leads to the computation of its expectation value as

$$\begin{aligned}\langle \mathrm{K}_S\rangle = \langle\Psi(t)|\, \mathrm{K}_S\, |\Psi(t)\rangle &= \sum_{n=0}^{m-1} (n+1)\, \langle\Psi(t)|k_n\rangle \langle k_n|\Psi(t)\rangle \\ = \sum_{n=0}^{m-1} (n+1)\alpha_n\alpha_n^* &= \sum_{n=0}^{m-1} (n+1)|\alpha_n(t)|^2 = \mathcal{K}_S(t). \end{aligned} \tag{III.29}$$

This shows that $\mathcal{K}_S$ can be interpreted as the expectation value of the Krylov operator K_S.

The definition of spread complexity can be generalized to arbitrary spaces. Let

$$\mathrm{U}_m = \mathrm{Span}\{|b_0\rangle, |b_1\rangle, \dots, |b_{m-1}\rangle\} = \mathrm{Span}(\mathcal{B})$$

be a space such that $|\Psi(t)\rangle \in \mathrm{U}_m$ for all $t \in \mathbb{R}$. Spread complexity with respect to this basis is written as

$$\mathcal{K}_S(\mathcal{B}, t) = \sum_{n=0}^{m-1} (n+1) |\langle b_n|\Psi(t)\rangle|^2. \tag{III.30}$$

In the paper [16], the authors claim that the Krylov basis $\mathcal{B}_K = \{|k_i\rangle\}_{i=0,\dots,m-1}$ is the basis that minimizes Krylov complexity $\mathcal{K}_S$, i.e.,

$$\mathcal{K}_S(\mathcal{B}_K, t) \leq \mathcal{K}_S(\mathcal{B}, t) \tag{III.31}$$

for all bases $\mathcal{B}$.

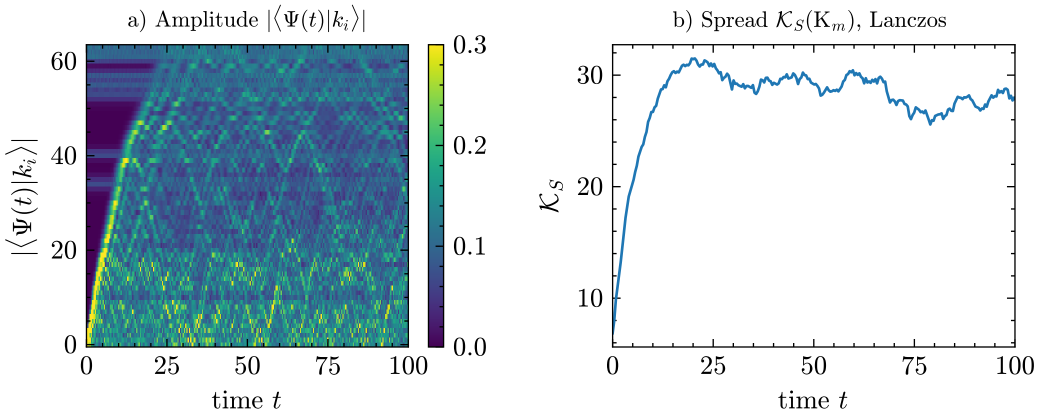


Figure III.1: Amplitudes $a_i = |\langle\Psi(t)|k_i\rangle|$, where $|k_i\rangle$ corresponds to the i-th Krylov subspace vector **a)** and the spread complexity $\mathcal{K}_S$ as a function of time in **b)**. In both plots the Lanczos algorithm is used for orthonormalization. The spread complexity increases and then saturates, similar to the behavior of the maximum amplitude $\max(a_i)_i$.

Figure III.1 shows the basis representation and spread complexity as a function of time t. The index with the maximum amplitude α_i increases over time until the highest value is reached. As t grows, the spread over the basis also increases. This is observed around $t \approx 25$, after which any time-evolved state consists of a superposition of all previous states. When the basis representation reaches maximal scrambling, spread complexity $\mathcal{K}_S$ reaches its maximum at approximately the same time (Figure III.1.b).

Before discussing this further, note the top-left corner of the amplitude plot in Figure III.1.a. This region shows that even at small times t, large amplitudes α_i exist beyond the maximum range (e.g., $\alpha_{30}, \dots, \alpha_{64}$). This is a numerical artifact caused by the unstable structure of the Lanczos algorithm, failing in representing time-evolved states which is vivid by the large reconstruction error $\delta \approx 0.47$, where δ is given by

$$\delta = \left| |\Psi(t)\rangle - \sum_{i=0}^{m-1} \langle k_i|\Psi(t)\rangle \, |k_i\rangle \right| . \tag{III.32}$$

This instability arises from the vectors not being entirely orthogonal in the reconstruction. To resolve this issue, the Arnoldi algorithm is used, and the results are shown in Figure III.2. This algorithm is more stable, as evidenced by a reconstruction error of approximately $\delta \approx 3 \times 10^{-5}$. The artifacts in the top-left corner are no longer visible, and the spread complexity $\mathcal{K}_S$ increases more slowly but reaches the same maximum.

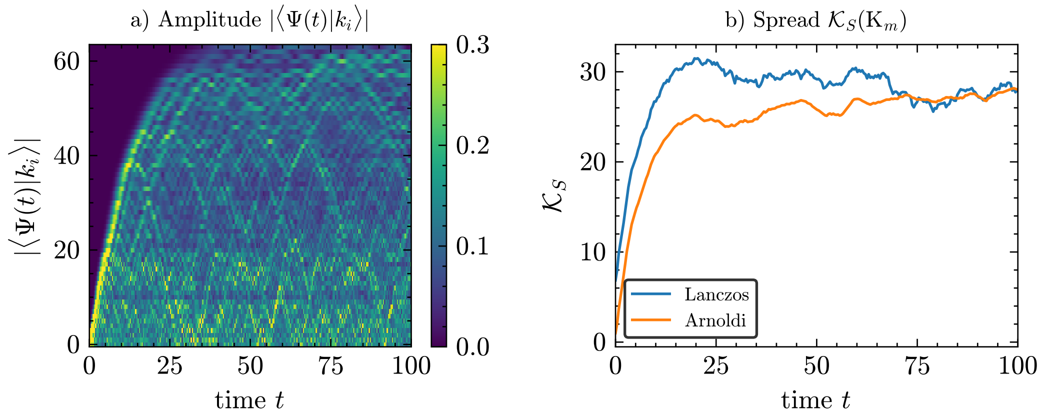


Figure III.2: Amplitudes $|\langle\Psi(t)|k_i\rangle|$, where $|k_i\rangle$ corresponds to the i-th Krylov subspace vector **a)** and the spread complexity $\mathcal{K}_S$ as a function of time in **b)** computed using the Arnoldi algorithm. The spread complexity increases and then saturates, similar to the behavior of the maximum amplitude $\max(\alpha_i)_i$.

The Krylov spaces and spread complexity, constructed using two distinct orthonormalization algorithms, have been discussed in an example, demonstrating that the Lanczos algorithm faces certain challenges. However, no statement has yet been made regarding the minimality of the number of basis vectors in the construction of K_m. The question is, whether there exists a basis $\mathcal{B}$ with $\dim(\mathcal{B}) < m$ such that $|\Psi(t)\rangle \in \mathrm{Span}\{\mathcal{B}\}$. The following lemma establishes that this is not the case.

Lemma 1 (K_m consists of the minimum number of basis states.)**.** *Assume* $|\Psi(t)\rangle$ *is a time-evolved state under Hamiltonian* H*, and assume that the corresponding Krylov space is given by* K_m*, such that* $|\Psi(t)\rangle \in \mathrm{K}_m$ *for all* $t \in \mathbb{R}$*. Then it holds that there exists no basis* $\mathcal{B}$ *with* $\dim(\mathcal{B}) < m$ *such that* $|\Psi(t)\rangle \in \mathrm{Span}\{\mathcal{B}\}$*.*

Proof. The proof proceeds as follows. Take $L > m$ time-evolved states $|\Psi(t_j)\rangle = e^{-iHt_j}|\Psi_0\rangle$ at times $0 = t_0 < t_1 < t_2 < \ldots < t_L < T_P$, where L is very large, such that

$$|\Psi(t_j)\rangle = \sum_{k=0}^{\infty}(-iH)^k \frac{t_j^k}{k!}|\Psi_0\rangle = \sum_{k=0}^{L-1}(-iH)^k \frac{t_j^k}{k!}|\Psi_0\rangle + \varepsilon, \tag{III.33}$$

where ε is negligible. Introduce the substitution:

$$\begin{aligned} h_j^{(L)}(|\Psi_0\rangle) &:= \sum_{k=0}^{L-1} f^k(|\Psi_0\rangle)\frac{t_j^k}{k!}, \\ \text{with } f^k(|\Psi_0\rangle) &:= (-iH)^k|\Psi_0\rangle . \end{aligned} \tag{III.34}$$

For simplicity $|\Psi_0\rangle$ is ignored in the following calculations. The vectors $h_j^{(L)}$ can be written as

$$h_i^{(L)} = \sum_{j=0}^{n-1} f^j \frac{t_i^j}{j!} = \begin{pmatrix} f^0 & f^1 & \ldots & f^{n-1}\end{pmatrix}\begin{pmatrix} 1 \\ t_i/1! \\ t_i^2/2! \\ \vdots \\ t_i^{n-1}/(n-1)! \end{pmatrix}. \tag{III.35}$$

Writing all vectors $h_0^{(L)}, \ldots, h_{n-1}^{(L)}$ yields

$$\begin{aligned} &\begin{pmatrix} h_0^{(L)} & h_1^{(L)} & \ldots & h_{n-1}^{(L)} \end{pmatrix} = \begin{pmatrix} f^0 & f^1 & \ldots & f^{n-1}\end{pmatrix}\Theta, \\ &\Theta = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ t_1/1! & t_2/1! & \ldots & t_{n-1}/1! \\ t_1^2/2! & t_2^2/2! & \ldots & t_{n-1}^2/2! \\ \vdots & \vdots & \ddots & \vdots \\ t_1^{n-1}/(n-1)! & t_2^{n-1}/(n-1)! & \ldots & t_{n-1}^{n-1}/(n-1)! \end{pmatrix}. \end{aligned} \tag{III.36}$$

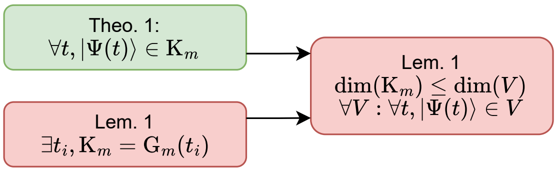

Theo. 1:
$\forall t, |\Psi(t)\rangle \in \mathrm{K}_m$
Lem. 1
$\exists t_i, \mathrm{K}_m = \mathrm{G}_m(t_i)$
Lem. 1
$\dim(\mathrm{K}_m) \leq \dim(V)$
$\forall V : \forall t, |\Psi(t)\rangle \in V$

Figure III.3: Sketch showing the properties of K_m. This shows that any time-evolved state lies in K_m, and using Lemma 1, which concludes that the number of basis states in K_m is minimal.

Since all times t_i are pairwise distinct by definition, the columns of Θ are linearly independent. Therefore, Θ is invertible, and its inverse Θ^{-1} exists:

$$\left(h_0^{(L)} \quad h_1^{(L)} \quad \dots \quad h_{n-1}^{(L)}\right)\Theta^{-1} = \left(f^0 \quad f^1 \quad \dots \quad f^{n-1}\right). \tag{III.37}$$

Since all vectors $\{f^0, f^1, \dots, f^{L-1}\}$ can be represented as linear combinations of vectors $\{h_0^{(L)}, h_1^{(L)}, \dots, h_{L-1}^{(L)}\}$, the spans of both sides are equal:

$$H_L^L = \mathrm{Span}\{h_0^{(L)}, h_1^{(L)}, \dots, h_{n-1}^{(L)}\} = \mathrm{Span}\{f^0, f^1, \dots, f^{L-1}\} = K_L. \tag{III.38}$$

Since K_L contains only m linearly independent vectors, H_L^L must also contain only m linearly independent vectors; that is, $H_m^L = K_m$. This implies that there cannot exist a basis $\mathcal{B}$ with $\dim(\mathcal{B}) < m$ such that $|\Psi(t)\rangle \in \mathrm{Span}\{\mathcal{B}\}$. If such a basis existed, then the L vectors $h_j^{(L)}$ would not be able to represent the vectors f^j, and the matrix Θ would not be invertible—contradicting the fact that Θ is indeed invertible. □

Figure III.3 shows a sketch illustrating Theo. 1 and Lemma 1, and how they interact.

III.1.2 Spread Complexity via Measurable Krylov Spaces

In this subsection, the connection between Krylov spaces and the eigenstates of the Hamiltonian is discussed. Afterwards, measurable Krylov spaces are proposed and explored in more detail.

The eigenvalue equation of the Hamiltonian is given by

$$H|\phi_n\rangle = \varepsilon_n |\phi_n\rangle. \tag{III.39}$$

For a Hermitian Hamiltonian, i.e., $H^\dagger = H$, it follows that all eigenstates $|\phi_n\rangle$ are orthogonal to each other. The spectral decomposition is thus given by

$$H = \sum_n \varepsilon_n |\phi_n\rangle \langle\phi_n| \,. \tag{III.40}$$

Inserting this into the time evolution and assuming $H \in \mathbb{C}^{N\times N}$ results in

$$|\Psi(t)\rangle = e^{-iHt} |\Psi_0\rangle = \sum_{n=0}^{N-1} e^{-i\varepsilon_n t} |\phi_n\rangle \langle\phi_n|\Psi_0\rangle \,. \tag{III.41}$$

It is apparent that $|\Psi(t)\rangle \in \mathrm{E}_N = \mathrm{Span}\{|\phi_0\rangle, \dots, |\phi_{N-1}\rangle\}$ for all $t \in \mathbb{R}$. The following theorem makes a connection between the number of pairwise distinct eigenvalues and the Krylov rank.

Theorem 2. *Given a Hamiltonian $H \in \mathbb{C}^{N\times N}$ with d pairwise distinct eigenvalues, satisfying the eigenvalue equation*

$$H |\phi_j\rangle = \varepsilon_j |\phi_j\rangle \,, \tag{III.42}$$

it follows that the time-evolved state $|\Psi(t)\rangle \in \mathrm{E}_d$, i.e.,

$$|\Psi(t)\rangle \in \mathrm{E}_d = \mathrm{Span}\{|\xi_0\rangle, |\xi_1\rangle, \dots, |\xi_{d-1}\rangle\}. \tag{III.43}$$

With $\langle\phi_j|\Psi_0\rangle = \alpha_j$, the vectors $|\xi_p\rangle$ are defined as

$$|\xi_p\rangle = \frac{1}{\sqrt{|J_p|}} \sum_{j\in J_p} \alpha_j |\phi_j\rangle \,, \tag{III.44}$$

where J_p is an index set such that all eigenstates $|\phi_j\rangle$ with $j \in J_p$ share the same eigenvalue ε_p, and $|J_p|$ is the number of indices in the set, used for normalization.

Proof. With $\alpha_j := \langle\phi_j|\Psi_0\rangle$, it follows:

$$|\Psi(t)\rangle = e^{-iHt} |\Psi_0\rangle = \sum_j e^{-i\varepsilon_j t} |\phi_j\rangle \langle\phi_j|\Psi_0\rangle = \sum_j e^{-i\varepsilon_j t} \alpha_j |\phi_j\rangle \,. \tag{III.45}$$

Given that $H \in \mathbb{C}^{N\times N}$ and is Hermitian, it possesses N eigenvalues, of which only d are distinct. Let these d distinct eigenvalues be $\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{d-1}$. Since H is Hermitian, its eigenstates $|\phi_j\rangle$ form an orthogonal basis of $\mathbb{C}^N$.

Suppose j_i eigenvectors correspond to the eigenvalue ε_i. Define the index sets $J_0 = \{0, \dots, j_0\}$ and $J_k = \{j_{k-1}+1, \dots, j_k\}$ for $k \in \{1, \dots, d-1\}$. These index sets categorize the eigenstates $|\phi_j\rangle$ such that all $|\phi_a\rangle$ with $a \in J_k$ share the eigenvalue ε_k.

The time evolution can be rewritten as:

$$|\Psi(t)\rangle = \sum_j e^{-i\varepsilon_j t}\alpha_j |\phi_j\rangle = \sum_{p=0}^{d-1}\sum_{j\in J_p} e^{-i\varepsilon_p t}\alpha_j |\phi_j\rangle = \sum_{p=0}^{d-1} e^{-i\varepsilon_p t}\sum_{j\in J_p}\alpha_j |\phi_j\rangle . \tag{III.46}$$

The inner sum over J_p is time-independent and can be represented by the vectors

$$|\xi_p\rangle = \sum_{j\in J_p}\alpha_j |\phi_j\rangle . \tag{III.47}$$

Normalizing the vectors via $|\xi_p\rangle \leftarrow \frac{|\xi_p\rangle}{\sqrt{|J_p|}}$, it follows that for all $t \in \mathbb{R}$,

$$|\Psi(t)\rangle \in \operatorname{Span}(|\xi_0\rangle, |\xi_1\rangle, \ldots, |\xi_{d-1}\rangle) =: \mathrm{E}_d. \tag{III.48}$$

Hence, the space E_d consists of a superposition of d linearly independent vectors $|\xi_0\rangle, \ldots, |\xi_{d-1}\rangle$. □

Lemma 2 (E_m consists of the minimum number of basis states.)**.** *Given an initial state* $|\Psi_0\rangle$*, a Hamiltonian* H*, and the corresponding space of eigenstates* E_d*, such that* $|\Psi(t)\rangle \in \mathrm{K}_m$ *for all* $t \in \mathbb{R}$*, there exists no basis* $\mathcal{B}$ *with* $\dim(\mathcal{B}) < d$ *such that any time-evolved state is in the span of* $\mathcal{B}$*.*

Proof. Assume times $t_0 < t_1 < \ldots < t_{d-1}$ and the states evolved at those times are given by

$$\begin{aligned}|\Psi(t_j)\rangle &= e^{-iHt_j}|\Psi_0\rangle = \sum_{k=0}^{N} e^{-i\varepsilon_k t_j}|\phi_k\rangle\langle\phi_k|\Psi_0\rangle \\ &= \sum_{p=0}^{d-1} e^{-i\varepsilon_p t_j}\sum_{j\in J_p}\alpha_j|\phi_j\rangle = \sum_{p=0}^{d-1} e^{-i\varepsilon_p t_j}|\xi_p\rangle \\ &= \begin{pmatrix}|\xi_0\rangle & |\xi_1\rangle & \ldots & |\xi_{d-1}\rangle\end{pmatrix}\begin{pmatrix} e^{-i\varepsilon_0 t_j} \\ e^{-i\varepsilon_1 t_j} \\ \vdots \\ e^{-i\varepsilon_{d-1} t_j}\end{pmatrix}\end{aligned} \tag{III.49}$$

Here, $|\xi_p\rangle = \sum_{j\in J_p} \alpha_j |\phi_j\rangle$ and $\alpha_j = \langle\phi_j|\Psi_0\rangle$ is used. Writing the d states in terms of the $|\xi_p\rangle$ basis results in

$$\begin{gathered}\Big(|\Psi(t_0)\rangle \quad |\Psi(t_1)\rangle \quad \dots \quad |\Psi(t_{d-1})\rangle\Big) = \Big(|\xi_0\rangle \quad |\xi_1\rangle \quad \dots \quad |\xi_{d-1}\rangle\Big)\,\Sigma \\ \Sigma = \begin{pmatrix} e^{-i\varepsilon_0 t_0} & e^{-i\varepsilon_0 t_1} & \dots & e^{-i\varepsilon_0 t_{d-1}} \\ e^{-i\varepsilon_1 t_0} & e^{-i\varepsilon_1 t_1} & \dots & e^{-i\varepsilon_1 t_{d-1}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-i\varepsilon_{d-1} t_0} & e^{-i\varepsilon_{d-1} t_1} & \dots & e^{-i\varepsilon_{d-1} t_{d-1}} \end{pmatrix}\end{gathered} \tag{III.50}$$

If the matrix Σ is invertible, then it follows that any d time-evolved states span E_d. If $|\varepsilon_j| < \pi$, we can write $x_j^t = (e^{-i\varepsilon_j})^t$. If this is not the case, the Hamiltonian can be rescaled as $H \leftarrow H/|\varepsilon_{\max}|$ and time as $t \leftarrow t \cdot |\varepsilon_{\max}|$, which results in the same dynamics. With this, the matrix in Eq. (III.50) can be rewritten as

$$\Big(|\Psi(t_0)\rangle \quad |\Psi(t_1)\rangle \quad \dots \quad |\Psi(t_{d-1})\rangle\Big) = \Big(|\xi_0\rangle \quad |\xi_1\rangle \quad \dots \quad |\xi_{d-1}\rangle\Big)\,\Sigma \tag{III.51}$$

where

$$\Sigma = \begin{pmatrix} x_0^{t_0} & x_0^{t_1} & \dots & x_0^{t_{d-1}} \\ x_1^{t_0} & x_1^{t_1} & \dots & x_1^{t_{d-1}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{d-1}^{t_0} & x_{d-1}^{t_1} & \dots & x_{d-1}^{t_{d-1}} \end{pmatrix} \tag{III.52}$$

For pairwise distinct times, the columns are linearly independent if $t_{d-1} < T_P$, where T_P is the period of the system. Σ is a generalized Vandermonde matrix, for which an inverse exists. Therefore,

$$\Big(|\Psi(t_0)\rangle \quad |\Psi(t_1)\rangle \quad \dots \quad |\Psi(t_{d-1})\rangle\Big)\,\Sigma^{-1} = \Big(|\xi_0\rangle \quad |\xi_1\rangle \quad \dots \quad |\xi_{d-1}\rangle\Big) \tag{III.53}$$

Since all time-evolved states lie in E_d, and since these d time-evolved states span the space, it follows that E_d consists of the minimum number of basis vectors. ☐

Theorem 3. *Given the previous definitions and notation, it holds that* $\mathrm{E}_m = \mathrm{K}_m$ *and* $m = d$.

Proof. Lemma 1 showed that m and d are the smallest possible dimensions such that all future states lie in $|\Psi(t)\rangle \in \mathrm{K}_m$ (Theo. 1) and $|\Psi(t)\rangle \in \mathrm{E}_d$ (Theo. 2). This implies that $m = d$, and thus $\mathrm{E}_m = \mathrm{K}_m$. ☐

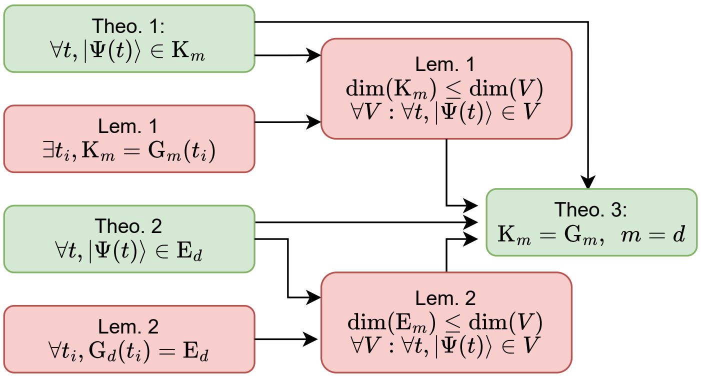

Theo. 1:
$\forall t, |\Psi(t)\rangle \in \mathrm{K}_m$
Lem. 1
$\exists t_i, \mathrm{K}_m = \mathrm{G}_m(t_i)$
Theo. 2
$\forall t, |\Psi(t)\rangle \in \mathrm{E}_d$
Lem. 2
$\forall t_i, \mathrm{G}_d(t_i) = \mathrm{E}_d$
Lem. 1
$\dim(\mathrm{K}_m) \leq \dim(V)$
$\forall V : \forall t, |\Psi(t)\rangle \in V$
Theo. 3:
$\mathrm{K}_m = \mathrm{G}_m, \quad m = d$
Lem. 2
$\dim(\mathrm{E}_m) \leq \dim(V)$
$\forall V : \forall t, |\Psi(t)\rangle \in V$

Figure III.4: The sketch is extended to show that E_d also consists of the minimum number of basis states. Theo. 3 concludes that $m = d$ and that $\mathrm{K}_m = \mathrm{E}_m$ holds.

Figure III.4 shows how the proven theorems and lemmas support this conclusion.

Theorem 4 (Measurable Krylov Spaces)**.** *Let us assume that our quantum system is described by a Hamiltonian* $H \in \mathbb{C}^{N\times N}$ *and that the initial state* $|\Psi_0\rangle$ *is given. Let* $\mathrm{K}_m = \mathrm{Span}\{H^0 |\Psi_0\rangle, \ldots, H^{m-1} |\Psi_0\rangle\}$ *be the corresponding Krylov space. Then it holds that*

$$\begin{aligned} \mathrm{K}_m &= \mathrm{Span}\{H^0 |\Psi_0\rangle, \ldots, H^{m-1} |\Psi_0\rangle\} \\ &= \mathrm{Span}\{e^{-iHt_0} |\Psi_0\rangle, \ldots, e^{-iHt_{m-1}} |\Psi_0\rangle\} =: \mathrm{G}_m, \end{aligned} \tag{III.54}$$

where $0 = t_0 < t_1 < \ldots < t_{m-1}$ *are times smaller than the period of the system, i.e.,* $t_{m-1} < T_P$*. From this, it follows that* $\mathrm{K}_m = \mathrm{E}_m = \mathrm{G}_m$*.*

Proof. This follows directly from Lemma 2, where it was shown that d time-evolved states can be used to construct the space of eigenstates, i.e., $\mathrm{G}_d = \mathrm{E}_d$. In Theo. 3, it was shown that $d = m$ and $\mathrm{K}_m = \mathrm{E}_m$. Thus, it follows that

$$\mathrm{K}_m = \mathrm{E}_m = \mathrm{G}_m, \tag{III.55}$$

and a visualization is given in Figure III.5. □

Lemma 3. *For equidistant times* $t_n = nT_G/m$*, the space* G_m *becomes a Krylov space and consists of* m *linearly independent basis states. While this result is already implied by Theo. 4, we provide an alternative proof here due to its significance for later sections of this work and to support the reader's understanding.*

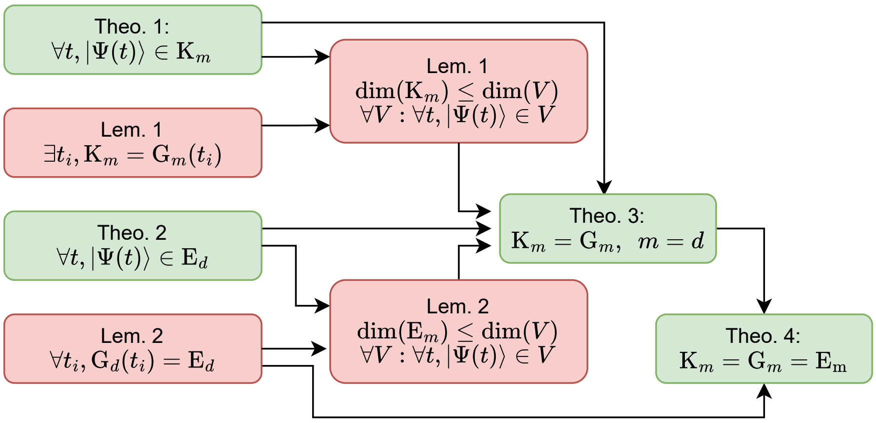

Theo. 1:
$\forall t, |\Psi(t)\rangle \in \mathrm{K}_m$
Lem. 1
$\exists t_i, \mathrm{K}_m = \mathrm{G}_m(t_i)$
Theo. 2
$\forall t, |\Psi(t)\rangle \in \mathrm{E}_d$
Lem. 2
$\forall t_i, \mathrm{G}_d(t_i) = \mathrm{E}_d$
Lem. 1
$\dim(\mathrm{K}_m) \leq \dim(V)$
$\forall V : \forall t, |\Psi(t)\rangle \in V$
Lem. 2
$\dim(\mathrm{E}_m) \leq \dim(V)$
$\forall V : \forall t, |\Psi(t)\rangle \in V$
Theo. 3:
$\mathrm{K}_m = \mathrm{G}_m, \quad m = d$
Theo. 4:
$\mathrm{K}_m = \mathrm{G}_m = \mathrm{E}_\mathrm{m}$

Figure III.5: The sketch is extended to show that E_d also consists of the smallest number of basis states. Theo. 3 concludes that $m = d$ and that $\mathrm{K}_m = \mathrm{E}_m$. From the previous theorems and lemmas, Theo. 4 concludes that the three spaces—constructed from powers of the Hamiltonian (K_m), eigenvectors (E_m), and time-evolved states (G_m)—are equivalent.

Proof. For equidistant times $t_n = nT_G/m$ with $n = 0, \ldots, m-1$, define the time evolution operator as $U = e^{-iHT_G/m}$. The space G_m can then be written as:

$$\mathrm{G}_m = \mathrm{Span}\{|\Psi_0\rangle, U|\Psi_0\rangle, U^2|\Psi_0\rangle, \ldots, U^{m-1}|\Psi_0\rangle\}. \tag{III.56}$$

If m is the grade such that

$$U^k|\Psi_0\rangle \in \mathrm{G}_m \tag{III.57}$$

for all $k \in \mathbb{N}$, then in the limit of small $\Delta t = t_n - t_{n-1} \to 0$, any time-evolved state $|\Psi(t)\rangle$ can be approximated as $|\Psi(t)\rangle = U^k|\Psi_0\rangle + \varepsilon$ for some k and a small error term $\varepsilon \to 0$ as $\Delta t \to 0$.

However, from earlier results, we know that all time-evolved states lie in the Krylov space

$$|\Psi(t)\rangle \in \mathrm{K}_m = \mathrm{Span}\{|\Psi_0\rangle, H|\Psi_0\rangle, \ldots, H^{m-1}|\Psi_0\rangle\}. \tag{III.58}$$

This implies that $\mathrm{G}_m = \mathrm{K}_m$, i.e., the space generated by time evolution at equidistant intervals is itself a Krylov space. □

It has been shown that as long as pairwise distinct times $t_0 < t_1 < \ldots < t_{m-1}$ are used, the space can be equivalently reconstructed through time-evolved states $|\Psi(t_i)\rangle$.

The times are arbitrary, so it makes sense to pick a total time T_G and then measure time-evolved states at equidistant points $t_i = iT_G/m$. This results in the space

$$\mathrm{G}_m = \mathrm{Span}\{\Psi(0), \Psi(T_G/m), \ldots, \Psi((m-1)T_G/m)\}. \tag{III.59}$$

Orthonormalization of this space yields

$$\mathrm{G}_m = \mathrm{Span}\{|g_0\rangle, |g_1\rangle, \ldots, |g_{m-1}\rangle\}. \tag{III.60}$$

Spread complexity can be computed in regards to the space G_m, which gives

$$\mathcal{K}_S(\mathrm{G}_m, t) = \sum_{n=0}^{m-1} (n+1) |\langle g_n|\Psi(t)\rangle|^2. \tag{III.61}$$

This is equivalent to the expectation value of the operator

$$\mathrm{G}_S := \sum_{n=0}^{m-1} (n+1) |g_n\rangle \langle g_n| , \tag{III.62}$$

i.e., $\mathcal{K}_S(\mathrm{G}_m, t) = \langle \mathrm{G}_S \rangle$.

In the following study, the basis representation and spread complexity for the spaces G_m will be shown. Equidistant state sampling is used with $t_n = nT_G/m$ and $T_G = 10,\ 40,\ 60$ for visualization. We show that the Krylov basis does not minimize spread complexity, contrary to the claim in [16].

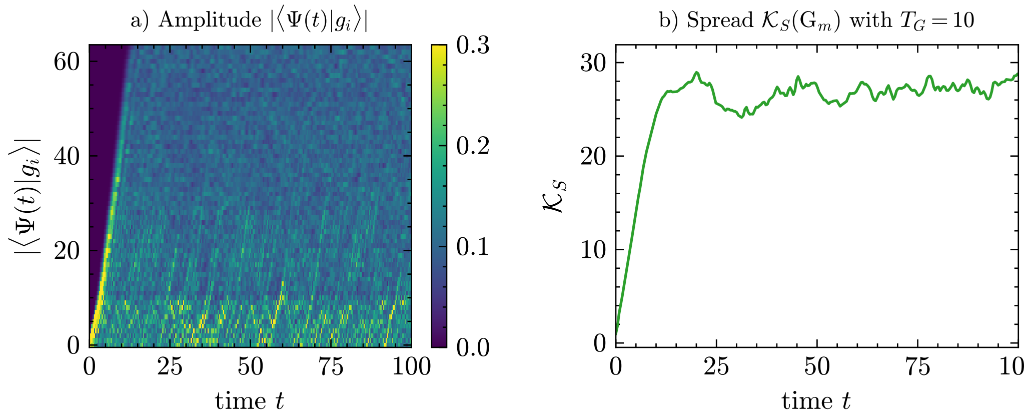


Figure III.6: Amplitudes $|\langle\Psi(t)|g_i\rangle|$, where $|g_i\rangle$ is the i-th basis state of G_m with $T_G = 10$ **a)**, and the spread complexity $\mathcal{K}_S$ as a function of time in **b)**, computed using the measurable Krylov basis. The spread complexity increases and then saturates.

Changing the total time in the construction of G_m to $T_G = 40$ gives Figure III.7. This behavior is somewhat similar to the Krylov space K_m in Figure III.2. The maximum is reached around $t \approx 75$. The basis representation is nearly linear up to $t = 40$, after which the state spreads across the whole basis. The slope of the maximum amplitude index change decreases with time.

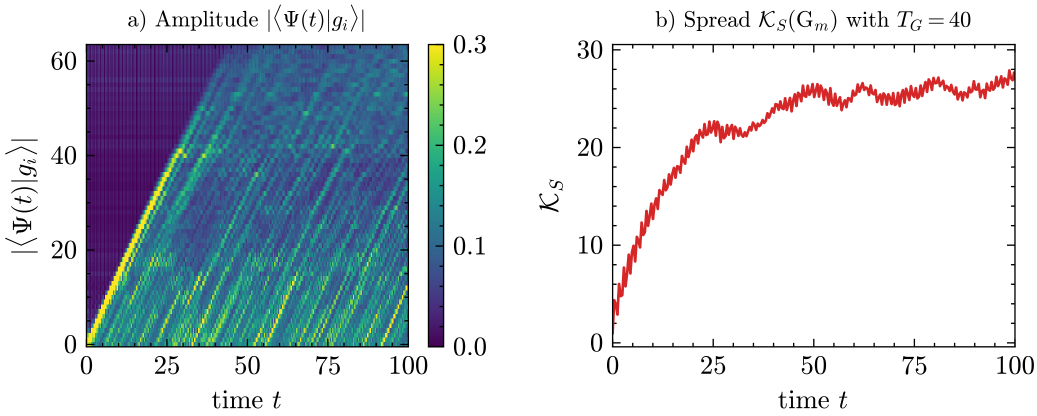


Figure III.7: Amplitudes $|\langle\Psi(t)|g_i\rangle|$, where $|g_i\rangle$ is the i-th basis state of G_m with $T_G = 40$ **a)**, and the spread complexity $\mathcal{K}_S$ as a function of time in **b)**, computed using the measurable Krylov basis. The spread complexity increases and then saturates.

The results for $T_G = 60$ are shown in Figure III.8. Here, a smaller slope is observed in the basis representation, with more distinct parallel lines. The spread complexity increases more slowly and reaches $\mathcal{K}_S = 27$ at $t = 100$, which was the saturation point for previous cases. Oscillations in the spread complexity are due to oscillations in the diagonal lines of the amplitude plot.

Figure III.9 compares spread complexity $\mathcal{K}_S$ computed via different powers of the Hamiltonian (using Lanczos and Arnoldi algorithms) in (a), and via the time-evolved states in (b). The spread complexity from Arnoldi is also shown in (b) for reference. It is observed that $\mathcal{K}_S$ using G_m with $T_G = 10$ (green) is larger than that from K_m. For $T_G = 40$ (red) and $T_G = 60$ (grey), the spread complexity is smaller and grows more slowly. The spread with $T_G = 60$ shows stronger oscillations and an overall smaller spread.

These findings contradict the claim in [16] that the Krylov basis minimizes spread complexity. While this work does not propose an alternative optimal basis, it disproves a foundational assumption in the field.

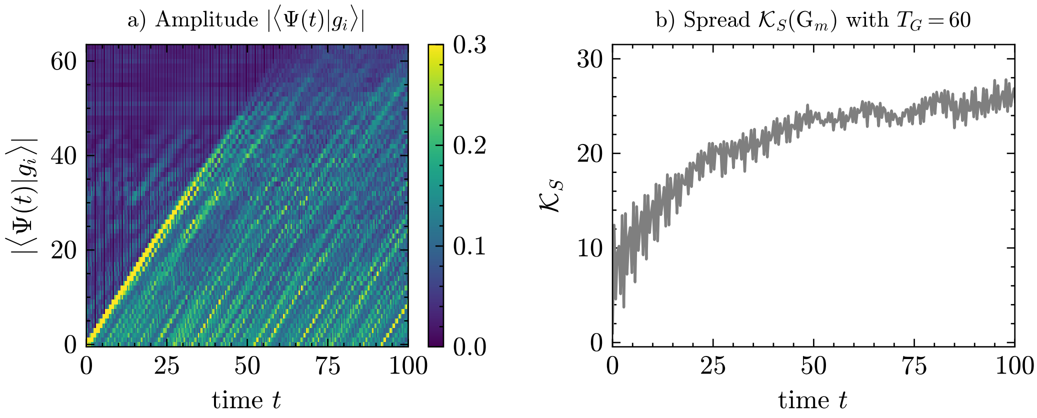


Figure III.8: Amplitudes $|\langle\Psi(t)|g_i\rangle|$, where $|g_i\rangle$ is the i-th basis state of G_m with $T_G = 60$ **a)**, and the spread complexity $\mathcal{K}_S$ as a function of time in **b)**, computed using the measurable Krylov basis. The spread complexity increases and then saturates.

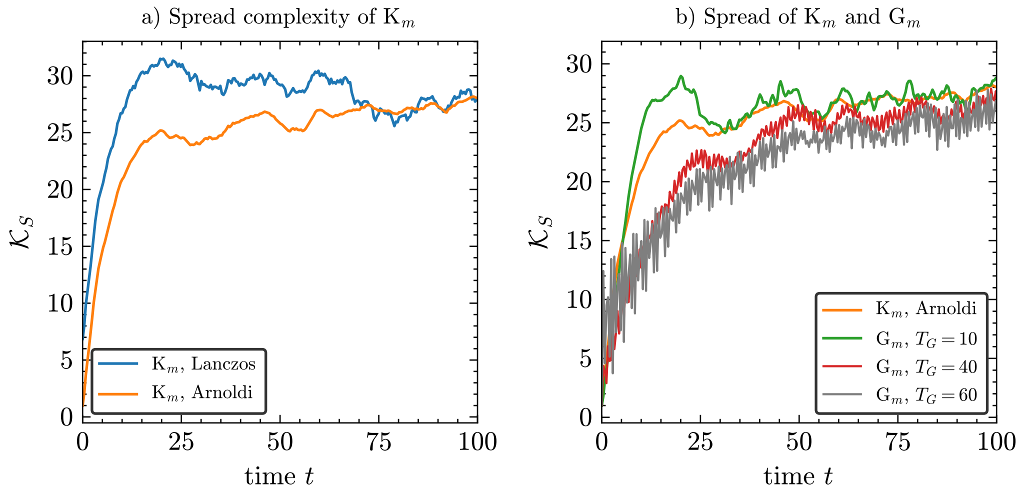


Figure III.9: Comparison of the spread complexities $\mathcal{K}_S$ computed via Hamiltonian powers H^k using Lanczos and Arnoldi algorithms (a). (b) shows $\mathcal{K}_S$ computed via powers of H (blue), and via time-evolved states in G_m with $T_G = 10, 40, 60$. We observe that for $T_G = 40$ and 60, the spread is smaller than for K_m, while it is larger for $T_G = 10$.

III.1.3 Krylov Expressivity: Phase Space Measure on the Krylov Space

Section III.1.1 introduced spread complexity, as discussed in [16]. In [81], it was shown that a measurable Krylov basis can be defined to compute spread complexity. Building upon this, Section III.1.2 further contradicts the assumption made in [16], where the authors claim that the Krylov basis minimizes the spread. This work demonstrates that this is not the case, as measurable Krylov spaces can be constructed that exhibit a slower increase in spread complexity.

This section proposes a new measure, *Krylov expressivity*, as an effective tool to capture the phase space dimension of the corresponding Krylov space.

Definition 7 (Krylov Expressivity)**.** *Let* H *be a Hamiltonian with initial condition* $|\Psi_0\rangle$*, and let* m *denote the grade of the Krylov space* K_m *or* G_m*. At any given point in time* t*, define a discretization* $t_n = nt/(m-1)$ *for* $n = 0, \ldots, m-1$*. The time-evolved states at these times, given by* $|\Psi(t_n)\rangle = e^{-iHt_n}|\Psi_0\rangle$*, are:*

$$|\Psi(t_0)\rangle\,,\ |\Psi(t_1)\rangle\,,\ \ldots,\ |\Psi(t_{m-1})\rangle\,.$$

Note that the discretization depends on the evolution time t*, and therefore these vectors are functions of it. Next, define a similarity function* f *between two pure states* $|\Phi_1\rangle$ *and* $|\Phi_2\rangle$ *such that:*

$$f(|\Phi_1\rangle\,, |\Phi_2\rangle) = 1 \quad \textit{if} \quad |\Phi_1\rangle = |\Phi_2\rangle\,.$$

Now compute the pairwise similarities between adjacent time-evolved states:

$$f_i = f(|\Psi(t_i)\rangle\,, |\Psi(t_{i+1})\rangle). \tag{III.63}$$

This function measures how similar two consecutive time-evolved states are. If f_i *is smaller than a threshold* λ*, the two vectors are considered linearly independent, and the effective phase space dimension is increased by 1. If* $f_i \geq \lambda$*, the increase is interpolated. In this work, linear interpolation is used. The cutoff parameter* λ *is inspired by the 3 dB threshold in signal processing, with* $\lambda = 1/\sqrt{2}$*. For pure states, fidelity is used as the similarity measure:*

$$f_i := \mathrm{F}\left(|\Psi(t_i)\rangle\langle\Psi(t_i)|\,, |\Psi(t_{i+1})\rangle\langle\Psi(t_{i+1})|\right) = |\langle\Psi(t_i)|\Psi(t_{i+1})\rangle|\,. \tag{III.64}$$

The ***Krylov expressivity*** $\mathcal{E}_K$ *for a given threshold* $\lambda \in [0,1]$ *is then defined as:*

$$\epsilon_i = \begin{cases} 1 & \textit{if } f_i < \lambda, \\ 1 - \frac{1}{1-\lambda}\cdot(f_i - \lambda) & \textit{if } f_i \geq \lambda, \end{cases}$$

$$\mathcal{E}_K = 1 + \sum_{i=1}^{m-1} \epsilon_i. \tag{III.65}$$

Note that different similarity functions f *and different cutoff values* λ *may be used depending on the application.*

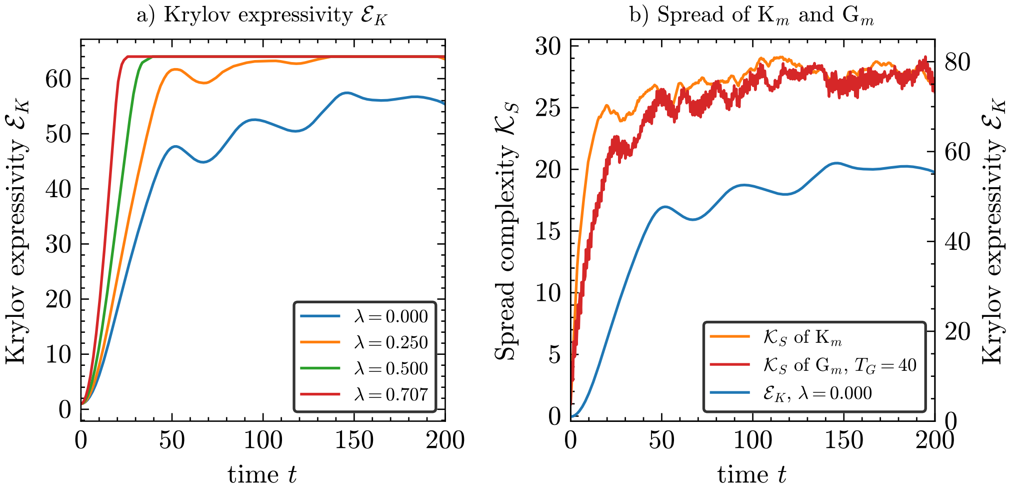


Figure III.10: Comparison of Krylov expressivities $\mathcal{E}_K$ for different cutoff parameters: $\lambda = 0$ (blue), $\lambda = 0.25$ (orange), $\lambda = 0.5$ (green), and $\lambda = 1/\sqrt{2}$ (red) in (a). Plot (b) shows the spread complexity computed via powers of the Hamiltonian (orange) and via the measurable Krylov space G_m with $T_G = 40$ (red).

Figure III.10.a shows Krylov expressivity as a function of evolution time t for various cutoff parameters $\lambda \in \{0, 0.25, 0.5, 1/\sqrt{2}\}$. Clearly, higher cutoff values result in an earlier increase in $\mathcal{E}_K$. The expressivity saturates at the Krylov space dimension $m = 64$ for all cutoff values except $\lambda = 0$.

For $\lambda = 0$, the Krylov expressivity simplifies to a parameter-independent form:

$$\mathcal{E}_K = 1 + \sum_{i=1}^{m-1} \left(1 - |\langle g_i | g_{i+1} \rangle|\right) = m - \sum_{i=1}^{m-1} |\langle g_i | g_{i+1} \rangle| \,. \tag{III.66}$$

Figure III.10.b presents the Krylov spread complexity $\mathcal{K}_S$ for the Krylov space K_m (orange) and for the time-evolved space G_m with $T_G = 40$ (red), as well as the Krylov expressivity for $\lambda = 0$ (blue, right axis). All three measures saturate around $t \approx 150$.

A deeper connection between Krylov spread complexity and Krylov expressivity requires further benchmarking in domains where spread complexity has been successful—most notably, quantum chaos. However, such an analysis lies beyond the scope of this chapter, which aims to provide conceptual intuition and introduce recent ideas.

The results on Krylov expressivity (especially the parameter-free case with $\lambda = 0$), alongside the earlier result showing smaller spread complexity for time-evolved bases, are meant to engage the reader and highlight promising directions for future research.

III.2 Krylov Spaces for Operators

This section addresses Krylov spaces for operators, with particular focus on their applications to observables, first discussed in [14]. The analysis begins by establishing that any time-evolved operator $O(t)$ lies within a Krylov space generated by successive powers of the Liouvillian, defined as $\mathcal{L}(O) = HO - OH$, such that

$$O(t) \in \mathrm{Span}\{O, \mathcal{L}(O), \mathcal{L}^2(O), \dots, \mathcal{L}^{M-1}(O)\} := \mathcal{L}_M, \tag{III.67}$$

where M denotes the grade of the corresponding Krylov space $\mathcal{L}_M$ [14].

Subsequently, it is shown that time-evolved operators $O(t_j)$ at discrete time points can be used to construct an equivalent space:

$$O(t) \in \mathrm{Span}\{O(t_0), O(t_1), O(t_2), \dots, O(t_{M-1})\} := \mathcal{F}_M. \tag{III.68}$$

It follows that $\mathcal{L}_M = \mathcal{F}_M$. The analysis is then extended to consider multiple observables $O_1, O_2, \dots, O_K$, and a space $\mathcal{F}$ is constructed containing the minimal number of basis elements such that

$$O_j(t) \in \mathcal{F} \tag{III.69}$$

for all $j \in \{1, 2, \dots, K\}$. This implies that no space $\mathcal{V}$ exists for which $O_j(t) \in \mathcal{V}$ and $\dim(\mathcal{V}) < \dim(\mathcal{F})$. Based on the space $\mathcal{F}$, the concept of Krylov observability is introduced to characterize the effective phase space dimension of multiple operators O_j with respect to time evolution governed by the Hamiltonian H.

Definition 8 (Operator Evolution)**.** *Let the quantum system be described by a Hamiltonian $H \in \mathbb{C}^{N\times N}$, where $H = H^\dagger$. Let $O, O(t) \in \mathbb{C}^{N\times N}$ represent an operator and its time-evolved form at time $t \in \mathbb{R}_{\geq 0}$, respectively. Given the initial condition O, the time evolution of the operator is defined as*

$$O(t) = e^{iHt} O e^{-iHt}. \tag{III.70}$$

To evaluate Krylov operator complexity, the framework is applied to an Ising model consisting of four spins, yielding $4^4 = 256$ total degrees of freedom. The corresponding Hamiltonian is defined as

$$H_{\mathrm{O}} = \sum_{i=1,j>i}^{4} J_{i,j} X_{i-1} X_i + \frac{1}{2} \sum_{i=0}^{4} Z_i, \tag{III.71}$$

where the coupling constants $J_{i,j}$ are selected from the set $\{0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.65\}$, ordered according to increasing indices (i, j).

III.2.1 Operator Complexity through the Powers of the Liouvillian

This section introduces how Krylov spaces can capture the time evolution of operators under the von Neumann or Heisenberg equation. It is structured similarly to the previous discussions but with a different focus. We first revisit the proof proposed in [14], which shows that any time-evolved operator lies in the Krylov space formed by powers of the Liouvillian $\mathcal{L}$.

Theorem 5. *Let $O \in \mathbb{C}^{N\times N}$ be an operator that evolves as $O(t) = e^{iHt} O e^{-iHt}$ (see Def. 8). Then, there exists $M \in \mathbb{N}$ such that for any time-evolved operator $O(t)$, the following holds:*

$$O(t) \in \mathrm{Span}\{O, \mathcal{L}(O), \mathcal{L}^2(O), \ldots, \mathcal{L}^{M-1}(O)\} := \mathcal{L}_M, \tag{III.72}$$

where the Liouvillian is defined as $\mathcal{L}(H, O) = HO - OH$.

Proof. The time evolution of the operator can be written as a series:

$$O(t) = e^{iHt} O e^{-iHt} = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} \mathcal{L}^k(O), \tag{III.73}$$

where $\mathcal{L}(O) = HO - OH$ is the Liouvillian. This shows that:

$$O(t) \in \mathrm{Span}\{\mathcal{L}^0(O), \mathcal{L}^1(O), \mathcal{L}^2(O), \ldots\}. \tag{III.74}$$

It is straightforward to verify that $\mathcal{L}$ is linear in O:

$$\mathcal{L}(O_1 + O_2) = \mathcal{L}(O_1) + \mathcal{L}(O_2), \tag{III.75}$$

$$\mathcal{L}(\mu O_1) = \mu \mathcal{L}(O_1). \tag{III.76}$$

Following [14], and by applying the Krylov space properties (see Def. 4), there exists a minimal $M \in \mathbb{N}$ such that

$$O(t) \in \mathrm{Span}\{\mathcal{L}^0(O), \ldots, \mathcal{L}^{M-1}(O)\} = \mathcal{L}_M. \tag{III.77}$$

□

Definition 9 (Operator Complexity via Liouvillian Powers)**.** *Given a Hamiltonian H and an operator O evolving under Def. 8, define the Krylov space*

$$\mathcal{L}_M = \mathrm{Span}\{O, \mathcal{L}(O), \mathcal{L}^2(O), \ldots, \mathcal{L}^{M-1}(O)\}. \tag{III.78}$$

After orthonormalizing the basis, we obtain:

$$\mathcal{L}_M = \mathrm{Span}\{\mathcal{K}_0, \mathcal{K}_1, \ldots, \mathcal{K}_{M-1}\}. \tag{III.79}$$

The time-evolved operator $O(t)$ can now be represented in this orthonormalized Krylov basis:

$$O(t) = \sum_{n=0}^{M-1} i^n \beta_n(t) \mathcal{K}_n, \tag{III.80}$$

where $\beta_n(t) = (\mathcal{K}_n, O(t))$ is the scalar product between the time-evolved operator and the basis operator $\mathcal{K}_n$.

The operator complexity $\mathcal{K}_O$ is then defined as:

$$\mathcal{K}_O(t) = \sum_{n=0}^{M-1} (n+1)|\beta_n(t)|^2. \tag{III.81}$$

While spread complexity $\mathcal{K}_S$ captures how a quantum state spreads over the Krylov basis formed from powers of the Hamiltonian H^n (see Def. 6, [16]), operator complexity measures how an operator spreads over the Krylov basis formed from the powers of the Liouvillian $[H, O]^n$ [14].

Similarly, the Krylov operator for observables can be defined as:

$$\mathrm{O}_{\mathcal{K}} := \sum_{n=0}^{M-1} (n+1)\mathcal{K}_n \mathcal{K}_n^d, \tag{III.82}$$

where $\mathcal{K}_n^d$ is the dual of $\mathcal{K}_n$, such that $\mathcal{K}_n^d O = (\mathcal{K}_n, O)$.

The expectation value of this operator is:

$$(\mathrm{O}_{\mathcal{K}}) = O(t)^d \mathrm{O}_{\mathcal{K}} O(t) = \sum_{n=0}^{M-1} (n+1)|\beta_n(t)|^2 = \mathcal{K}_O(t). \tag{III.83}$$

Lemma 4 ($\mathcal{L}_M$ Consists of the Minimum Number of Basis States)**.** *Let O be an operator and H the Hamiltonian under which O evolves. Then, there exists no space W such that $O(t) \in W$ for all t and $\dim(W) < \dim(\mathcal{L}_M)$.*

Proof. Let $0 = t_0 < t_1 < \ldots < t_Q$ be a time discretization with $t_Q < T_P$, where T_P is the period of the system. Define:

$$\tilde{O}(t_a) = \sum_{k=0}^{Q} \frac{(it_a)^k}{k!} \mathcal{L}^k(O)$$

as an approximation of the time-evolved operator at time t_a. Then,

$$\tilde{O}(t_a) = (i^0\mathcal{L}^0(O), i^1\mathcal{L}^1(O), \ldots, i^Q\mathcal{L}^Q(O)) \begin{pmatrix} 1 \\ t_a \\ \vdots \\ \frac{t_a^Q}{Q!} \end{pmatrix} .$$

Writing all $\tilde{O}(t_a)$ with pairwise distinct t_a gives:

$$(\tilde{O}(t_0), \ldots, \tilde{O}(t_Q)) = (i^0\mathcal{L}^0(O), \ldots, i^Q\mathcal{L}^Q(O))\Theta, \tag{III.84}$$

with

$$\Theta = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ t_0 & t_1 & \ldots & t_Q \\ t_0^2/2! & t_1^2/2! & \ldots & t_Q^2/2! \\ \vdots & \vdots & \ddots & \vdots \\ t_0^Q/Q! & t_1^Q/Q! & \ldots & t_Q^Q/Q! \end{pmatrix} . \tag{III.85}$$

Since the times t_a are pairwise distinct, Θ is a generalized Vandermonde matrix and hence invertible. Therefore:

$$(\tilde{O}(t_0), \ldots, \tilde{O}(t_Q))\Theta^{-1} = (i^0\mathcal{L}^0(O), \ldots, i^Q\mathcal{L}^Q(O)). \tag{III.86}$$

This shows that the time-evolved operators can reconstruct the powers of the Liouvillian. Since by Theo. 5, all time-evolved operators lie in $\mathcal{L}_M$, and each Liouvillian power can be expressed using the time-evolved operators, it follows that $\mathcal{L}_M$ has the minimum possible number of basis elements.

Therefore, no space W of lower dimension than $\mathcal{L}_M$ can contain all $O(t)$ for arbitrary t. □

Up to this point, only a single operator has been considered. However, in general, quantum systems involve multiple operators. While individual observables may be compared to one another to gain insight, it is often more informative to analyze the behavior of a set of observables $O_1, O_2, \ldots, O_K$. In particular, it is relevant to study the vector space spanned by multiple operators and to explore how operator complexity can be generalized in such scenarios.

In [79], a framework was proposed for constructing operator complexity from multiple operators, termed **multiseed operator complexity**.

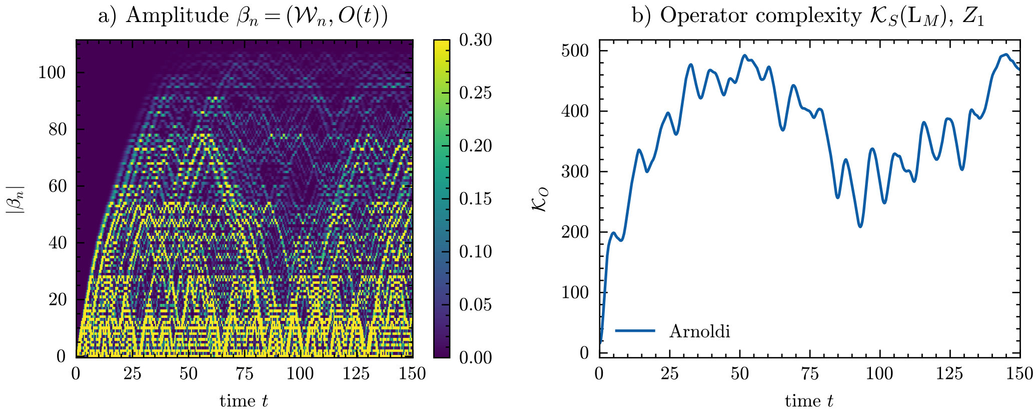


Figure III.11: Amplitudes $(\mathcal{K}_n, O(t))$, where $\mathcal{K}_n$ denotes the n-th basis state of $\mathcal{L}_M$, computed using the Arnoldi algorithm in **a)**, and the operator complexity $\mathcal{K}_O$ as a function of time in **b)**. The complexity is computed using the basis formed by the powers of the Liouvillian, i.e., $\mathcal{L}^k$.

Figure III.11.a presents the basis representation within the space spanned by powers of the Liouvillian $\mathcal{L}^k$, as a function of time t, computed using the Arnoldi algorithm. A comparable pattern is observed in the context of spread complexity (see Figure III.2), highlighting conceptual similarities between operator and spread complexity.

However, a key distinction lies in the treatment of initial conditions. For operator complexity, the initial operator is fixed—specifically, the observable Z_1 is analyzed in terms of its expansion in the Krylov space $\mathcal{L}_M$. In contrast, spread complexity involves randomly chosen initial states $|\Psi_0\rangle$, and the results are averaged. The absence of averaging in operator complexity, along with the use of a four-qubit system (as opposed to six qubits in the spread complexity analysis), results in stronger oscillatory behavior in the complexity values. Nevertheless, many insights derived from spread complexity remain applicable to the operator setting.

To mitigate these oscillations, one strategy is to average operator complexity over a set of observables. However, such an approach reduces the ability to track the dynamics of specific operators, which is often essential in this context. Therefore, the task becomes

Algorithm 1 Construction of Krylov Spaces for Multiple Operators

```
$\mathcal{I}_O = \{1, 2, \dots, K\}$
$\mathcal{L}^{(B)} = \emptyset$
$\mathcal{L} = \mathrm{Span}(\mathcal{L}^{(B)})$
$\tilde{\mathcal{L}}_1^{(B)}, \tilde{\mathcal{L}}_2^{(B)}, \dots, \tilde{\mathcal{L}}_K^{(B)} = \emptyset$
$i = 0$
while $i \in \mathcal{M} - 1$ do
    for $k \in \mathcal{I}_O$ do
        if $\mathcal{L}^i O_k \notin \mathcal{L}$ then
            $\mathcal{L}^{(B)} = \mathcal{L}^{(B)} \cup \mathcal{L}^i O_k$
            $\mathcal{L}_k^{(B)} = \mathcal{L}_k^{(B)} \cup \mathcal{L}^i O_k$
            $\mathcal{L} = \mathrm{Span}(\mathcal{L}^{(B)})$
            $\mathcal{L}_k = \mathrm{Span}(\mathcal{L}_k^{(B)})$
        end if
    end for
end while
$\mathcal{L} \leftarrow \mathrm{orthonormalize}(\mathcal{L})$
```

identifying a minimal vector space $\mathcal{L}$ such that all time-evolved operators $O_k(t)$ lie within it, i.e., $O_k(t) \in \mathcal{L}$. This problem is addressed via the concept of multiseed operator complexity, defined below.

Definition 10 (Multiseed Operator Complexity [79])**.** *Let $O_1, \dots, O_K$ be a collection of K operators, each evolving under $O_j(t) = e^{iHt} O_j e^{-iHt}$, with H as the system Hamiltonian. The following algorithm constructs a Krylov space $\mathcal{L}$ that captures all time-evolved operators, i.e., $O_j(t) \in \mathcal{L}$.*

In the original formulation in [79], orthonormalization is included within the loop to enable a termination condition. For clarity, the process is presented here as occurring at the end. To determine whether $\mathcal{L}^i O_k \in \mathcal{L}$, a rank comparison can be used. If $\mathrm{rank}(\mathcal{L}^{(B)}) = \mathrm{rank}(\mathcal{L}^{(B)} \cup \mathcal{L}^i O_k)$, then the element is already in the span.

The orthonormalized basis elements are indexed as:

$$\mathcal{L}_k = \mathrm{Span}\{\mathcal{K}_{0,k}, \mathcal{K}_{1,k}, \dots, \mathcal{K}_{M_k-1,k}\}, \tag{III.87}$$

$$\mathcal{L} = \mathrm{Span}\Big(\mathcal{K}_{0,1}, \dots, \mathcal{K}_{M_1-1,1}, \mathcal{K}_{0,2}, \dots, \mathcal{K}_{M_2-1,2}, \dots, \mathcal{K}_{0,K}, \dots, \mathcal{K}_{M_K-1,K}\Big), \tag{III.88}$$

where $M_k = \dim(\mathcal{L}_k)$, and the union of all $\mathcal{L}_k$ defines the full space $\mathcal{L}$.

Each time-evolved operator $O_n(t)$ can then be expressed as:

$$O_n(t) = e^{i\mathcal{L}t} O_n = \sum_{k=1}^{K} \sum_{j=0}^{M_k-1} (\mathcal{K}_{j,k}, O_n(t)) \mathcal{K}_{j,k}. \tag{III.89}$$

Defining $\gamma_{j,k}^{(n)} = (\mathcal{K}_{j,k}, O_n(t))$*, the operator complexity for* $O_n(t)$ *becomes:*

$$\mathcal{K}_O(O_n) = \sum_{k=1}^{K} \sum_{j=0}^{M_k - 1} (j+1) \left| \gamma_{j,k}^{(n)} \right|^2 . \tag{III.90}$$

Finally, the ***multiseed operator complexity*** $\mathcal{K}_{mult}$ *is defined as the average:*

$$\mathcal{K}_{mult}(t) = \frac{1}{K} \sum_{n=1}^{K} \mathcal{K}_O(O_n). \tag{III.91}$$

The block-Lanczos algorithm is used for orthonormalization.

At this stage, numerical results for multiseed operator complexity are omitted. The rationale is that different operator combinations describe distinct measurement regimes and experimental configurations, requiring case-specific analysis. A detailed discussion of these variations lies beyond the scope of this chapter, which focuses primarily on formal mathematical foundations. The broader goal is to demonstrate how Krylov-based complexity measures can be employed in quantum machine learning, particularly in quantum reservoir computing. In this context, a shift from Liouvillian powers to time-evolved operators is explored for Krylov space construction.

III.2.2 Operator Complexity through Time-Evolved Operators

The construction of Krylov spaces using time-evolved operators follows a similar approach to that used in Section III.1.2, where it is demonstrated that time-evolved states can be used to compute the Krylov space. First, Theo. 6 establishes the equivalence between time-evolved operator spaces and Krylov spaces constructed via the Liouvillian. Subsequently, operator complexity on this new space $\mathcal{F}$ is defined in Def. 13.

Theorem 6 (Equidistant Time-Evolved Operators Construct $\mathcal{F}_M = \mathcal{L}_M$)**.** *Let* $O \in \mathbb{C}^{N\times N}$ *be an operator that evolves under unitary time evolution governed by a Hamiltonian* $H \in \mathbb{C}^{N\times N}$*. Then, the first* M *time-evolved operators in an equidistant time discretization* $t_i = i\tau$ *span the Krylov space* $\mathcal{L}_M$*.*

Proof. By Lemma 4, there exist at least M time-evolved operators that construct the Krylov basis $\mathcal{L}_M$. In the equidistant case, the unitary time evolution is given by $U = e^{-iH\tau}$, so the time evolution becomes

$$O(\tau) = U^\dagger O U. \tag{III.92}$$

The time-evolved operator at time $t_n = n\tau$ is then

$$O(n\tau) = (U^\dagger)^n O U^n. \tag{III.93}$$

The space spanned by all time-evolved operators is

$$\mathcal{F}_Q = \mathrm{Span}\{O, U^\dagger O U, (U^\dagger)^2 O U^2, \ldots\} = \mathcal{L}_Q, \tag{III.94}$$

where the equality $\mathcal{F}_Q = \mathcal{L}_Q$ follows from Lemma 4. To confirm this, the linearity of $g(O) = U^\dagger O U$ is verified:

$$g(O_1 + O_2) = U^\dagger(O_1 + O_2)U = g(O_1) + g(O_2), \tag{III.95}$$

$$g(\mu O_1) = \mu g(O_1). \tag{III.96}$$

By the definition of a Krylov space, there exists a minimal K such that the set

$$O, U^\dagger O U, \ldots, (U^\dagger)^{K-1} O U^{K-1} \tag{III.97}$$

is linearly independent, and the inclusion of $(U^\dagger)^K O U^K$ makes the set linearly dependent. According to Lemma 4, the maximum number of linearly independent time-evolved operators is M. Therefore, $K = M$, and the first M equidistantly sampled time-evolved operators span the Krylov space $\mathcal{L}_M$. Thus,

$$\mathcal{F}_M = \mathrm{Span}\{O, O(\tau), O(2\tau), \ldots, O((M-1)\tau)\}, \tag{III.98}$$

$$\mathcal{L}_M = \mathrm{Span}\{O, \mathcal{L}(O), \mathcal{L}^2(O), \ldots, \mathcal{L}^{M-1}(O)\}, \tag{III.99}$$

and it follows that

$$\mathcal{F}_M = \mathcal{L}_M. \tag{III.100}$$

□

The space $\mathcal{F}_M$ can be used as a basis for computing operator complexity.

Definition 11 (Operator Complexity Using Time-Evolved Operators $O(t)$)**.** *Given a Hamiltonian H and an operator O evolving under Def. 8, consider the Krylov space*

$$\mathcal{F}_M = \mathrm{Span}\{O, O(\tau), O(2\tau), \ldots, O((M-1)\tau)\}. \tag{III.101}$$

This space can be orthonormalized to yield the basis $\mathcal{G}_i$*, such that*

$$\mathcal{F}_M = \mathrm{Span}\{\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_{M-1}\}. \tag{III.102}$$

Any time-evolved operator $O(t)$ *can then be represented in this basis as*

$$O(t) = \sum_{n=0}^{M-1} i^n \beta_n(t) \mathcal{G}_n, \tag{III.103}$$

where $\beta_n(t) = (\mathcal{G}_n, O(t))$ *denotes the scalar product. The operator complexity* $\mathcal{K}_O(\mathcal{F}_M, t)$ *is defined as*

$$\mathcal{K}_O(\mathcal{F}_M, t) = \sum_{n=0}^{M-1} (n+1)|\beta_n(t)|^2. \tag{III.104}$$

The associated Krylov operator is given by

$$\mathrm{O}_{\mathcal{G}} := \sum_{i=0}^{M-1} (i+1)\mathcal{G}_i \mathcal{G}_i^d, \tag{III.105}$$

where $\mathcal{G}_i^d$ *is the dual of* $\mathcal{G}_i$ *such that* $\mathcal{G}_i^d(O) = (\mathcal{G}_i, O)$*.*

The results of operator complexity $\mathcal{K}_O$ computed on spaces $\mathcal{F}_M$ with construction times $T_G = 10$, $T_G = 40$, and $T_G = 150$ are illustrated in Figure III.12, Figure III.13, and Figure III.14, respectively. In Figure III.12, the basis representation (Figure III.12.a) demonstrates rapid spreading of the observable within $\mathcal{F}_M$.

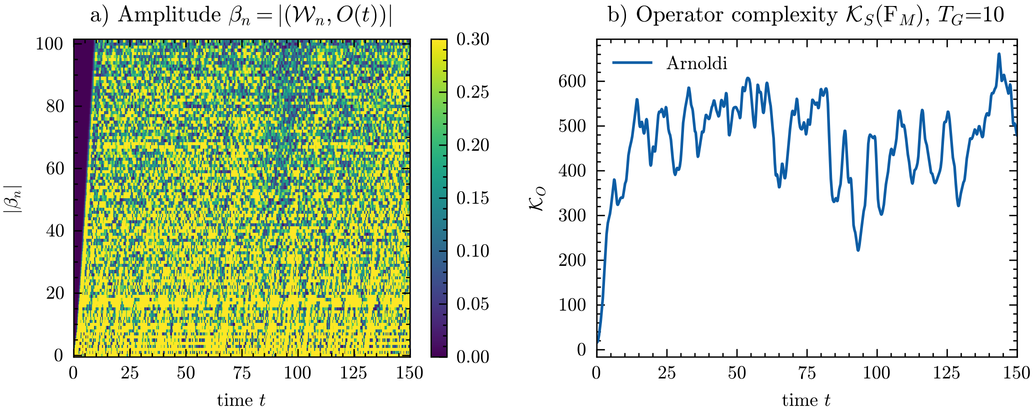


Figure III.12: Amplitudes $(\mathcal{K}_n, O(t))$, where $\mathcal{K}_n$ denotes the n-th basis state of $\mathcal{F}_M$ with $T_G = 10$ using QR decomposition in **a)** and the corresponding operator complexity $\mathcal{K}_O$ over time in **b)**.

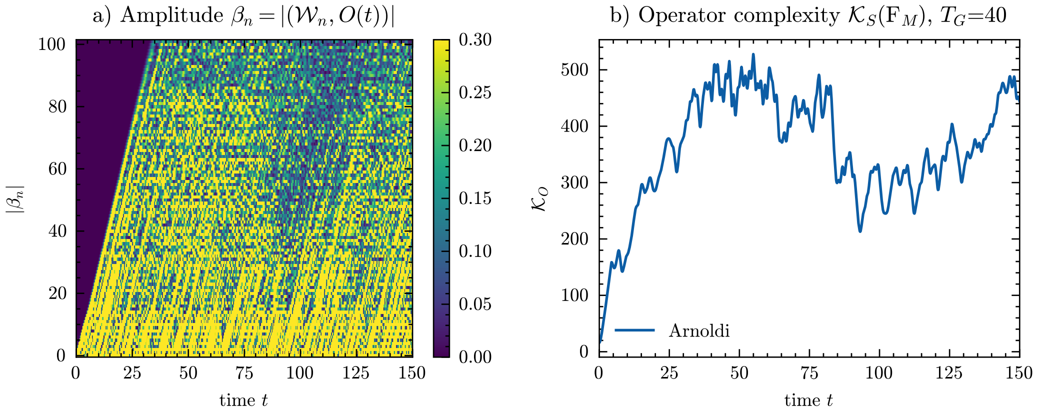


Figure III.13: Amplitudes $(\mathcal{K}_n, O(t))$, where $\mathcal{K}_n$ denotes the n-th basis state of $\mathcal{F}_M$ with $T_G = 40$ using QR decomposition in **a)** and the corresponding operator complexity $\mathcal{K}_O$ over time in **b)**.

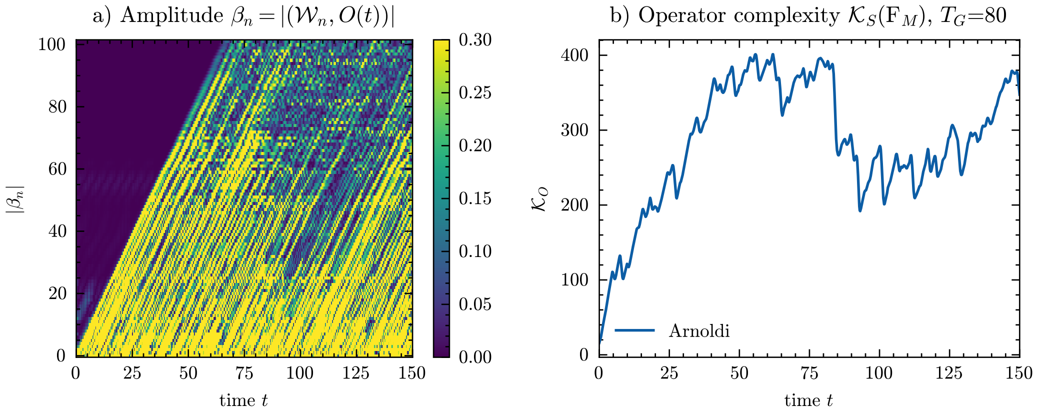


Figure III.14: Amplitudes $(\mathcal{K}_n, O(t))$, where $\mathcal{K}_n$ denotes the n-th basis state of $\mathcal{F}_M$ with $T_G = 80$ using QR decomposition in **a)** and the corresponding operator complexity $\mathcal{K}_O$ over time in **b)**.

To provide further insight, Figure III.15 presents a comparison between operator complexities computed using the Liouvillian Krylov space $\mathcal{L}_M$ and the spaces $\mathcal{F}_M$ constructed from time-evolved operators at $T_G = 10$, $T_G = 40$, and $T_G = 80$. The operator complexity $\mathcal{K}_O$ for $T_G = 10$ (green) exceeds the value obtained from the conventional Liouvillian construction. At $T_G = 40$, results from time-evolved operators (red) and Liouvillian powers (orange) are nearly identical. For $T_G = 80$, the time-evolved operator space yields lower operator complexity across all times compared to the Liouvillian-based approach.

These results raise questions about whether operator and spread complexity are appropriate measures. In general, any complexity measure should ideally be constructed

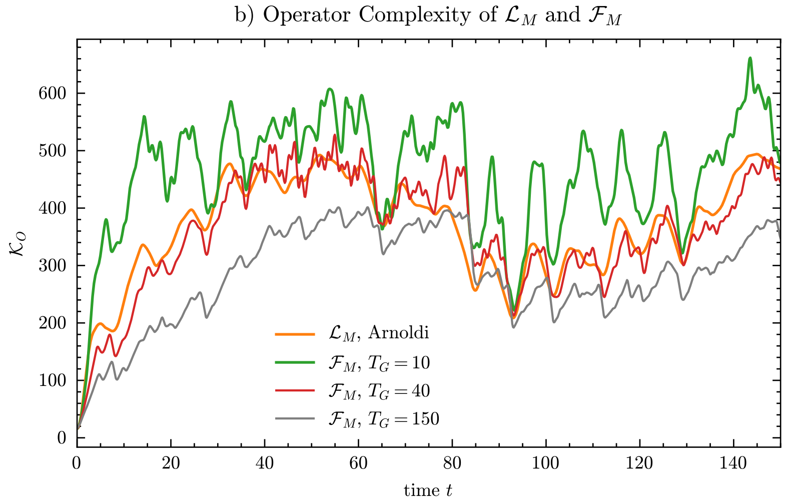


Figure III.15: Comparison of operator complexities $\mathcal{K}_O$ computed using the powers of the Liouvillian $\mathcal{L}^k$ via the Arnoldi algorithm (orange), and operator complexities computed on the spaces consisting of time-evolved operators $\mathcal{F}_M$ using three different construction times: $T_G = 10$ (green), $T_G = 40$ (red), and $T_G = 80$ (grey).

to be minimal. Therefore, the minimization of Krylov operator complexity and Krylov spread complexity serves as a foundational assumption. With respect to Krylov spread complexity, these findings directly contradict the proof presented in [16].

Such a contradiction suggests the need to re-evaluate existing literature that derives physical insights based on this complexity measure. Many works that compute spread complexity rely on the assumption that the Krylov basis minimizes the measure, using this minimization as justification for its utility.

Definition 12 (Multiseed Operator Complexity Using Time-Evolved Operators)**.** *Instead of using powers of the Liouvillian, as outlined in Theo. 5, time-evolved operators can be utilized for constructing Krylov spaces when equidistant sampling is employed. Let* $O_1, O_2, \dots, O_K$ *be operators that evolve under the time evolution* $O_k(t) = e^{iHt} O_k e^{-iHt}$*, and define the space* $\mathcal{F}$ *constructed from time-evolved operators such that for all* $k \in \{1, \dots, K\}$ *and all* $t \in \mathbb{R}$*,*

$$O_k(t) \in \mathcal{F} := \mathrm{Span}\{O_1(t_1), \dots, O_1(t_{M_1}), O_2(t_1), \dots, O_2(t_{M_2}), \dots, O_K(t_1), \dots, O_K(t_{M_K})\}. \tag{III.106}$$

The corresponding subspaces $\mathcal{F}_k$ *are defined as*

$$\mathcal{F}_k = \mathrm{Span}\{O_k(t_1), O_k(t_2), \dots, O_k(t_{M_k})\}. \tag{III.107}$$

Algorithm 2 Construction of Observability Spaces

```
1: $\mathcal{I}_O = \{1, 2, \ldots, K\}$
2: $\mathcal{T} = \{t_1, t_2, \ldots, t_R\}$
3: $\mathcal{F}^{(B)} = \emptyset$
4: $\tilde{\mathcal{F}}_1^{(B)}, \tilde{\mathcal{F}}_2^{(B)}, \ldots, \tilde{\mathcal{F}}_K^{(B)} = \emptyset$
5: while $t_j \in \mathcal{T}$ do
6:     for $k \in \mathcal{I}_O$ do
7:         if $O_k(t_j) \notin \mathcal{F}$ then
8:             $\mathcal{F}^{(B)} = \mathcal{F}^{(B)} \cup O_k(t_j)$
9:             $\mathcal{F}_k^{(B)} = \mathcal{F}_k^{(B)} \cup O_k(t_j)$
10:            $\mathcal{F} = \mathrm{Span}\{\mathcal{F}^{(B)}\}$
11:            $\mathcal{F}_k = \mathrm{Span}\{\mathcal{F}_k^{(B)}\}$
12:        end if
13:    end for
14: end while
15: $\mathcal{F}^{(B)} \leftarrow \mathrm{orthonormalize}(\mathcal{F}^{(B)})$
```

The following algorithm can be used to construct these spaces:

The index set $\mathcal{I}_O$ and the time set $\mathcal{T}$ are initialized in lines 1 and 2. It is required that $R > M$, and for the purposes of this implementation, $R = \dim(H)$ is used. Lines 3 and 4 initialize the basis sets as empty. The while-loop in line 5 iterates over all time steps t_j, and the for-loop in line 6 processes each operator O_k. If a time-evolved operator $O_k(t_j)$ is not already in the space $\mathcal{F} = \mathrm{Span}(\mathcal{F}^{(B)})$, it is added to the basis in lines 8 and 9.

Verification can be performed via orthonormalization or matrix rank comparison: if

$$\mathrm{rank}(\mathcal{F}^{(B)}) = \mathrm{rank}(\mathcal{F}^{(B)} \cup O_k(t_j)),$$

then it follows that $O_k(t_j) \in \mathcal{F}$. Lines 10 and 11 update the spaces $\mathcal{F}$ and $\mathcal{F}_k$, and the final orthonormalization is applied to $\mathcal{F}^{(B)}$. This construction ensures the returned spaces satisfy:

$$\mathcal{F} = \bigcup_{k=1}^{K} \mathcal{F}_k = \bigcup_{k=1}^{K} \tilde{\mathcal{F}}_k, \quad \dim\left(\bigcup_{j=1}^{l} \tilde{\mathcal{F}}_j\right) = \sum_{j=1}^{l} \dim(\tilde{\mathcal{F}}_j).$$

This ensures that the spaces $\mathcal{F}_k$ are linearly independent and that each observable O_k is mapped to its corresponding space. After orthonormalization, the bases are denoted as

$\mathcal{G}_{j,k} = O_k(t_j)$:

$$\mathcal{F} = \text{Span}\{\mathcal{G}_{1,1}, \dots, \mathcal{G}_{M_1,1}, \dots, \mathcal{G}_{1,K}, \dots, \mathcal{G}_{M_K,K}\}, \tag{III.108}$$

$$\mathcal{F}_k = \text{Span}\{\mathcal{G}_{1,k}, \mathcal{G}_{2,k}, \dots, \mathcal{G}_{M_k,k}\}. \tag{III.109}$$

Each time-evolved operator $O_n(t)$ *can then be represented in* $\mathcal{F}$ *as*

$$O_n(t) = e^{iHt} O_n e^{-iHt} = \sum_{k=1}^{K} \sum_{j=0}^{M_k-1} (\mathcal{G}_{j,k}, O_n(t)) \mathcal{G}_{j,k}. \tag{III.110}$$

With $\eta_{j,k}^{(n)} = (\mathcal{G}_{j,k}, O_n(t))$, *the operator complexity of* $O_n(t)$ *is defined by*

$$\mathcal{K}_O(O_n) = \sum_{k=1}^{K} \sum_{j=0}^{M_k-1} (j+1) |\eta_{j,k}^{(n)}|^2. \tag{III.111}$$

Finally, the multiseed operator complexity is defined as

$$\mathcal{K}_{mult,(\mathcal{F})}(t) = \frac{1}{K} \sum_{n=1}^{K} \mathcal{K}_O(O_n). \tag{III.112}$$

III.2.3 Capturing the Phase Space Dimension of Operators with Krylov Observability

The previous section discussed operator complexity and its construction using time-evolved operators. Subsequently, multiseed operator complexity was introduced, along with its counterpart based on time-evolved operators. In analogy to Krylov expressivity for pure states, this section proposes an effective phase space measure for multiple operators $O_1, \dots, O_K$ measured V times. This measure is referred to as **Krylov Observability**.

Definition 13 (Krylov Observability). *Let* K *observables* $O_1, \dots, O_K$ *be given. For each observable* O_k, *compute the space* $\mathcal{F}_k$ *with* $\dim(\mathcal{F}_k) = M_k$, *and let* $T > 0$. *Two cases are distinguished:*

1. *When considering full Krylov observability for understanding system evolution, set* $R_k = M_k$ *for each observable.*

2. *When observables* O_k *are measured* V_k *times, set* $R_k = \min(V_k, M_k)$. *This case is particularly relevant in applications such as quantum reservoir computing (where* $V_k < M_k$ *is common) and quantum machine learning (often* $V_k = 1$*). This ensures*

that Krylov observability reflects only the minimally necessary number of linearly independent basis elements.

Define the discretization up to time t as $\tau_k = t/R_k$. The time-evolved operators are listed as:

$$\mathcal{P}_k = \{O_k(\tau_1), O_k(\tau_2), \dots, O_k(\tau_{M_k})\}. \tag{III.113}$$

The observability of the k-th operator is defined by:

$$p_k(T) = 1 + \sum_{j=1}^{R_k - 1} \left(1 - f(O_k(\tau_j), O_k(\tau_{j+1}))\right), \tag{III.114}$$

where $f(A, B) = 1$ if $A = B$. In this context, the normalized fidelity function is used:

$$f(A, B) := \mathrm{F}(A, B) = \left| \mathrm{Tr}\left(\frac{A^\dagger B}{\|A\| \|B\|} \right) \right|. \tag{III.115}$$

The ***Krylov Observability*** *$\mathcal{O}_K(T)$ for the multiplexed observables $O_1, \dots, O_K$ is defined as:*

$$\mathcal{O}_K(T) = \sum_{k=1}^{K} p_k(T). \tag{III.116}$$

Figure III.16 presents the Krylov observability for the Ising Hamiltonian in Eq. (III.71), considering measurements at one (a), two (b), three (c), and all four (d) sites in the Pauli-Z direction. A consistent trend is observed: increasing the number of observables leads to higher overall Krylov observability. When only one site is measured, the highest values of $\mathcal{O}_K$ appear in the upper right corner, as indicated by the brightest region in (a). With more observables included, this region shifts towards smaller times t and fewer measurement steps V. This behavior reflects the fact that the number of basis states depends on both the number of observables and the number of measurement instances. This chapter introduced several theoretical results and proposed novel quantum information measures that quantify the phase space dimension of the Krylov space. These developments raise a number of open questions, the exploration of which lies beyond the scope of this thesis.

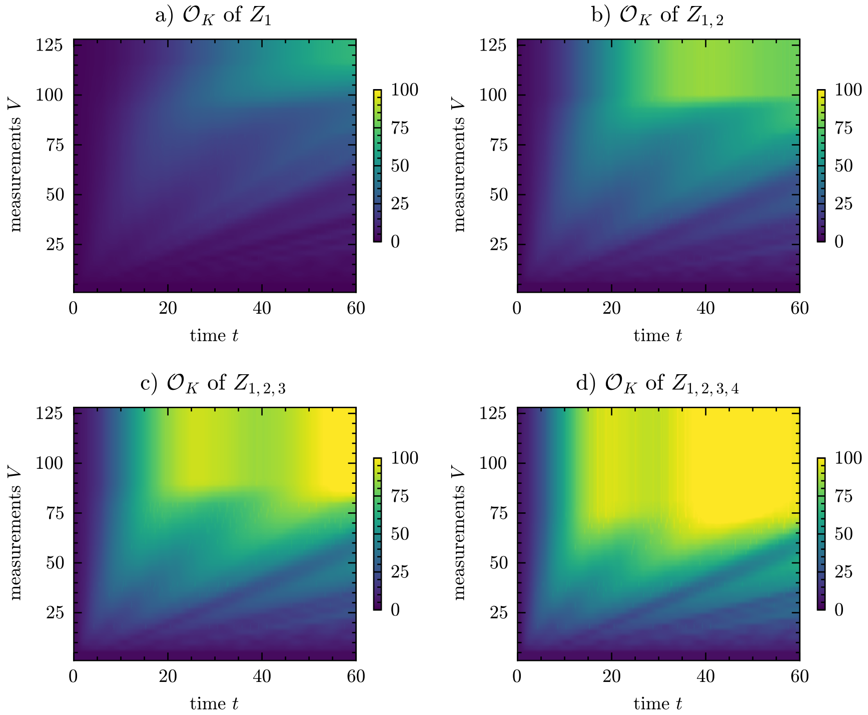


Figure III.16: Krylov observability $\mathcal{O}_K$ computed for the Ising Hamiltonian H_O (Eq. (III.71)), considering one (a), two (b), three (c), and four (d) observables. The observability is shown as a function of the number of measurements V and time t.

IV Chapter: Quantum Reservoir Computing

The previous chapter discussed how Krylov spaces can be used to understand time evolution in quantum mechanics. Krylov spread and operator complexity were defined and examined in terms of how a time-evolved quantum state or operator evolves distributes over the Krylov basis [14, 16]. It was then shown that time-evolved states and time-evolved operators can be used to construct spaces equivalent to the Krylov spaces involved in defining Krylov complexity. In the context of quantum states and spread complexity, it was demonstrated that such spaces can be reconstructed via quantum measurements, making them experimentally accessible. Additionally, this method allows for the determination of the number of pairwise distinct eigenenergies present in the system. The findings further challenge the widely accepted assumption that the Krylov basis is the optimal basis for minimizing Krylov spread complexity [16]. In fact, it was shown that multiple alternative bases exist that yield a smaller spread than the Krylov basis.

Similar conclusions were drawn for Krylov operator complexity: the constructed time-evolved spaces can exhibit reduced complexity, provided that the evolution times are chosen appropriately. Typically, a complexity measure should be defined in such a way that it is minimized—this is the main argument in favor of Krylov spread complexity over operator complexity. One might assume that this weakens the concept of Krylov complexity; however, it instead highlights an important aspect of this novel complexity measure, thereby deepening our understanding of its limitations, potential, and applications. These time-evolved spaces were then used to introduce a measure of the effective phase space dimension of the Krylov space, namely Krylov expressivity for states and Krylov observability for operators and mixed states [81, 82, 83]. This chapter will explore the extent to which these measures can be employed to explain task performance in quantum reservoir computing [26].

Before delving into the interplay between quantum reservoir computing and Krylov-based information measures, a brief recap will be provided. This serves two purposes: it helps the reader refresh their memory and provides context.

IV.1 Summary of Krylov Spaces and Quantum Reservoir Computing

The first information measure that will be discussed as a predictor for quantum reservoir computing task performance is the fidelity F of two quantum states ρ and σ.

$$\mathrm{F}(\rho,\sigma)=\left(\mathrm{Tr}\left[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right]\right)^2. \tag{IV.1}$$

In the case of pure states, i.e., if $\rho=|\psi\rangle\langle\psi|$ and $\sigma=|\phi\rangle\langle\phi|$, the fidelity simplifies to:

$$F=|\langle\psi|\phi\rangle|^2. \tag{IV.2}$$

For Hermitian operators, a normalized inner product is defined as

$$F(A,B):=\left|\mathrm{Tr}\left(\frac{A^\dagger B}{\|A\|\|B\|}\right)\right|. \tag{IV.3}$$

Table IV.1 summarizes how Krylov spaces are used to describe pure quantum states. First, the Schrödinger equation is presented as the equation of motion, followed by the unitary time evolution. Then, the Krylov space property of the time evolution is shown; that is, it is demonstrated that there exists a Krylov space K_m such that $|\psi(t)\rangle\in\mathrm{K}_m$ holds.

Based on this, Krylov spread complexity $\mathcal{K}_S$ is introduced and interpreted as the expectation value of the corresponding Krylov spread operator K_S. The evolved Krylov spaces G_m are presented as an alternative construction for Krylov complexity. Finally, the definition of Krylov expressivity $\mathcal{E}_K$ is given. Compared to Krylov spread complexity $\mathcal{K}_S$, Krylov expressivity $\mathcal{E}_K$ can be understood as an effective phase space dimension at each point in time.

Table IV.1: Summary: Krylov spaces for pure state evolution

Equation of motion	$\partial_t \lvert\Psi(t)\rangle = -iH \lvert\Psi(t)\rangle$ $\lvert\Psi(0)\rangle := \lvert\Psi_0\rangle$
Time evolution	$\lvert\Psi(t)\rangle = e^{-iHt} \lvert\Psi_0\rangle = \sum_{k=0}^{\infty}(-iH)^k \frac{t^k}{k!} \lvert\Psi_0\rangle$
Krylov space	$\lvert\Psi(t)\rangle \in \mathrm{K}_m := \mathrm{Span}\{H^0 \lvert\Psi_0\rangle, \ldots, H^{m-1} \lvert\Psi_0\rangle\}$ Orthonormalizing results in $\mathrm{K}_m = \mathrm{Span}\{\lvert k_0\rangle, \lvert k_1\rangle \ldots, \lvert k_{m-1}\rangle$
Spread complexity $\mathcal{K}_S$	$\mathcal{K}_S(t) = \sum_{n=0}^{m-1}(n+1)\lvert\alpha_n(t)\rvert^2.$ with $\alpha_n := \langle k_n \vert \Psi(t)\rangle$
Interpretation of $\mathcal{K}_S$	Krylov spread operator: $\mathrm{K}_S := \sum_{i=0}^{m-1}(n+1) \lvert k_n\rangle \langle k_n\rvert$ $\mathcal{K}_S$ is the expectation value of the Krylov spread operator $\mathcal{K}_S(t) = \langle\psi(t)\rvert\, \mathrm{K}_S \lvert\psi(t)\rangle$
Evolved Krylov space	$\lvert\Psi(t)\rangle \in \mathrm{G}_m := \mathrm{Span}\{e^{iHt_0} \lvert\Psi_0\rangle, \ldots, e^{iHt_{m-1}} \lvert\Psi_0\rangle\}$ Orthonormalizing results in $\mathrm{G}_m = \mathrm{Span}\{\lvert g_0\rangle, \lvert g_1\rangle \ldots, \lvert g_{m-1}\rangle$
Krylov expressivity $\mathcal{E}_K$	For any t discretize with $t_n = nt/(m-1)$ to get the states $\lvert\Psi(t_0)\rangle,\ \lvert\Psi(t_1)\rangle, \ldots,\ \lvert\Psi(t_{m-1})\rangle.$ Compute overlap: $f_i = \lvert\langle\Psi(t_i)\vert\Psi(t_{i+1})\rangle\rvert$ and the effective dimension $epsilon_i$ with $\lambda = \frac{1}{\sqrt{2}}$ $\epsilon_i = \begin{cases} 1 & \text{if } f_i < \lambda \\ 1 - \frac{1}{1-\lambda}\cdot(f_i - \lambda) & \text{if } f_i \geq \lambda \end{cases}$ resulting in Krylov expressivity $\mathcal{E}_K = 1 + \sum_{i=1}^{m-1}\epsilon_i$

Table IV.2: Summary: Krylov spaces for operator evolution

Equation of motion	$\partial_t \mathrm{O}(t) = i[H, \mathrm{O}]$ $\mathrm{O}(0) =: \mathrm{O}$
Time evolution	$O(t) = e^{iHt} O e^{-iHt} = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} \mathcal{L}^k(O),$ with $\mathcal{L}(O) = [H, O]$
Krylov space	$O(t) \in \mathrm{Span}\{\mathcal{L}^0(O), \mathcal{L}^1(O), \ldots, \mathcal{L}^{M-1}(O)\} = \mathcal{L}_M$ Orthonormalizing: $\mathcal{L}_M = \mathrm{Span}\{\mathcal{K}_0, \mathcal{K}_1, \ldots, \mathcal{K}_{M-1}\}$
Operator complexity $\mathcal{K}_S$	$\mathcal{K}_O(t) = \sum_{n=0}^{M-1} (n+1) \lvert\beta_n(t)\rvert^2$ with $\beta_n(t) := (\mathcal{K}_n, O(t)) = \mathrm{Tr}\Big(W_n^\dagger O\Big)$
Interpretation	Krylov operator: $\mathrm{K}_O := \sum_{i=0}^{M-1} (n+1) \mathcal{K}_i \mathcal{K}_i^d$ $\mathcal{K}_S$ is the expectation value of the operator K_O $\mathcal{K}_O(t) = \langle \mathrm{K}_O \rangle$
Evolved Operator spaces	$O(t) \in \mathcal{F}_M = \mathrm{Span}\{O, O(t_1), O(t_2), \ldots, O(t_{M-1})\}$
Krylov observability	For operators $O_1, \ldots, O_K$, calculate the spaces $\mathcal{F}_k(t)$ and $\mathcal{F}(t) = \cup_k \mathcal{F}_k(t)$. Define Krylov observability $\mathcal{O}_K$ on $\mathcal{F}(t)$.

Table IV.3: Summary: Quantum Reservoir Computing

Input	Time series $\mathbf{u} = (u_1, u_2, \ldots, u_{N_U})$
Encoding	$\lvert\Psi_n\rangle = \sqrt{\frac{1-u_n}{2}}\,\lvert 0\rangle + \sqrt{\frac{1+u_n}{2}}\,\lvert 1\rangle$
Input state	$\rho_n = \lvert\Psi_n\rangle\langle\Psi_n\rvert \otimes \mathrm{Tr}_1(\rho_{n-1}(T))$
Measurement	Compute the expectation values $O_1, O_2, \ldots, O_K$ at times $\tau_j = jT/V$: $\langle O_{n,i}(\tau_j)\rangle = \mathrm{Tr}(O_i \rho_n(\tau_j))$ Construct state matrix $\mathbf{S} \in \mathbb{R}^{N_U \times N_R}$. N_R is the readout dimension.
Output	$\mathbf{Y} = \mathbf{SW}$
Training	Weights $\mathbf{W}$ are computed using a training set $\mathbf{u}_{\mathrm{Tr}}$ and target $\hat{\mathbf{Y}}_{\mathrm{Tr}}$: $\mathbf{W} = (\mathbf{S}_{\mathrm{Tr}}^T \mathbf{S}_{\mathrm{Tr}})^{-1} \mathbf{S}_{\mathrm{Tr}}^T \hat{\mathbf{Y}}_{\mathrm{Tr}}$
Evaluation	Evaluate NRMSE on a testing $\mathbf{u}_{\mathrm{Te}}$ and target $\hat{\mathbf{Y}}_{\mathrm{Te}}$.
Lorenz Task	$\sigma = 10.0,\ \rho = 28.0,\ \beta = \frac{8}{3}$ Equation: $\frac{dx}{dt} = \sigma(y - x),$ $\frac{dy}{dt} = x(\rho - z) - y,$ $\frac{dz}{dt} = xy - \beta z$ Discretization with $t_n = n\Delta t$ $x \to x$ Input: $x(t_n)$; Target: $x(t_{n+p})$ $x \to z$ Input: $x(t_n)$; Target: $x(z_{n+p})$
IPC	The IPC measures how the reservoir can map and retain data from a uniform time-series $u \in \mathcal{U}([-1, 1])$ non-linearly.

The summary for Krylov operator complexity is given in Table IV.2. This table is organized similarly to Table IV.1, where the equation of motion considered is the Heisenberg equation for operators, resulting in unitary time evolution. The corresponding Krylov space $\mathcal{L}_M$ consists of different powers of the Liouvillian $\mathcal{L}^k$, with $\mathcal{L} := [H, \cdot]$. Any time-evolved operator is then represented in the orthonormalized basis of $\mathcal{L}_M$, and Krylov operator complexity is defined based on this basis representation in the Krylov space $\mathcal{L}_M$. The interpretation that Krylov operator complexity can be understood as the expectation value of the corresponding Krylov operator K_O is provided. Similar to the case of quantum states, it is shown that time-evolved operators can be used to construct a Krylov space $\mathcal{F}_M$ equivalent to $\mathcal{L}_M$, i.e., $\mathcal{F}_M = \mathcal{L}_M$. The discussion regarding Krylov spread and operator complexity of these time-evolved spaces is not revisited, but their significance in exhibiting smaller Krylov complexity is noted. Lastly, the idea behind Krylov operator complexity—which aims to quantify an effective phase space dimension across multiple operators measured V times—is presented. The last aspect to be revisited is quantum reservoir computing, which is summarized in Table IV.3. The time series $\mathbf{u} = (u_1, u_2, \dots, u_{N_U})$ is encoded into one qubit as

$$|\Psi_n\rangle = \sqrt{\frac{1-u_n}{2}}\,|0\rangle + \sqrt{\frac{1+u_n}{2}}\,|1\rangle\,. \tag{IV.4}$$

The full system state is represented by the density matrix

$$\rho_n = |\Psi_n\rangle\langle\Psi_n| \otimes \mathrm{Tr}_1\left(\rho_{n-1}(T)\right), \tag{IV.5}$$

where T (referred to as the clock cycle) is the evolution time before the next input is introduced. The system evolves under the Hamiltonian H according to

$$\rho_n(t) = e^{-iHt}\rho_n e^{iHt}. \tag{IV.6}$$

To extract features from the reservoir, the expectation values of K observables $O_1, O_2, \dots, O_K$ are measured at discrete time points $\tau_j = jT/V$. This yields the values

$$\langle O_{n,i}(\tau_j)\rangle = \mathrm{Tr}\left(O_i \rho_n(\tau_j)\right), \tag{IV.7}$$

resulting in a total of $N_R = VK$ readout values per input. These are arranged row-wise for each input n to form the state matrix $\mathbf{S} \in \mathbb{R}^{N_U \times N_R}$. To account for shot noise and provide regularization, Gaussian noise $\mathcal{N}$ is added as $\mathbf{S} \leftarrow \mathbf{S} + \eta\mathcal{N}$. The final output of the reservoir is computed as $\mathbf{Y} = \mathbf{S}\mathbf{W}$, where $\mathbf{W}$ denotes the readout weights, trained using a labeled training set $\mathbf{u}_{\mathrm{Tr}}$. The weights are calculated via ridge regression as

$$\mathbf{W} = \left(\mathbf{S}_{\mathrm{Tr}}^{T}\mathbf{S}_{\mathrm{Tr}}\right)^{-1}\mathbf{S}_{\mathrm{Tr}}^{T}\hat{\mathbf{Y}}_{\mathrm{Tr}}. \tag{IV.8}$$

The evaluated measures are the Lorenz prediction tasks and the information processing capacity and the quantum reservoir is described by the Ising Hamiltonian

$$H_I = \sum_{i=1,j>i}^{N_S} J_{ij}X_iX_j + \sum_{i=1}^{N_S} hZ_i \tag{IV.9}$$

, where N_S is the number of sites.

IV.2 Quantum Reservoir Computing Maps Data onto the Krylov Space

The definitions of Krylov expressivity and observability might seem interesting up to this point, but the question arises whether it is necessary to introduce them. It may well be that Krylov complexity can fully describe task performance in quantum reservoir computing, and that introducing a measure of the Krylov space dimension does not provide additional insight, thus resulting in redundant propositions. In this regard, scientific rigor requires a test of whether, and to what extent, quantum reservoir computing task performance can be explained by Krylov complexity, expressivity, and observability. Since the initial motivation of this work was to explain the saturation behavior as a function of the clock cycle T and the number of measurements V, observed in Section II.4, it only makes sense to examine whether quantum fidelity, Krylov complexity, Krylov expressivity, and Krylov observability can explain this behavior.

Krylov spread complexity was first explored to gain understanding in quantum reservoir computing in [84]. The authors showed that quantum reservoirs with larger mean spread complexity tended to perform better. However, this observation was valid only for the same reservoir with different clock cycles T. The authors analyzed different classes of reservoirs, which made effective comparison difficult. The results were shown for one clock cycle T, where Krylov spread complexity is a measure of how spread complexity evolves over the clock cycle. Therefore, it would be reasonable to discuss the late time behavior of spread complexity in relation to task performance, requiring different clock cycles for each quantum reservoir. However, this would not make sense from the perspective of reservoir computing. Ideally, one defines quantum reservoirs that exhibit similar complexity growth. In this work, the Ising model with slightly modified inter-spin couplings is analyzed as a quantum reservoir. Initially, our results show that Krylov complexity and fidelity exhibit an anticorrelation with task performance and fail to explain performance in quantum reservoir computing beyond the initial increase in time. Additionally, it is shown that the introduced measures of Krylov expressivity and Krylov observability successfully capture this behavior of task saturation and outperform Krylov spread complexity in this regard.

Saturation in task performance was also discussed in [85] for quantum extreme learning machines. The authors show that in quantum extreme learning machines, state estimation can occur beyond the scrambling time, which challenges the common belief that information cannot be retrieved after this point [85]. A similar pattern is observed in quantum reservoir computing, where task performance initially increases and then saturates [86]. This phenomenon can be explained by Krylov expressivity and Krylov observability, where the data is encoded into the initial state. The initial state then evolves into the corresponding Krylov space, where Krylov expressivity and observability increase before reaching a saturation point. Beyond this point, further increases in time do not enhance the system's expressivity, resulting in stable task performance.

IV.2.1 Understanding Saturation in Quantum Reservoir Computing

Figure IV.1 shows the NRMSE for the Lorenz prediction and cross-prediction tasks (a), Krylov spread complexity $\mathcal{K}_S$ and fidelity F (b), and Krylov expressivity $\mathcal{E}_K$ and observability $\mathcal{O}_K$ (c), as functions of the clock cycle T.

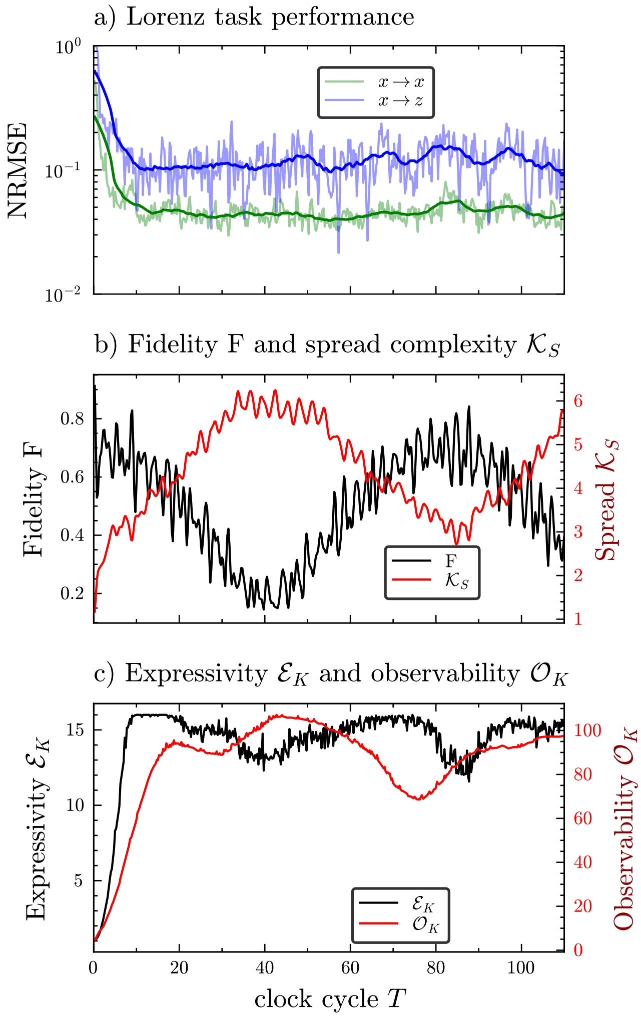


Figure IV.1: Lorenz task performance for the five-step-ahead prediction (with discretization $\Delta t = 0.02$) of the x-variable (green) and cross-prediction of the z-variable (blue) are shown in (a) as functions of the clock cycle T, when measuring all sites $V = 30$ times under the H_{I4} Hamiltonian. The bold lines represent the first-order Savitzky–Golay filtered NRMSE, used to better illustrate the saturation in task performance (computed using `scipy.signal.savgol_filter`). Fidelity F (black) and spread complexity $\mathcal{K}_S$ (red) are shown in (b), while Krylov observability $\mathcal{O}_K$ (red) and expressivity $\mathcal{E}_K$ (blue) are shown in (c).

The normalized root mean square error (NRMSE) is shown in Figure IV.1(a) for the Lorenz tasks. The green curve represents the NRMSE for the five-step-ahead prediction of the x-variable, while the blue curve represents the cross-prediction of the z-variable. An initial decline in error is followed by saturation in performance around $T \approx 14$.

The first question to be addressed is whether spread complexity and fidelity can explain this saturation in task performance. Figure IV.1(b) shows fidelity and spread complexity, averaged over twenty randomly selected initial states. The spread complexity $\mathcal{K}_S$ increases with the clock cycle T, while fidelity F exhibits a decreasing trend. An inverse relationship emerges: as fidelity reaches a minimum around $T \approx 36$, spread complexity peaks. Conversely, when fidelity increases, spread complexity declines.

This behavior captures some aspects of the NRMSE. A decrease in fidelity F indicates how far the system state has deviated from its initial configuration, aligning with the expectation that the system must evolve from its initial condition to perform computation. At $T = 0$, the reservoir behaves like an identity operation. In contrast, spread complexity $\mathcal{K}_S$ characterizes how the evolved state distributes across the Krylov basis over time. Higher spread complexity implies a broader distribution within the Krylov basis at a given clock cycle.

However, neither fidelity nor spread complexity alone fully account for the saturation observed in NRMSE. The oscillatory behavior of these quantities in Figure IV.1(b) is not reflected in task performance in Figure IV.1(a).

Figure IV.1(c) presents Krylov expressivity $\mathcal{E}_K$ and observability $\mathcal{O}_K$. Both metrics show an initial increase followed by a plateau, around which oscillations occur. This behavior closely resembles the saturation trend in task performance, suggesting that $\mathcal{E}_K$ and $\mathcal{O}_K$ are more indicative of performance saturation than fidelity and spread complexity.

IV.2.2 Krylov Observability in Quantum Reservoir Computing

After discussing the saturation behavior and demonstrating that Krylov spread complexity and fidelity cannot fully account for it, it is reasonable to investigate to what extent the data generalization capabilities of quantum reservoir computing can be explained by the proposed Krylov-based measures.

In reservoir computing, data generalization is commonly quantified by the *information processing capacity* (IPC), which depends on the encoding, the reservoir dynamics, and the readout dimension. The encoding defines how macroscopic input data is mapped into the system's quantum states. The reservoir dynamics are governed by unitary time evolution, determined by the system Hamiltonian. The readout dimension is characterized by the choice of observables and the number of measurements performed.

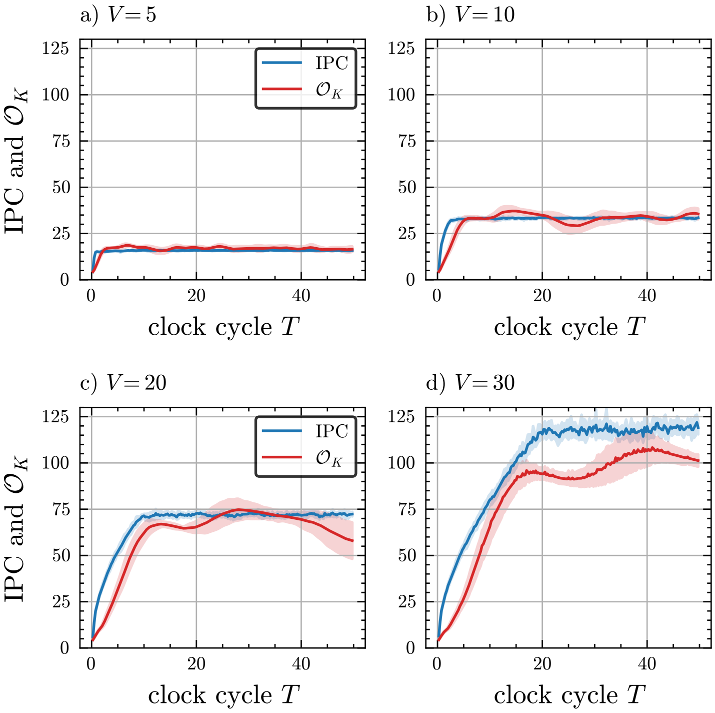


Figure IV.2: Krylov observability $\mathcal{O}_K$ and information processing capacity (IPC), averaged over ten Hamiltonians, are shown as functions of clock cycle T, shown for different numbers of measurements: $V = 5$ (a), $V = 10$ (b), $V = 20$ (c), and $V = 30$ (d). These plots illustrate how both metrics evolve with increasing T, highlighting the onset of saturation and oscillatory behavior in both observability and generalization performance.

In quantum reservoir computing, the evolution of the reservoir and the readout play a crucial role in determining the system's ability to express data. This is especially important because standard encoding methods often produce mixed quantum states, as the system evolves continuously over time. Krylov observability, which describes a set of observables $O_1, \ldots, O_K$ under unitary time evolution and across multiple measurements V, offers a compelling framework for characterizing the data generalization capacity of quantum reservoirs, while also providing a physical interpretation.

In this context, Figure IV.2 shows the information processing capacity IPC and Krylov observability $\mathcal{O}_K$ as functions of the clock cycle T, for $V = 5$, 10, 20, and 30 measurements. Plotting both quantities as functions of T enables evaluation of how effectively Krylov observability $\mathcal{O}_K$ captures the expressivity of the quantum reservoir, as indicated by the IPC, under continuous unitary time evolution defined by $U = \exp(-iHT)$.

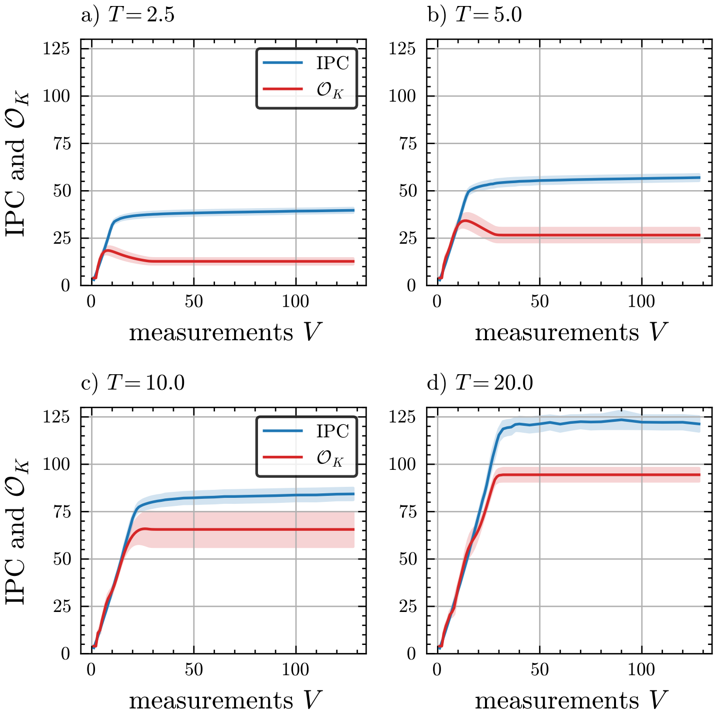


Figure IV.3: Krylov observability $\mathcal{O}_K$ and information processing capacity (IPC), averaged over ten Hamiltonians, are shown as functions of the number of measurements V, shown for different clock cycles: $T = 2.5$ (a), $T = 5.0$ (b), $T = 10.0$ (c), and $T = 20.0$ (d). These plots illustrate how both metrics evolve with increasing V.

In all cases, an almost identical behavior between IPC and $\mathcal{O}_K$ is observed, where both increase and then saturate. This is quite surprising, as IPC represents a black-box approach that quantifies how well data can be projected onto the Legendre basis, while Krylov observability $\mathcal{O}_K$ quantifies an effective dimension at each clock cycle T. In the case of $V = 30$ (Figure IV.2d), a more noticeable deviation is observed; however, the overall trend and saturation values still remain close.

The second aspect that Krylov observability $\mathcal{O}_K$ must capture is which observables are measured and how often they are measured. Our analysis initially focuses solely on the number of measurements V. Figure IV.3 shows both the information processing capacity IPC and Krylov observability $\mathcal{O}_K$ as functions of V for clock cycles $T = 2.5$ (a), $T = 5$ (b), $T = 10$ (c), and $T = 20$ (d). Note that IPC increases and then saturates for all values of T, whereas $\mathcal{O}_K$ exhibits a small peak when $T = 2.5$ and $T = 5.0$.

For simplicity, assume a single observable is measured V_1 and V_2 times when constructing the state matrix, so that for each input the output vector is

$$S_i = \Big(\langle O(\tfrac{T}{V_j})\rangle, \langle O(\tfrac{2T}{V_j})\rangle, \ldots, \langle O(T)\rangle\Big), \quad j = 1, 2. \tag{IV.10}$$

If

$$\mathrm{IPC}(V_2) = \mathrm{IPC}(V_1) \quad \text{even though} \quad V_2 > V_1,$$

then the reservoir with V_1 measurements already captures all of the information contained in $\langle O(t)\rangle$ for $t \in [0, T]$. Further increasing the number of measurements yields expectation values that are linearly dependent, and so IPC remains unchanged.

The time evolution of the observable is then

$$O(t) = e^{iHt}\, O\, e^{-iHt} = \sum_{k=0}^{\infty} \frac{(i\tau)^k}{k!}\, \mathcal{L}^k(H, O(t)) \approx \sum_{k=0}^{n_T} \frac{(i\tau)^k}{k!}\, \mathcal{L}^k(H, O(t)), \tag{IV.11}$$

where the series is truncated at some order n_T depending on the clock cycle. Assume this map yields N_T linearly independent measurements. If V_1 already samples all N_T independent modes, then increasing to $V_2 > V_1$ measurements will not improve task performance, as observed. This argument explains the saturation of IPC at larger values of V.

In the case of Krylov observability, the reasoning is slightly different. We consider the sequence of time evolved observables

$$\Big(O(\tfrac{T}{V_j}), O(\tfrac{2T}{V_j}), \ldots, O(T)\Big) \tag{IV.12}$$

when computing $\mathcal{O}_K$. We compute the normalized fidelity between two neighboring elements with $\tau_j = T/V_j$ as

$$O(t_n) = O\Big(n\tau_1\Big), \quad O(t_{n+1}) = O\Big((n+1)\tau_1\Big),,$$

to quantify their linear independence, and then sum over all contributions. If for some V_1 and $V_2 > V_1$ one finds

$$\sum_{n=1}^{V_1} \Big[1 - \mathrm{F}\Big(O(n\tau_1), O((n+1)\tau_1)\Big)\Big] = \sum_{n=1}^{V_2} \Big[1 - \mathrm{F}\Big(O(n\tau_2), O((n+1)\tau_2)\Big)\Big], \tag{IV.13}$$

then this implies a linear relation of the form

$$O(t+\tau) = O(t) + \tau\, L\Big(O(t)\Big), \tag{IV.14}$$

where L is a linear map acting on $O(t)$ and $\tau \in [0, \max(\tau_1], \tau_2)]$ must hold. For sufficiently small τ, the series expansion yields

$$O(t+\tau) = \sum_{k=0}^{\infty} \frac{(i\tau)^k}{k!} \mathcal{L}^k(H,\, O(t)) \approx O(t) + i\tau\,[H,\, O(t)] + \mathcal{O}(\tau^2), \tag{IV.15}$$

where $[H,\, O(t)]$ is the commutator. For small τ, the term of order τ^2 is negligible, and since $\mathcal{L}$ is linear in $O(t)$, the desired linear relation is recovered. This explains the saturation of Krylov observability $\mathcal{O}_K$ and information processing capacity.

IV.2.3 Quantum Reservoir Computing Maps Data onto the Krylov Space

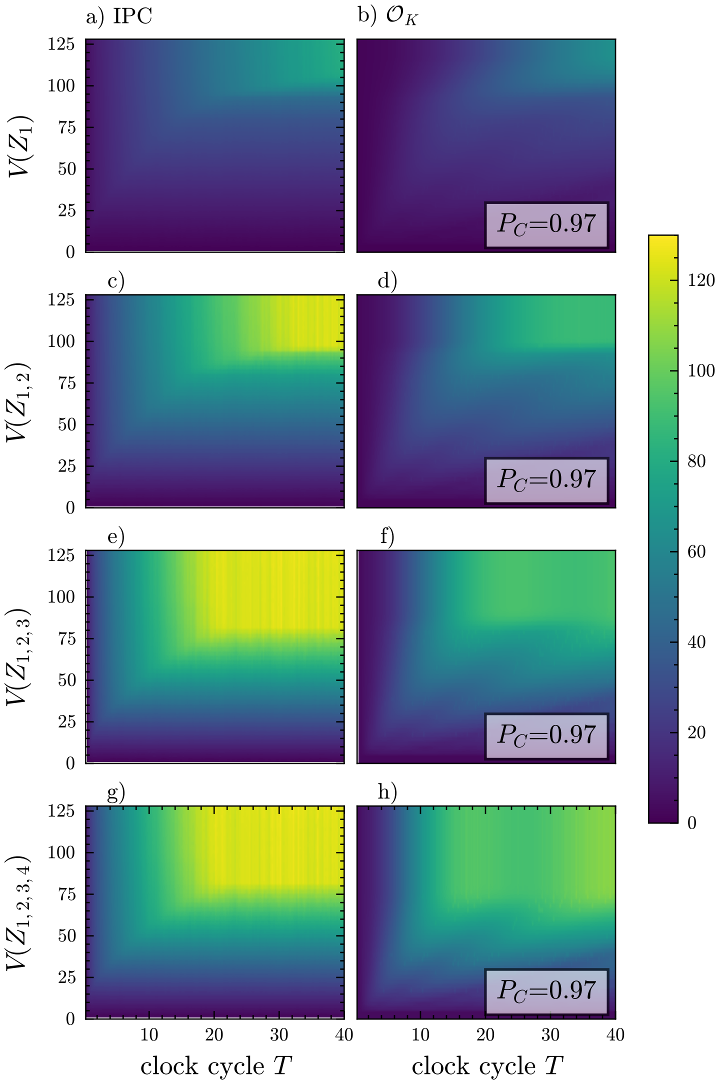


Figure IV.4: Krylov observability $\mathcal{O}_K$ and information processing capacity (IPC), averaged over ten Hamiltonians, are shown as functions of multiplexing V and clock cycle T. The rows correspond to cases where one (first row), two (second row), three (third row), and four (fourth row) sites are measured in the σ_z basis for the computation of IPC and Krylov observability.

Examining individual clock cycles T and numbers of measurements V provides valuable insights and reveals notable behavior. Initially, the saturation in task performance is explained. As a subsequent step, it is relevant to investigate how Krylov observability adapts to different sets of observables. In this context, Figure IV.4 presents the information processing capacity IPC (first column) and Krylov observability $\mathcal{O}_K$ (second column) as functions of the clock cycle T and the number of measurements V, when measuring one site (first row), two sites (second row), three sites (third row), and all sites (fourth row).

To evaluate the capability of Krylov observability $\mathcal{O}_K$ in quantifying the computational power of a quantum system, a comparison is made with the information processing capacity (IPC). The analysis considers ten different Ising Hamiltonians, each consisting of four sites and uniformly distributed coupling constants $J_{i,j}$ within the interval $[0.25, 0.75]$.

Figure IV.4 shows the IPC (a, c, e, g) and $\mathcal{O}_K$ (b, d, f, h) as functions of the number of measurements V and the clock cycle T. Each row corresponds to a different number of measured observables, where the first, second, third, and fourth row represent measurements on one (Z_1), two ($Z_{1,2}$), three ($Z_{1,2,3}$), and four ($Z_{1,2,3,4}$) sites, respectively. An increase in T and V results in an increase in the information processing capacity IPC, until saturation is reached at $T_{\text{sat}}(Z_1) \approx 36$ and $V_{\text{sat}}(Z_1) \approx 110$. This trend closely matches the behavior of Krylov observability $\mathcal{O}_K$ (Figure IV.4.b). The saturation of $\mathcal{O}_K$ suggests that increasing the number of measurements V beyond a certain point no longer yields linearly independent expectation values. Since linearly dependent outputs do not enhance the computational power of the reservoir, saturation of the information processing capacity follows. To quantify the similarity between IPC and $\mathcal{O}_K$, the Pearson correlation factor is computed for each case:

$$P_C = \frac{\text{Cov}(\text{IPC}, \mathcal{O}_K)}{\sigma(\text{IPC})\sigma(\mathcal{O}_K)} \tag{IV.16}$$

For the case of one measured observable (first row), a Pearson correlation factor of $P_C = 0.97$ is obtained, indicating an almost identical relationship between IPC and $\mathcal{O}_K$ (Figure IV.4.a and Figure IV.4.b).

Figure IV.4.c–d illustrate the behavior of IPC and Krylov observability when two sites are measured. The saturation thresholds shift to $T_{\text{sat}}(Z_{1,2}) \approx 30$ and $V_{\text{sat}}(Z_{1,2}) \approx 95$, with a Pearson correlation factor of $P_C = 0.97$, again showing close agreement. Qualitatively, the region of optimal performance moves toward lower values of both T and V compared to the single-site measurement case. Extending the analysis to three measured sites (Figure IV.4.e–f) continues this trend, with saturation occurring at $T_{\text{sat}}(Z_{1,2,3}) \approx 20$ and $V_{\text{sat}}(Z_{1,2,3}) \approx 80$, which aligns closely with the behavior observed when all four sites are measured (Figure IV.4.g–h).

In all cases, Krylov observability accurately captures the expressive power of the data, as evidenced by consistently high correlation values of $P_C = 0.97$. Krylov observability offers a way to capture the data generalization capabilities of quantum systems. Yet fully grasping its significance requires more than formal analysis. One must develop a physical interpretation of the observed effects, integrate these findings with existing theories, and situate them within the broader context of the field. We now address these challenges.

Krylov Observability Quantifies Data Expressivity in Quantum Reservoir Computing

The results indicate that Krylov observability effectively captures the generalization capabilities of a quantum reservoir, as evidenced by its strong agreement with the information processing capacity (IPC), showing correlation factors of $\mathcal{P}_C = 0.97$. This supports the interpretation of quantum reservoir computing as a mapping of classical data onto Krylov spaces, offering both a theoretical framework for understanding quantum reservoirs and a practical, physically grounded tool based on quantum information measures.

An additional advantage of Krylov observability is its computational efficiency. Even with unoptimized code, the computation takes only a few seconds, which is four orders of magnitude faster than highly optimized code for calculating the information processing capacity. Moreover, Krylov observability is not a black box technique. It offers a physically interpretable framework that reveals the structure of the spaces through which measured observables evolve.

Quantum Reservoir Computing Showcases the Utility of Krylov-Based Information Measures

Quantum reservoir computing tests the physical relevance of Krylov spaces in quantum mechanics by evaluating the system's ability to generalize macroscopic data encoded into the system externally. This ability is quantified by the information processing capacity IPC. The central question is how well a measure defined on the Krylov space, namely Krylov observability $\mathcal{O}_K$, can predict a quantum system's ability to retain and process macroscopic data such as encoded input sequences, as evaluated by the information processing capacity. Another question is whether insights from Krylov spaces can apply to other areas such as quantum machine learning. The almost identical behavior of Krylov observability and the information processing capacity shows that Krylov complexity is an effective and insightful metric for studying time evolution in quantum systems.

The Information Processing Capacity Captures Quantum Evolution

From a quantum information perspective, a quantum reservoir computer encodes macroscopic data into quantum states. A central question is how effectively a quantum mechanical system can perform complex operations on this encoded data or act as a surrogate for another physical system, such as the Lorenz attractor. In this context, the information processing capacity (IPC) is especially important. The linear component, or memory capacity IPC_1, quantifies how well the system can reconstruct past inputs. If the input u is encoded at time t, yielding the quantum state $\rho(u,t)$, the question is: up to what time T can the macroscopic data u be reliably recovered?

This question mirrors those in quantum dynamics, where measures such as Krylov complexity, out of time order correlators (OTOCs), and Loschmidt echoes estimate how long information about the initial state remains accessible. The time after which this information becomes unrecoverable is known as the scrambling time.

Higher order components of the IPC, for example IPC_2, also have clear physical interpretations. The term IPC_2 measures how well the system can process pairs of inputs (u_i, u_j) encoded at times (t_i, t_j). First, u_1 is encoded into the quantum state, starting from an initial state ρ_0:

$$\rho(u_1) = U_E\Big(u_1, \rho_0\Big). \tag{IV.17}$$

Then the system evolves for time t_2 before the second input is encoded:

$$\rho\Big(t_2, u_1\Big) = e^{-iHt_2}\, \rho(u_1)\, e^{iHt_2}, \tag{IV.18}$$

$$\rho\Big(t_2, u_1, u_2\Big) = U_E\Big(u_2, \rho(t_2, u_1)\Big). \tag{IV.19}$$

Finally, the system evolves for an additional time t_3:

$$\rho\Big(t_3, u_1, u_2\Big) = e^{-iHt_3}\, \rho\Big(t_2, u_1, u_2\Big)\, e^{iHt_3}. \tag{IV.20}$$

Reconstructing a target function such as $y^{\mathrm{targ}} = u_1\, u_2$ at time t_3 reveals how the system correlates.

Although this view may seem interpretive, Figure IV.4 shows that the connection is direct. Krylov observability was devised to quantify an effective dimensionality for a chosen set of observables, not specifically to explain the IPC. In the quantum regime, one can set the number of measurements V equal to the rank M of the Krylov space $\mathcal{F}$. The fact that the IPC closely follows the behavior of $\mathcal{O}_K$ indicates that the IPC is a systematic and powerful tool for understanding quantum systems. A natural direction for future work would be to investigate whether OTOCs can be interpreted as components of the information processing capacity. This analysis lies beyond the scope of the present thesis but offers a promising avenue for further research.

IV.3 Engineering Quantum Reservoir Computers

In the previous section, an analysis based on random Hamiltonians was conducted and then the Krylov based information measures were applied in that regard. This is consisted in most literature, where the error is averaged over many realizations of some Hamiltonian. The following section focuses on evaluating Krylov-based information measures in the context of unique Hamiltonians. This evaluation aims to determine whether certain characteristics are better captured by Krylov observability or Krylov expressivity, and whether these measures can serve as efficient pre-processing tools for selecting suitable quantum systems for computational tasks.

As before, transverse field Ising Hamiltonians with varying inter-spin couplings are employed as quantum reservoirs. These systems are analyzed with respect to spread complexity, fidelity, Krylov observability, Krylov expressivity, and task performance. The Hamiltonian is defined as:

$$H_{I\alpha} = \sum_{i=1,j>i}^{N_S} J_{ij} X_i X_j + \sum_{i=1}^{N_S} h Z_i, \tag{IV.21}$$

where $\alpha \in \{1, 2, 3, 4\}$ indexes the four different Hamiltonians, N_S is the number of qubits, and X_i, Y_i, Z_i denote the Pauli-x, y, and z matrices acting on the ith qubit, defined as:

$$X_i, Y_i, Z_i = \left(\bigotimes_{k=1}^{i-1} I_2 \right) \otimes \sigma_{x,y,z} \otimes \left(\bigotimes_{k=i+1}^{N_S} I_2 \right), \tag{IV.22}$$

with $h = 0.5$ fixed for all Hamiltonians. The number of pairwise distinct eigenvalues d, along with the inter-spin coupling values J_{ij}, are listed in Table IV.4.

	H_{I1}	H_{I2}	H_{I3}	H_{I4}
# of eigenvalues d	9	16	15	16
$J_{1,2}$	0.5	0.4	0.35	0.35
$J_{1,3}$	0.5	0.5	0.4	0.4
$J_{1,4}$	0.5	0.5	0.45	0.45
$J_{2,3}$	0.5	0.5	0.5	0.5
$J_{2,4}$	0.5	0.5	0.55	0.55
$J_{3,4}$	0.5	0.5	0.6	0.65

Table IV.4: The number of pairwise distinct eigenvalues d and the inter-spin couplings J_{ij} for the different Ising reservoirs.

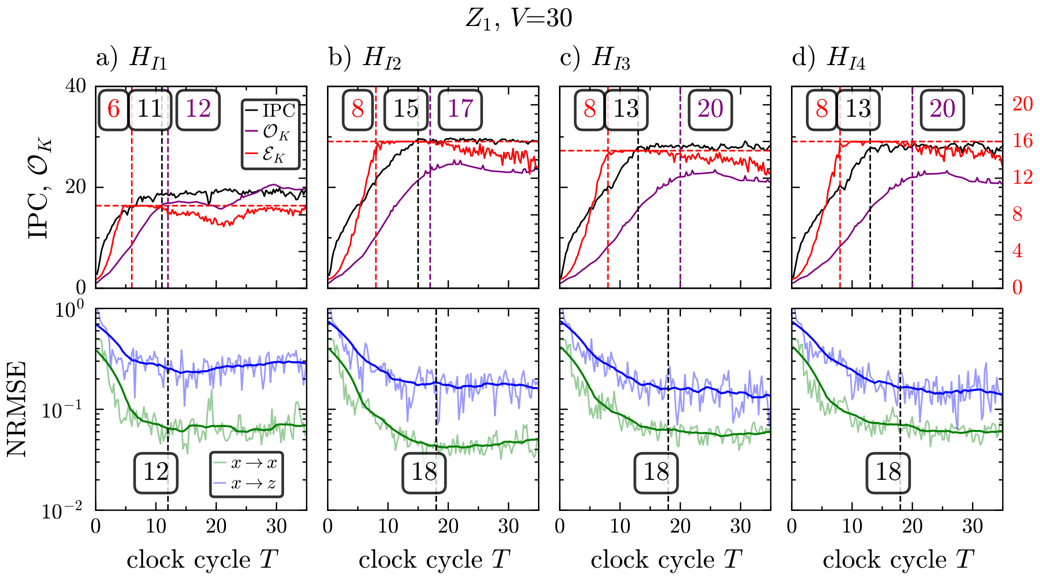


Figure IV.5: The first row displays the information processing capacity (IPC, black), Krylov expressivity $\mathcal{E}_K$ (red), and Krylov observability $\mathcal{O}_K$ (purple). The right axis indicates the value of Krylov expressivity $\mathcal{E}_K$. The second row shows Lorenz task performance: the five-step-ahead prediction ($\Delta t = 0.1$) of the x-variable (green) and the cross-prediction of the z-variable (blue), both as functions of the clock cycle T for the quantum reservoirs **a)** H_{I1}, **b)** H_{I2}, **c)** H_{I3}, and **d)** H_{I4}. The state matrix is constructed by measuring the Z_1 observable $V = 30$ times. The bold lines of the NRMSE in the second row represent the first-order Savitzky–Golay filtered NRMSE, used to better illustrate the saturation in task performance (computed using `scipy.signal.savgol_filter`).

Figure IV.5 presents results for the case in which only the first observable Z_1 is measured, while Figure IV.6 displays the outcomes when all four observables Z_1, Z_2, Z_3, and Z_4 are included. In all experiments, the observables are time-multiplexed $V = 30$ times.

The Krylov expressivity $\mathcal{E}_K$ for the Hamiltonians H_{I1}, H_{I2}, H_{I3}, and H_{I4} is bounded above by 9, 16, 15, and 16, respectively, corresponding to the dimension of the Krylov space, or equivalently, the number of distinct eigenvalues[81].

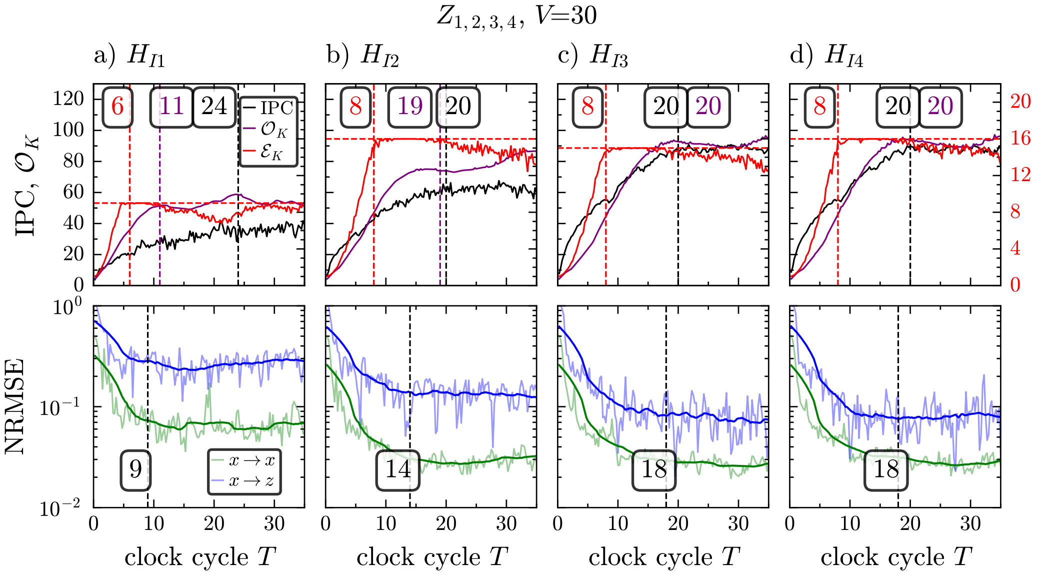


Figure IV.6: As in Figure IV.5, but with the state matrix constructed through measurement of all four sites.

Table IV.5 lists the NRMSE values at saturation for the Lorenz task, comparing scenarios in which only Z_1 is measured versus all four sites. For H_{I1}, task performance remains unchanged regardless of the number of measured observables. In the case of H_{I2}, a slight reduction in error is observed when all sites are included. Hamiltonians H_{I3} and H_{I4} exhibit identical performance, with errors reduced by half when additional observables are incorporated.

Figure IV.5.a illustrates the behavior for H_{I1}, where the normalized root mean square error (NRMSE) for the Lorenz task, the information processing capacity (IPC), and Krylov observability $\mathcal{O}_K$ all saturate around $T = 12$. In contrast, the Krylov expressivity $\mathcal{E}_K$ reaches saturation earlier, at $T = 6$. A modification to one of the inter-spin couplings in H_{I1} yields H_{I2}, which results in increased values for IPC, $\mathcal{E}_K$, and $\mathcal{O}_K$, along with improved task performance (Figure IV.5.b).

The results for H_{I3}, constructed using evenly spaced but distinct inter-spin couplings, are shown in Figure IV.5.c. In this configuration, $\mathcal{E}_K$, IPC, and $\mathcal{O}_K$ saturate at approximately $T = 8$, $T = 13$, and $T = 20$, respectively. Task performance saturates at around $T = 18$, closely aligning with the behavior of $\mathcal{O}_K$, suggesting that Krylov observability is the most accurate performance predictor in this case.

Taking a closer look at IPC and $\mathcal{O}_K$ reveals similar behavior for H_{I1} and H_{I2}, whereas for H_{I3}, IPC and $\mathcal{O}_K$ appear to behave differently. This discrepancy is attributed to IPC being upper-bounded by the readout dimension $N_R = 30$, i.e., IPC $\leq N_R$, leading to premature saturation. In contrast, $\mathcal{O}_K$ continues to track the system's dynamics more effectively, making it a better indicator of task performance in cases where the reservoir is undersampled.

Next, a closer look at Krylov expressivity and observability is taken. In the case of H_{I1}, the lowest performance for both Lorenz tasks is explained by consistently lower values across all three metrics: IPC, $\mathcal{E}_K$, and $\mathcal{O}_K$.

Comparing H_{I2} and H_{I3} reveals that $\mathcal{E}_K(H_{I3}) < \mathcal{E}_K(H_{I2})$ and $\mathcal{O}_K(H_{I3}) > \mathcal{O}_K(H_{I2})$. Furthermore, H_{I3} outperforms H_{I2} in cross-prediction tasks, while H_{I2} shows slightly better results for predicting the x-component. To probe this further H_{I4} is introduced. H_{I4} is nearly identical to H_{I3} except for one modified inter-spin coupling, increasing its Krylov expressivity to match that of H_{I2}, i.e., $\mathcal{E}_K(H_{I4}) = \mathcal{E}_K(H_{I2})$. Despite the enhanced expressivity, the Lorenz task performance, IPC, and $\mathcal{O}_K$ of H_{I4} remain almost identical to those of H_{I3}.

Next, the typical QRC scheme will be employed, to test Krylov expressivity and observability. In that regard, Figure IV.6 considers a typical quantum reservoir configuration in which all observables are measured, resulting in a readout dimension of $N_R = 120$. Under this configuration, H_{I1} continues to exhibit the lowest task performance among the tested Hamiltonians. Although the Lorenz task performance remains unchanged, the information processing capacity increases from IPC $= 20$ in Figure IV.5.a to IPC $= 40$ in Figure IV.6.a.

For H_{I2}, H_{I3}, and H_{I4}, as shown in Figure IV.6.b–d, both IPC and $\mathcal{O}_K$ provide consistent and reliable explanations for the saturation behavior of the Lorenz task NRMSE. The improved performance of IPC in Figure IV.6 compared to Figure IV.5 is attributed to the increased number of readout nodes, with $N_R = 120$. In this case, IPC is upper-bounded by IPC $\leq N_R = 120$, which is sufficiently large to avoid saturation effects that could otherwise limit the observed capacity. The results for other configurations of quantum reservoirs are shown in Appendix B.

Hamiltonian	Z_1 $x \to x$ $\mathrm{NRMSE}_{\mathrm{sat}}$	All sites $x \to x$ $\mathrm{NRMSE}_{\mathrm{sat}}$	Z_1 $x \to z$ $\mathrm{NRMSE}_{\mathrm{sat}}$	All sites $x \to z$ $\mathrm{NRMSE}_{\mathrm{sat}}$
H_{I1}	0.08	0.08	0.30	0.30
H_{I2}	0.04	0.03	0.20	0.15
H_{I3}	0.06	0.03	0.16	0.08
H_{I4}	0.06	0.03	0.16	0.08

Table IV.5: Saturation values of NRMSE for the Lorenz task for Hamiltonians H_{I1} to H_{I4}, comparing single-site measurement Z_1 (rows one and three) to measurement of all sites (rows two and four).

To explore the differences in expressivity and observability between H_{I2} and H_{I3}, task performance under full measurement is examined in Figure IV.6. The cross-prediction task for H_{I3} results in an error of $\mathrm{NRMSE} = 0.08$, while H_{I2} yields a higher error of $\mathrm{NRMSE} = 0.15$. Notably, the Krylov expressivity of H_{I2}, with $\mathcal{E}_K(H_{I2}) = 16$, exceeds that of H_{I3}, which has $\mathcal{E}_K(H_{I3}) = 15$. This was observed in the case, when the system was undersampled, but is now seen in the typical QRC scheme. One would assume that, if the Krylov space onto which data is mapped is larger, then better task performance is achieved. However, the ability to access this information depends on the measurement process. In this instance, the number of linearly independent measurements in H_{I2} is lower than in H_{I3}, as indicated by the reduced Krylov observability: $\mathcal{O}_K(H_{I2}) = 60$ versus $\mathcal{O}_K(H_{I3}) = 85$. This discrepancy likely results from partial commutation between the measured observables and the time-evolution under H_{I2}. The use of pairwise distinct inter-spin couplings in H_{I3} appears to allow the observables to probe a larger portion of the quantum state, thereby enhancing overall observability.

To examine this effect in greater detail, both IPC and $\mathcal{O}_K$ are evaluated across varying clock cycles T, numbers of virtual nodes V, and measurement configurations for H_{I2} (Figure IV.7) and H_{I3} (Figure IV.8). The first and second rows display results based on measurements from one and two observables, respectively, while the third row corresponds to measurements of all four observables.

For H_{I2}, the observed values of IPC and $\mathcal{O}_K$ do not exceed 85. In contrast, H_{I3} reaches a maximum value of 105, suggesting that the increased IPC is a result of improved Krylov observability.

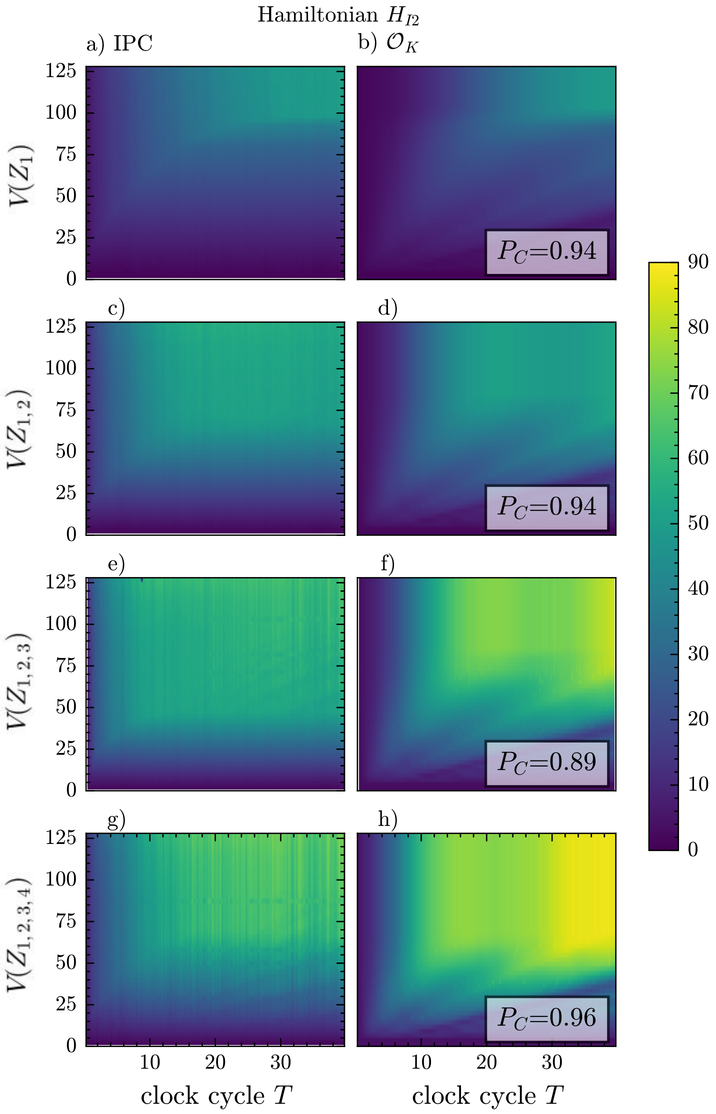


Figure IV.7: Observability $\mathcal{O}_K$ (right column) and IPC (left column) color-coded from 0 to 90, depending on the number of measurements V and the clock cycle T for H_{I2}. The rows correspond to cases where one (first row), two (second row), three (third row), and four (fourth row) sites are measured in the σ_z basis for the computation of IPC and Krylov observability. The Pearson correlation coefficient P_C is calculated between the two images in each row, indicating a strong correlation between $\mathcal{O}_K$ and IPC.

Additionally, increasing the number of measured observables shifts the optimal performance region. This shift reflects a transition from requiring longer clock cycles and more measurements to achieving similar performance with shorter cycles and fewer measurements. Each plot includes the Pearson correlation coefficient P_C between IPC and $\mathcal{O}_K$. For H_{I2}, the correlation is consistently strong: $P_C = 0.94$ for one or two observables and $P_C = 0.96$ when all four are measured. For H_{I3}, the corresponding coefficients are $P_C = 0.91$ for one observable and $P_C = 0.95$ for two or four observables.

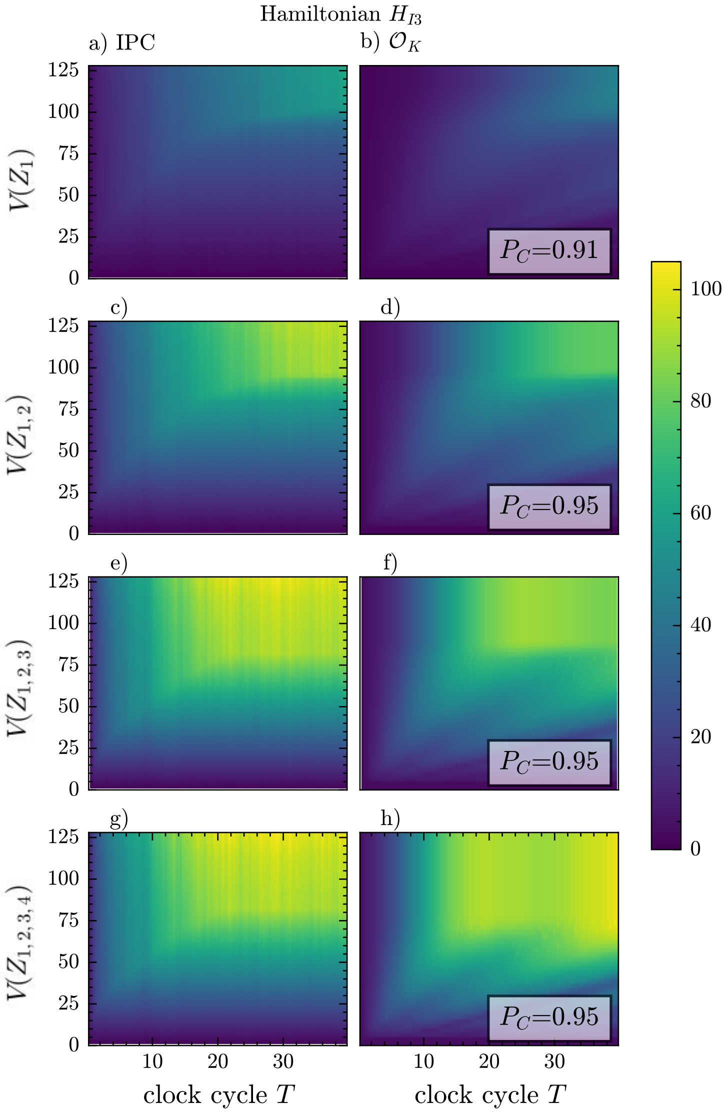


Figure IV.8: Same as in Figure IV.7, shown here for H_{I3}.

Conclusion

Krylov-based information measures are analyzed to evaluate their ability to explain and predict task performance in quantum reservoir computing, specifically applied to chaotic time-series prediction using the Lorenz63 system.

The analysis focuses on Krylov expressivity $\mathcal{E}_K$ [81] and Krylov observability $\mathcal{O}_K$ [82], both of which exhibit saturation behavior at larger clock cycles. These measures are compared to the information processing capacity (IPC), which quantifies the system's ability to retain and non-linearly transform input data. Four quantum reservoirs are evaluated in an undersampled regime, revealing that Krylov observability $\mathcal{O}_K$ more effectively correlates with task performance than IPC. In such scenarios, IPC reaches its upper bound while the prediction error for the Lorenz tasks continues to decrease with increasing clock cycle duration T.

Further simulations involve scenarios in which all qubit sites are measured, as illustrated in Figure IV.6. Results indicate that IPC and $\mathcal{O}_K$ display nearly identical behavior under full observability. Additionally, H_{I3} demonstrates superior Lorenz task performance and higher IPC compared to H_{I2}, despite having lower Krylov expressivity. Although Krylov expressivity provides insights into the mapping of input data onto the system's state space, the volume of information extracted from this space is smaller in the case of H_{I2}, as evidenced by its reduced Krylov observability.

Krylov expressivity provides a framework for analyzing how input states are projected into the Krylov space, offering valuable insight into input encoding strategies. While this concept may have limited applicability in quantum reservoir computing—where system dynamics are governed by temporal evolution—it holds significant potential in variational quantum machine learning and quantum extreme learning machines, where input encoding plays a central role [87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97]. In such frameworks, inputs x_i may be encoded into a quantum system via a unitary transformation $U_E(x_i)$, yielding a set of initial states defined by $|x\rangle \in \{U_E(x_i)\,|s\rangle \mid x_i \in \mathcal{X}\}$, where $\mathcal{X}$ denotes the input set. Krylov expressivity can then be utilized to evaluate and compare different encoding strategies in variational quantum algorithms and quantum extreme learning models.

Existing research on the explainability and expressivity of quantum machine learning remains limited, with most approaches adapting techniques from classical machine learning [98, 99, 100, 101]. A notable exception is [102], which investigates the development of explanation methods tailored to parameterized quantum circuits. However, many of these methods lack physical interpretability and are often data-driven or resemble black-box models, similar to the information processing capacity (IPC). In contrast, Krylov observability and expressivity provide interpretable insights into the structure of the underlying Hilbert space and are independent of the input data. This physical interpretability may contribute to a more comprehensive understanding of phenomena in quantum machine learning, such as barren plateaus, and offers a meaningful quantum-mechanical measure for examining system dynamics.

Finally, the superior performance of Krylov observability and expressivity, compared to Krylov complexity in the context of quantum reservoir computing, highlights their potential as effective tools for advancing research in areas where Krylov complexity has traditionally been applied. These include quantum chaos, driven and open quantum systems, integrability, and SYK models, among others [14, 16].

IV.4 Memory-Restricted Quantum Reservoir Computing

The previous chapter discussed how Krylov-based information measures can be used to understand the generalization capabilities of quantum reservoir computing. In that context, the standard quantum reservoir computing scheme was employed, specifically the one explored with the Ising model in [26, 27, 28, 29, 30, 31, 32, 33]. Early implementations of quantum reservoir computing were carried out on IBM's quantum processor [34, 35, 36, 37]. One particularly interesting aspect is that quantum reservoir computing appears capable of leveraging dissipation and noise as computational resources—whereas, in contrast, these are major challenges in variational quantum machine learning [38, 39, 40].

However, a key challenge in quantum reservoir computing remains largely unaddressed: the issue of measurement and the associated collapse of the quantum state. In quantum mechanics, the measurement of a quantum state $|\phi\rangle$ with respect to an observable O, defined by the eigenvalue equation $O\,|o_i\rangle = o_i\,|o_i\rangle$, results in the collapse

$$|\phi\rangle \xrightarrow{\text{measurement}} |o_i\rangle\,, \quad \text{with probability} \quad p_i = |\langle o_i|\phi|o_i|\phi\rangle|^2 \tag{IV.23}$$

where the measurement outcome is the eigenvalue o_i, occurring with probability p_i. This collapse causes a loss of information about the original state.

In the first quantum reservoir computing scheme proposed by the authors of [26], the state matrix is constructed through successive measurements of the evolved system. This approach necessitates re-initializing the system with the complete input time series up to the current time step. The procedure is outlined in Algorithm 3.

The algorithm takes as input a unitary operator U, a set of observables $\{O_\alpha\}_\alpha$ to be measured, an initial quantum state $|\Psi\rangle$, a time-series input $\mathbf{u} = \{u_i\}_{i=1}^{M}$ of length M, and an encoding scheme that maps each input value to a quantum state $|\Psi_\epsilon(u_i)\rangle$. At each time step i, the reservoir is reset to $|\Psi\rangle$ and sequentially fed all inputs up to u_i. Once the system has evolved through these inputs, the observables $\{O_\alpha\}_\alpha$ are measured.

$$T_1(M) = \sum_{i=1}^{M} i = \frac{M(M+1)}{2} \in \mathcal{O}(M^2). \tag{IV.24}$$

This quadratic scaling renders the method impractical for very long or continuous input sequences ($M \to \infty$). As a result, much of the existing research in quantum reservoir computing for time-series analysis focuses on understanding reservoir behavior under this computationally intensive scheme.

Algorithm 3 Typical Quadratic complexity quantum algorithm (QCQA).

```
Require: $U$, $|\Psi\rangle$, $|\Psi_\epsilon(u_m)\rangle$, $\{O_\alpha\}$
Require: $u \leftarrow [u_1, u_2, \ldots, u_M]$                         ▷ Input series
$i \leftarrow 1$
while $i \leq M$ do
    $\rho(0) \leftarrow |\Psi\rangle\langle\Psi|$
    $j \leftarrow 1$
    while $j \leq i$ do
        $\hat{\rho}(j) \leftarrow |\Psi_\epsilon(u_j)\rangle\langle\Psi_\epsilon(u_j)| \otimes \mathrm{Tr}_1[\rho(j-1)]$
        $\rho(j) \leftarrow U\hat{\rho}(j)U^\dagger$
        $j \leftarrow j + 1$
    end while
    $S_{i,\alpha} \leftarrow \mathrm{Tr}[\rho(i)O_\alpha]$                      ▷ Collapse of state
    $i \leftarrow i + 1$
end while
return $S$
```

To address the time complexity for time-series tasks, weak measurements [41, 42] and reinitialization schemes [103, 43] have been proposed. Other approaches include reintroducing measured outputs [44, 45] and continuous measurements [46].

The key insight explored in this section is that introducing artificial memory restriction in quantum reservoir computing can not only reproduce the performance of classical approaches, but can even enhance it. This improvement is attributed to the ability to tune the system's sensitivity to nonlinear features in the input data. These findings build upon the work presented in

- *Enhancing the performance of quantum reservoir computing and solving the time-complexity problem by artificial memory restriction*, S. Čindrak, B. Donvil, K. Lüdge, and L. C. Jaurigue, Phys. Rev. Res. 6, 013051 (2024).

This work proposes an alternative scheme for reservoir computing, based on the commonly assumed *fading memory property* of reservoir systems [19, 20], along with the well-established trade-off between memory capacity and nonlinearity [52, 104, 54].

The fading memory property describes the tendency of a reservoir to gradually lose sensitivity to inputs from the distant past. This property can be quantified via the linear contribution to the information processing capacity, denoted as IPC_1. The information processing capacity, which generalizes the concept of linear memory capacity [51], captures the reservoir's ability to perform nonlinear transformations of past inputs and has been used to predict performance on various tasks [105, 106]. In the proposed scheme, only the most recent n inputs, i.e., $(u_{i-n+1}, u_{i-n+2}, \ldots, u_i)$, are injected into the system at the t_i-th time step, where u_i corresponds to the input at time step t_i. The memory window n may be considerably shorter than the full fading memory length of the reservoir. Consequently, each input requires n unitary operations, resulting in a total computational cost of

$$T_2(M) = nM \in \mathcal{O}(M), \tag{IV.25}$$

which scales linearly with the length of the time series M. This linear scaling makes the approach suitable for large or continuous input sequences ($M \to \infty$), as only n unitary operations are needed per input, enabling real-time computation. The linear-time algorithm is formally described in Algorithm 4.

Algorithm 4 Linear complexity quantum algorithm (LCQA).

```
Require: $U$, $|\Psi\rangle$, $|\Psi_\epsilon(u_m)\rangle$, $\{O_\alpha\}$
Require: $u \leftarrow [u_1, u_2, \ldots, u_M]$                              ▷ Input series
$i \leftarrow 1$
while $i \leq M$ do
    $\rho(0) \leftarrow |\Psi\rangle \langle\Psi|$
    $k \leftarrow 0$
    while $k \leq n$ do
        $j \leftarrow i - n + k$
        $k \leftarrow k + 1$
        $\hat{\rho}(j) \leftarrow |\Psi_\epsilon(u_j)\rangle \langle\Psi_\epsilon(u_j)| \otimes \mathrm{Tr}_1[\rho(j-1)]$
        $\rho(j) \leftarrow U\hat{\rho}(j)U^\dagger$
    end while
    $S_{i,\alpha} \leftarrow \mathrm{Tr}[\rho(i)O_\alpha]$                      ▷ Collapse of state
    $i \leftarrow i + 1$
end while
return $S$
```

Figure IV.9.d shows the linear memory capacity as a function of steps into the past. When the system is reset using the complete input history—corresponding to the classical scheme (QCQA)—the recall capability of the Ising model reservoir declines to zero after approximately $d_F = 15$ steps into the past, which indicates the fading memory limit of the reservoir. This suggests that restricting the system using the LCQA algorithm should result in identical performance if the reservoir is limited to this many past inputs, i.e., $n = d_F$.

Since, in classical reservoir computing, data expressivity is quantified by the information processing capacity (IPC), it is meaningful to examine how IPC values change as a function of the reset length n. This provides a more quantitative understanding than evaluating performance across multiple specific tasks. The first- to third-order information processing capacities (IPC_1–IPC_3) are shown in Figure IV.9.a, while Figure IV.9.b displays the fourth- to sixth-order capacities (IPC_4–IPC_6). Figure IV.9.c shows the total IPC, i.e., $\mathrm{IPC} = \sum_{i=1}^{6} \mathrm{IPC}_i$. In each case, the IPC values approach the QCQA limit (dashed lines) as n increases toward the reservoir's maximum recall depth ($n \to d_F$).

For small values of n, IPC values are lower due to the artificially limited memory of the reservoir. However, within a specific intermediate range of n, both the total IPC and higher-order components (IPC_2–IPC_6) exceed the QCQA benchmark. Notably, the higher-order nonlinear terms IPC_4, IPC_5, and IPC_6 show significant increases: IPC_4 increases by half, while IPC_5 and IPC_6 increase by approximately four- and five-fold, respectively.

This implies that restricting the reset length to $n < d_F$ is not only computationally more efficient but also enhances the nonlinear dynamics of the reservoir. To gain more insight into this effect, in Figure IV.9.d the linear memory in dependence of the reconstructed input d steps into the past is plotted for $n = 6, 10, 15$. Here, it can be seen that the distribution of the linear IPCs is changed compared with the QCQA case. Although the summed linear IPC (IPC_1) is decreased for $n = 6$ (see Figure IV.9.a IPC_1 at $n = 6$), the capacities for the past inputs which can be reconstructed are higher. For $n = 10 < d_F$ the IPC_1 reaches the QCQA limit, as can be seen in Figure IV.9.a. This seems at first surprising, since only ten inputs are considered. However, in Figure IV.9.d the case of $n = 10$ is shown to have an increase around $d = 10$ as well, resulting in the same area under curve and therefor IPC_1 as the QCQA limit. Further increasing to $n = 15$ results in almost identical behavior to the QCQA limit, with again a small increase at $d = 15$.

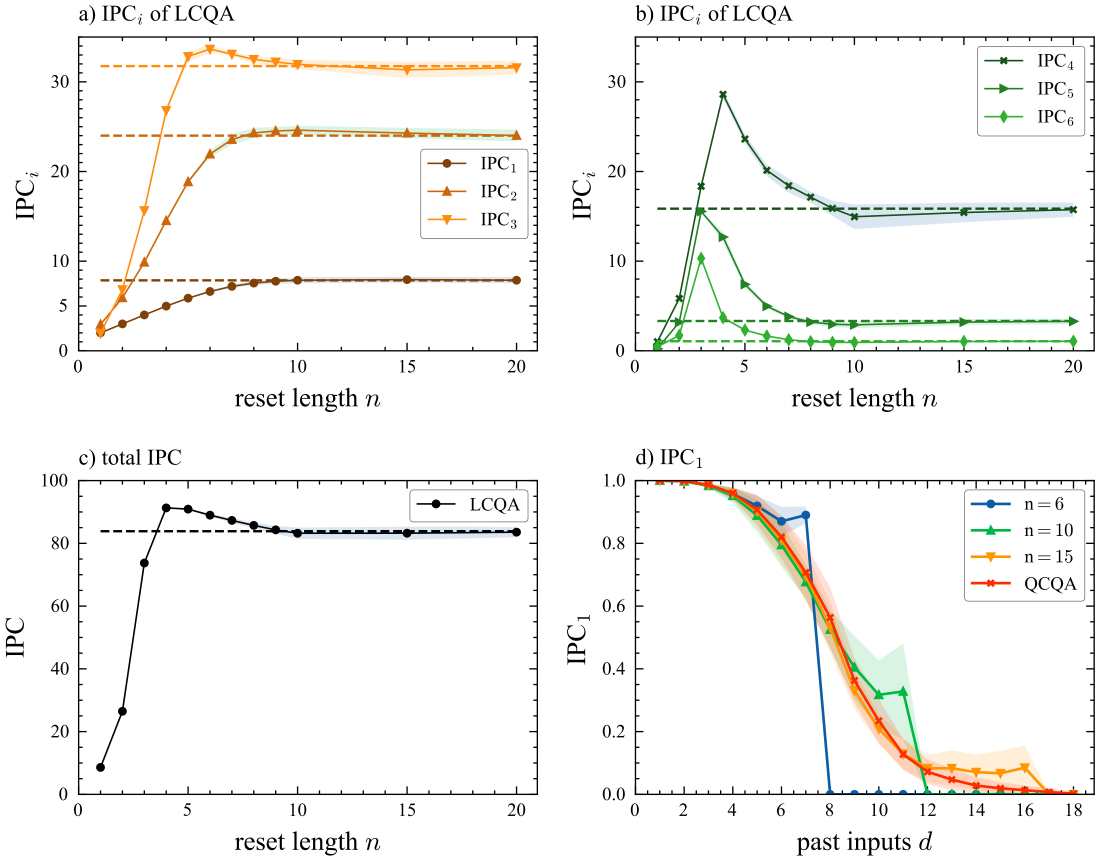


Figure IV.9: Information processing capacities (IPCs) of the Ising model reservoir as a function of the reset length n. a) IPC$_1$–IPC$_3$ (linear to third-order), (b) IPC$_4$–IPC$_6$ (higher-order nonlinearities), .c total IPC across all six orders, and (d) IPC$_1$ as a function of past input delay d, highlighting the fading memory property. The QCQA baseline is shown as dashed lines. As n increases toward the fading memory depth $d_F = 15$, IPC values approach the QCQA limit. At intermediate reset lengths, higher-order IPCs (especially IPC$_4$–IPC$_6$) exceed QCQA values, demonstrating enhanced nonlinear memory and suggesting an optimal regime for performance and efficiency.

Using the LCQA scheme introduces two key changes in the reservoir dynamics that may contribute to improved task performance:

- For small values of n, information about the initial state is still present at the time of measurement. This could indicate that the system is better able to generalize to this type of behavior. In the QCQA limit, by contrast, each new input effectively corresponds to a random initial state, as the current state depends on inputs beyond the fading memory window, and no information about the original state can be recovered.

- In the QCQA limit, the repeated overwriting and re-encoding of multiple inputs leads to highly mixed quantum states. In contrast, the LCQA scheme, with a short reset length n, begins from a pure initial state and does not have sufficient time to become as mixed as in the QCQA scheme.

To analyze this influence, consider using the LCQA scheme with a small reset length $n = 3$. Based on the encoding scheme described in Algorithm 4, define the map

$$U_E(\rho(j-1)) = |\Psi_\epsilon(u_j)\rangle\langle\Psi_\epsilon(u_j)| \otimes \mathrm{Tr}_1[\rho(j-1)]U_R, \tag{IV.26}$$

with the initial state $\rho(0) = |0000\rangle\langle 0000|$. Let the reservoir evolve under the unitary operator $U_R = e^{-iHT}$, so the time evolution of the state is given by

$$\rho(j) = U_R U_E(\rho(j-1))U_R^\dagger. \tag{IV.27}$$

With $n = 3$, the state after three time steps becomes

$$\begin{aligned}\rho(n) &= U_R U_E\left(U_R U_E\left(\rho(n-2)\right)U_R^\dagger\right)U_R^\dagger\\ &= U_R U_E\left(U_R U_E\left(U_R U_E\left(|0000\rangle\langle 0000|\right)U_R^\dagger\right)U_R^\dagger\right)U_R^\dagger.\end{aligned} \tag{IV.28}$$

The key point is that the larger n is, the more frequently the operations U_R and U_E are applied during the system's evolution, leading to greater loss of information about the initial state $|0000\rangle\langle 0000|$. The improved performance observed with a restricted reset length is attributed to the preservation of some information about the initial state, which enables better generalization to new data during training.

To discuss this in the context of the fading memory capacity cutoff d_F, consider the following cases:

- $n > d_F$: This results in $\mathrm{IPC}(d = n) = 0$, i.e., the reset length exceeds the fading memory capacity cutoff. Now, imagine the point in time n steps in the past with input u_{i-n}, when the system was first initialized as $\rho(0) = |0000\rangle\langle 0000|$. This implies that any information encoded into the reservoir at that time (u_{i-n}) will be forgotten after $n > d_F$ steps. If the data is randomly generated or derived from a chaotic time series, then from the perspective of the future, the current encoding will appear to originate from a random initial state. Since new data is encoded at every time step and is uncorrelated with future inputs, the reservoir effectively encounters a random initial state at each point in time. Additionally, this leads to repeated overwriting of pure states, progressively increasing the mixedness of the system. In this case, the quantum state becomes highly mixed over time.

- $n < d_F$: Now consider the opposite case—a small n—where the data is encoded such that $\mathrm{IPC}_1(d = n) > 0$. This implies that some information about the initial state $\rho(0)$ is still retained in the future. As a result, the system exhibits improved training capabilities, as each input sequence begins from the same initial state, allowing the reservoir to learn a consistent transformation. For small n, the information about the initial state, including its purity, is still preserved. This could therefore be an indicator that the reservoir generalizes better when starting from a pure state.

Both points are currently only a hypothesis based on interpretation. Therefore, tests are required to verify whether the reduced performance of the QCQA scheme is indeed due to poorer trainability and increased state mixture.

Name	**Starting State**
Up	$\rho(0) = \lvert 0000\rangle \langle 0000\rvert$
Entangled	$\rho(0) = \lvert\Psi\rangle \langle\Psi\rvert, \quad \lvert\Psi\rangle = \frac{1}{\sqrt{2}}(\lvert 0000\rangle + \lvert 1111\rangle)$
Mixed	$\rho(0) = \frac{1}{16}\mathbb{I}_{16}$
Same Random	$\rho(0) = \rho_{\text{rand}}$, same at each step
New Random	$\rho(0) = \rho_{\text{rand}}(t)$, different at each step
QCQA	The reservoir evolves according to Algorithm 3.

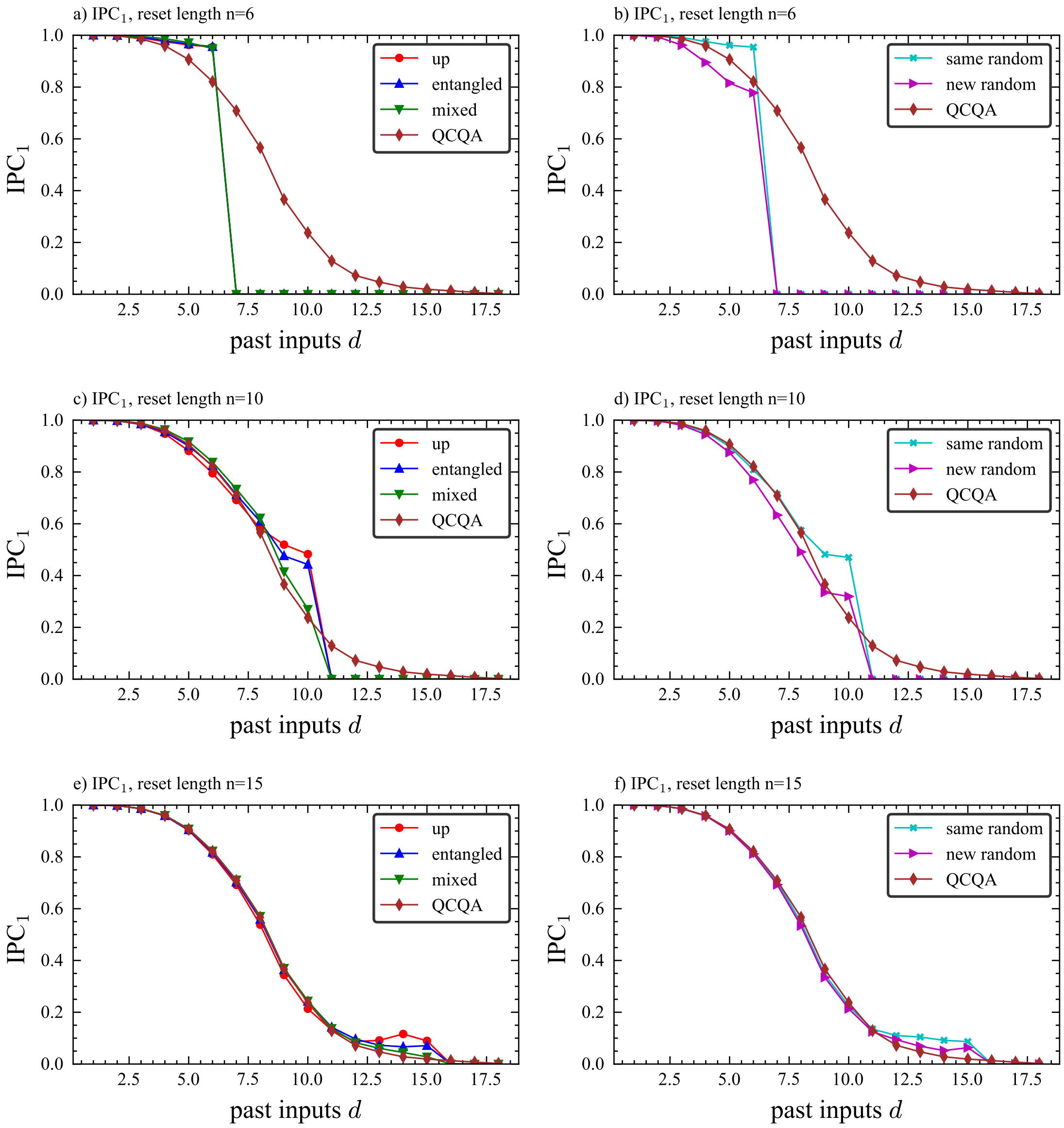
Table IV.6: Initial states used for the reservoir, with corresponding definitions of the starting density matrix $\rho(0)$.

To test this hypothesis, Table IV.6 summarizes the different initial states $\rho(0)$ used to initialize the reservoir in order to evaluate task performance under the LCQA scheme. These include a pure computational basis state (Up), a maximally entangled state (Entangled), a fully mixed state (Mixed), and two random initialization schemes: one where the same random state is reused at each step (Same Random), and another where a new random state is generated at every step (New Random). Additionally, the behavior of the QCQA scheme is also included for comparison. These variations are used to study the influence of the reservoir's initial condition on its memory and processing capabilities.

Figure IV.10 shows how well the system can reconstruct past inputs d steps into the past, as indicated by IPC_1, for $n = 6$ (a, b), $n = 10$ (c, d), and $n = 15$ (e, f). The first row shows results when the up state (red), the entangled state (blue), and the mixed state (green) are used, while the second row shows results when the same random (turquoise) and new random (violet) states are used at each step. In all plots, the performance of the QCQA scheme is also shown for comparison. For $n = 6$ (Figure IV.10.a–b), the up, entangled, mixed, and same-random states all show nearly identical performance, outperforming the QCQA limit up to $d = 6$. The only outlier is the new random initial state (violet), which shows lower performance than the QCQA scheme—already indicating that trainability depends on the initial state.

When increasing to $n = 10$, the up, entangled, and same random states outperform the QCQA limit, while the mixed state performs slightly better than QCQA. The new random state performs worse for all d except at $d = 10$, where it slightly outperforms QCQA. At $n = 15$ (Figure IV.10.e–f), it is clearly visible that the up, entangled, and same random states again exhibit a slight increase in performance compared to the QCQA limit. The new random and mixed states behave more similarly to the QCQA scheme.

Since the up, entangled, and same random states consistently perform better across all realizations, this suggests that the performance gain is due to the use of a fixed initial state. In contrast, the new random, mixed, and QCQA schemes all result in highly mixed initial states, which make training more difficult. This indicates that using a quantum reservoir where information about the initial state is retain results in better memory capacity of the reservoir.


a) IPC_1, reset length n=6

b) IPC_1, reset length n=6

c) IPC_1, reset length n=10

d) IPC_1, reset length n=10

e) IPC_1, reset length n=15

f) IPC_1, reset length n=15

Figure IV.10: Reconstruction performance of past inputs measured by IPC_1 for different reset lengths $n = 6$, $n = 10$, and $n = 15$, across various initial reservoir states. The first row (a–c) compares fixed initial states: up (red), entangled (blue), and mixed (green). The second row (d–f) shows performance with random initializations: same random state (turquoise) and new random state at each step (violet). The QCQA baseline is included in all plots. Results show that fixed initial states (up, entangled, same random) consistently outperform QCQA, especially for small n, while random and mixed states exhibit similar or degraded performance, highlighting the importance of initial state consistency for effective training.

To test the hypothesis of whether the nature of the quantum state—i.e., mixed or pure—affects the reservoir's capabilities, a direct comparison between mixed states and the QCQA scheme must be made. Similar to the case of a new random state at each step, the mixed state performs poorly and exhibits behavior close to the QCQA limit as n increases. For $n = 6$, the mixed state appears to perform slightly better than the QCQA limit, but this performance advantage diminishes at $n = 10$. Therefore, the mixed state performs marginally better than the QCQA limit for small n, but worse than pure states for any n, suggesting that the nature of the quantum state contributes to the performance of quantum reservoir computing. Since the up, entangled, and fixed random pure states consistently outperform other configurations across all realizations, this suggests that the performance gain is due to the use of a fixed pure initial state. A fixed mixed state performs worse, while a new random initial state at each step leads to performance similar to the QCQA limit.

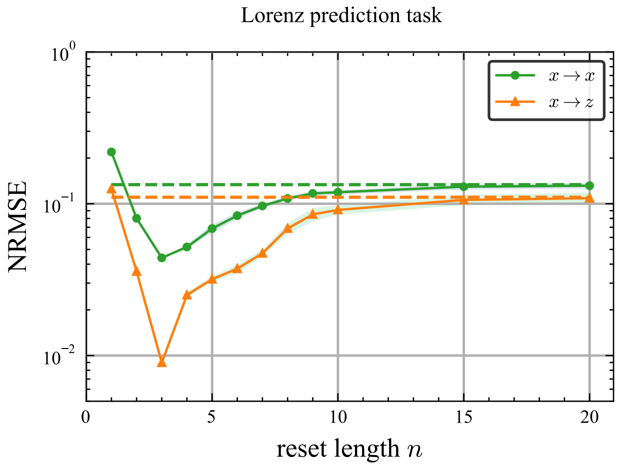


Figure IV.11: Normalized root-mean-squared error (NRMSE) for two Lorenz time-series prediction tasks as a function of reset length n. The tasks include: predicting the x-variable one step ahead (LXX) and cross-predicting the z-variable from the x-input (LXZ). The QCQA baseline is indicated by dashed lines. Both tasks achieve their minimum error at $n = 3$, which coincides with the peak of higher-order nonlinear capacities IPC_5 and IPC_6 [see Figure IV.9.c]. These results highlight the importance of nonlinear processing and demonstrate that optimal reset lengths can enhance task performance without altering the physical reservoir.

To demonstrate that the observed increases in IPCs can translate into improved performance on a time-series prediction task, two tasks related to the Lorenz chaotic attractor are evaluated. In both tasks, the x-variable of the Lorenz system serves as the input to the reservoir.

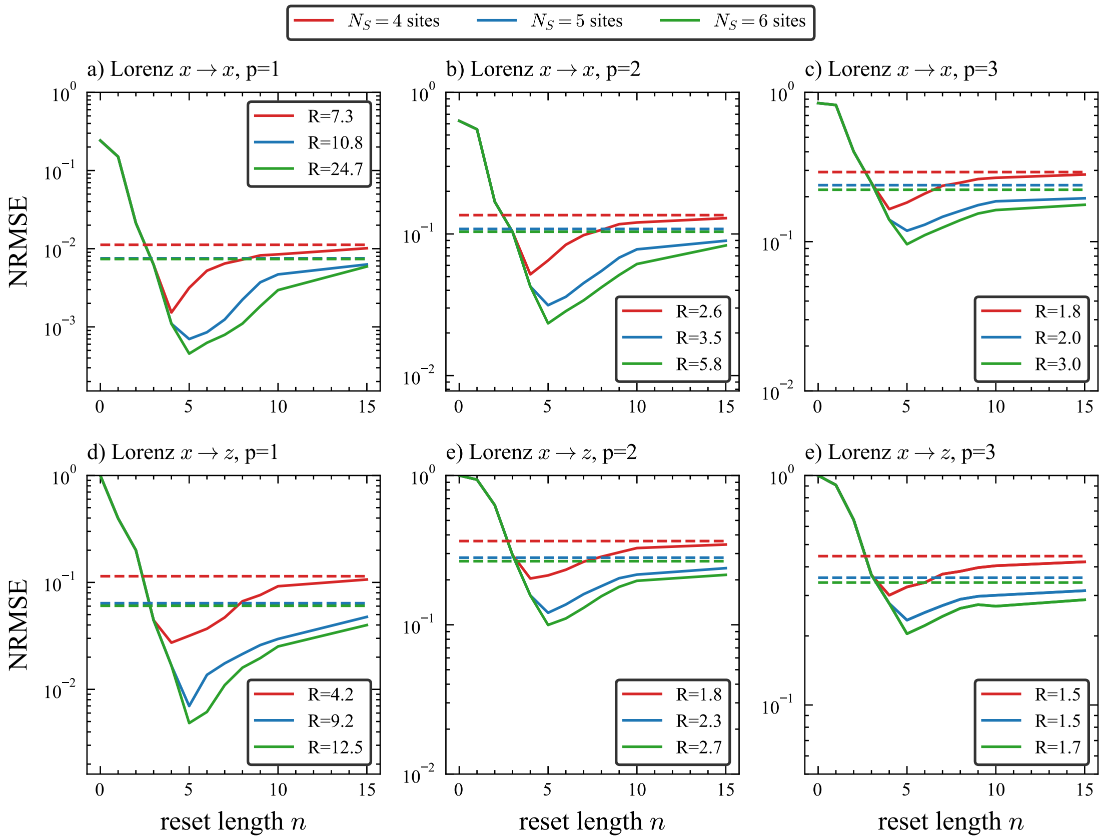


Figure IV.12: Task performance improvement for LCQA compared to QCQA across different quantum reservoir sizes $N_S = 4$ (red), $N_S = 5$ (blue), and $N_S = 6$ (green). Panels (a–c) show the normalized root-mean-squared error (NRMSE) for predicting the Lorenz x-variable one, two, and three steps ahead, respectively. Panels (d–e) display results for cross-predicting the Lorenz z-variable from the x-input for the same prediction horizons. In all cases, the ratio $R = \text{NRMSE(QCQA)} / \min_n \text{NRMSE}(\text{LCQA}, n)$ quantifies performance gain. Significant improvements are observed for all reservoir sizes, with larger gains for bigger systems and shorter prediction horizons. Even as tasks become more complex, LCQA consistently outperforms QCQA, demonstrating its effectiveness and scalability.

The first task (Lorenz $x \to x$) involves predicting the x-variable one step ahead. The second task (Lorenz $x \to z$) involves cross-predicting the z-variable one step ahead. The discretization parameter of the Lorenz time series is set to $\Delta t = 0.1$.
Figure IV.11.a shows the normalized root-mean-squared error (NRMSE) for the LXX and LXZ tasks as a function of the reset length n. For both tasks, lower NRMSE values are achieved at small n, compared to the QCQA limit (dashed lines). For this particular reservoir, the minimum NRMSE for both tasks occurs at $n = 3$, which corresponds to

the reset length at which IPC_5 and IPC_6 reach their maxima [see Figure IV.9.c]. These results demonstrate that the reset length n can be used as a tuning parameter without requiring any changes to the physical reservoir. In general, the optimal reset length will depend on both the task and the specific reservoir being used. Notably, examining the information processing capacities in Figure IV.9 reveals that optimal performance coincides with the peak of IPC_5 and IPC_6. This suggests that higher-order nonlinear IPCs play a more significant role in the Lorenz task, which is reasonable given the chaotic nature of the Lorenz attractor.

To quantify the utility of the LCQA scheme and how it scales with larger numbers of qubits, Figure IV.12 shows the results for quantum reservoirs consisting of $N_S = 4$ (red), $N_S = 5$ (blue), and $N_S = 6$ (green) sites. Figure IV.12.a-c display results for predicting the x-variable one (Figure IV.12.a), two (Figure IV.12.b), and three (Figure IV.12.c) steps into the future, while Figure IV.12.d-e show corresponding results for predicting the z-variable.

In all cases, the ratio

$$\mathrm{R} = \frac{\mathrm{NRMSE(QCQA)}}{\min_n \mathrm{NRMSE}(\mathrm{LCQA}, n)} \tag{IV.29}$$

is computed, representing the relative improvement in task performance at the optimal reset length n compared to the QCQA limit.

For the one-step-ahead prediction of the Lorenz $x \rightarrow x$ task (Figure IV.12.a), performance increases of $R = 7.3$, $R = 10.8$, and $R = 24.7$ are observed for four, five, and six qubits, respectively. This shows a significant performance boost for all system sizes, with improvements increasing alongside reservoir size.

Increasing the task difficulty through multi-step prediction results in a larger NRMSE. Yet even for the two-step prediction task, performance increases of $R = 2.6$, $R = 3.5$, and $R = 5.8$ are observed. Notably, the QCQA limit exhibits errors of approximately $\mathrm{NRMSE} \approx 0.1$, while optimal LCQA configurations achieve errors around $\mathrm{NRMSE} \approx 3 \cdot 10^{-2}$. Although the relative improvement is smaller for more complex tasks, the absolute error reduction is more substantial. A similar trend is observed for $p = 3$, where the task becomes even more difficult, yet performance improvements of up to $R = 3$ are still achieved.

Analyzing the results for the cross-prediction task $x \rightarrow z$ reveals similar patterns, although with smaller improvement ratios R and higher NRMSE overall. The cross-prediction task $x \rightarrow z$ is typically more challenging, as it requires not only predicting future values but also doing so without any direct information about the z variable, making the task inherently more complex. Nevertheless, the LCQA scheme consistently outperforms QCQA across all three prediction horizons ($p = 1$, $p = 2$, and $p = 3$). For $p = 1$, the largest performance improvements are observed, ranging from $R = 4.2$ to $R = 12.5$, while for $p = 2$, improvements of $R = 1.5$ to $R = 1.7$ are recorded. Notably, even for $p = 3$, Figure IV.12.e shows a meaningful absolute NRMSE reduction of approximately ΔNRMSE ≈ 0.1.

To conclude, this chapter addressed the time-complexity problem arising from state collapse due to measurement in quantum reservoir computing and introduced the LCQA algorithm in this context (Algorithm 4). The LCQA scheme for quantum reservoir computing is introduced as a method that enables the reset length to serve as a tuning parameter, controlling the nonlinearity of the reservoir response. This approach significantly reduces the time complexity of quantum reservoir computing for time-series tasks, lowering it from quadratic (Algorithm 3) to linear and thereby facilitating physical implementations for long time series. A comparison between LCQA (Algorithm 4) and the established QCQA (Algorithm 3, [26]) was conducted using a fully connected Ising model. The results indicate that LCQA outperforms the QCQA scheme in both information processing capacity and Lorenz time series prediction tasks.

.

V Chapter: Probing Quantum Phase Transitions

As discussed in Section II.1, chaos is quantified by a positive Lyapunov exponent. Quantum systems, on the other hand, evolve unitarily and are therefore non-chaotic in the strict sense. Nevertheless, certain observables in quantum systems can exhibit behavior that appears chaotic. As the system size increases, it is expected that chaotic features eventually emerge.

To explore this further, the following section delves into key aspects of quantum mechanics, with a particular focus on quantum measures used to characterize quantum phases and quantum chaos. Along the way, historical context is included. In particular, the 150-year-old debate on the arrow of time involving Loschmidt, Boltzmann, and Kelvin remains especially insightful.

The classical limit of some quantum systems can follow chaotic paths. However, the quantum evolution remains unitary and not chaotic, which has led to the development of several tools to connect classical and quantum mechanics [107].

Although chaos is fundamentally a dynamical phenomenon, it is often described using spectral and eigenvector statistics [1, 2, 3]. Recent studies show that isolated quantum systems can thermalize locally, consistent with the eigenstate thermalization hypothesis (ETH) [2, 3]. However, ETH fails in systems exhibiting many-body localization (MBL), where strong disorder increases the number of local integrals of motion and prevents thermalization [4]. The transition between localization and thermal regimes, known as a dynamical phase transition [5, 6], has gained substantial interest.

In quantum computation, it has been shown that MBL can improve trainability in quantum machine learning [108] while it is unfavorable in quantum annealing [109, 110] and quantum random walks [111, 112]. [29] found that task performance in quantum reservoir computing is optimal at the transition between localized and ergodic phases.

Quantum phase transitions remain an active research area with many conflicting views on quantum chaos, MBL and thermalization. Some studies suggest that quantum chaos appears at the transition from localized to the ergodic phase. Others argue that the ergodic phase is both chaotic and thermal which seems counterintuitive from a macroscopic perspective. The debate is further complicated by questions over whether MBL truly exists. Some suggest that MBL disappears in the thermodynamic limit and is merely an artifact of small system sizes. Others argue that MBL is a robust and intrinsic phenomenon. Defining quantum chaos as a phase transition becomes problematic if the underlying phase does not persist in the thermodynamic limit. If the MBL phase ultimately vanishes and all systems thermalize as size increases, then every quantum system would approach thermalization. In that case, chaos might not be a distinct concept but merely describe unusual trajectories on the path to thermal equilibrium.

Despite these uncertainties, this work proceeds with current knowledge and caution. For clarity, the thermalizing phase will be referred to as ergodic or thermal. The transition from localization to ergodicity is described as *more chaotic*, even though the thermal phase itself is treated as non-chaotic throughout this work. The model of interest is the transverse field Ising model (TFIM) with on-site disorder across the MBL, spin glass (SG), and ergodic regimes, originally explored in [29] as a quantum reservoir. The quantities used in this work to probe phase transitions are level spacing statistics, Loschmidt echo, out of time order correlators and Krylov complexity. Krylov complexity in particular has not been fully explored as a phase transition indicator.

The most well-established quantifier of phase transitions used here is level spacing statistics. By analyzing adjacent gaps in the energy spectrum, one can probe quantum chaos, thermalization, and localization. This work computes the level spacing ratio and demonstrates that it successfully captures thermalization and, to some extent, localization. However, for small system sizes, its ability to resolve phase transitions is limited due to the small number of eigenvalues.

A dynamical measure is the Loschmidt Echo, which quantifies the overlap between a forward-evolved quantum state and a perturbed backward-evolved state. Perfect reconstruction implies non-chaotic behavior, while a strong response to small perturbations suggests chaos. The results show that the initial decay rate (Lyapunov exponent) is largest in the SG phase, and essentially zero in both the ergodic and MBL phases.

Operator dynamics can also be analyzed through Out-of-Time-Order Correlators (OTOCs), which measure how the commutator of two operators grows with time. While typically evaluated at finite temperature, in this work OTOCs are considered for a pure state, as would be relevant in a quantum reservoir computing setup. The results indicate that the Lyapunov exponent of the OTOC reaches it's maximum near the phase transition.

The last quantity examined is Krylov complexity which captures how a state or operator evolves in its Krylov basis. It shows clear differences between ergodic and localized phases. After rescaling time by the system's Heisenberg time one can distinguish not only between ergodic and localized systems but also between the two localized regimes, MBL and SG. As the system moves from the SG phase toward ergodicity the late-time average of Krylov spread complexity increases and shifts to a behavior with a peak at early times. In the MBL phase rescaling shows that all trajectories follow the same growth rate and reach the same saturation value. These findings indicate that Krylov complexity may serve as a useful tool to distinguish between MBL, SG, and ergodic behavior, but require a fundamental analysis in future work.

Finally, these same quantum systems are used as quantum reservoirs to compute their information processing capacities (IPCs). The aim is to assess their expressivity in handling input data. The results show that linear memory capacity—or first-order IPC—is highest in the localized regime, particularly in the SG phase. Higher-order IPCs, however, peak near the transition from SG to ergodic, mirroring the behavior of the critical Krylov time introduced in this work.

Below is a short summary of the four methods discussed and how they can be interpreted.

- *Loschmidt Echo* measures the decay of overlap between time evolved states under perturbed and unperturbed Hamiltonians.
- *Level Spacing Statistics* studies the distribution of energy level spacings to distinguish between chaotic and regular behavior.
- *OTOCs* track the growth of operator commutators over time to detect signatures of quantum chaos.
- *Krylov complexity* follows how the operator or state evolves within its Krylov basis where deeper levels contribute more to the overall complexity.

V.1 Level spacing statistics, Loschmidt Echo and Correlators

V.1.1 Transverse-Field Ising Model

The **Transverse-Field Ising Model (TFIM)** describes systems of interacting spins under the influence of a transverse magnetic field. This model exhibits quantum phase transitions, which will also be discussed. The Hamiltonian is given by:

$$H = \sum_{i>j} J_{ij} \sigma_x^{(i)} \sigma_x^{(j)} + \frac{1}{2} \sum_{i=1}^{N} (h + D_i) \sigma_z^{(i)}, \tag{V.1}$$

where $\sigma_{x,z}^{(i)}$ are the Pauli matrices acting on the i-th site, defined as:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In a system with N sites (also referred to as qubits), the spin operators are expressed as:

$$\sigma_\alpha^{(j)} = \bigotimes_{i=1}^{j-1} I_2 \otimes \sigma_\alpha \otimes \bigotimes_{i=j+1}^{N} I_2 \tag{V.2}$$

for $\alpha \in \{x, y, z\}$, where $\otimes$ denotes the tensor product. The commutation relations between the Pauli matrices are:

$$[\sigma_x^m, \sigma_y^n] = 2i\sigma_z^m \delta_{m,n}, \quad [\sigma_y^m, \sigma_z^n] = 2i\sigma_x^m \delta_{m,n}, \quad [\sigma_z^m, \sigma_x^n] = 2i\sigma_y^m \delta_{m,n}.$$

The first term of the Hamiltonian, $\sum_{i>j} J_{ij} \sigma_x^{(i)} \sigma_x^{(j)}$, models the interaction between sites with coupling strength J_{ij}. These interactions encourage the spins to either align or anti-align, depending on whether J_{ij} is positive or negative.

The second term, $\frac{1}{2} \sum_{i=1}^{N} (h + D_i) \sigma_z^{(i)}$, models the influence of a transverse magnetic field and local disorder. The parameter h represents the strength of a uniform transverse field applied to all spins, while D_i introduces site-dependent disorder. The Pauli-Z operators $\sigma_z^{(i)}$ measure spin orientation along the z-axis. This term competes with the spin-spin interaction by encouraging spins to flip due to the transverse field.

To study phase transitions in the Ising model the coupling constants are sampled uniformly from $J_{ij} \in \mathcal{U}\left(-\frac{1}{2}, \frac{1}{2}\right)$ and the external field from $h \in \mathcal{U}[10^{-2}, 10^{2}]$. The on-site disorder is sampled from a uniform distribution $D_i \in \mathcal{U}([-W, W])$ where W is varied within $W \in [10^{-2}, 10^{2}]$. For each pair of W and h values, an average over 50 Hamiltonians is taken unless stated otherwise.

V.1.2 Level Spacing Statistics

As already discussed, analyzing quantum chaos was mostly achieved using spectral and eigenvector statistics [1, 2, 3]. One such tool are level spacing statistics, where the adjacent gaps of the eigenvalues are analyzed as a quantifier of quantum chaos and quantum phases. It holds that in the localized phase, the adjacent gaps exhibit a Poisson distribution and a Wigner-Dyson distribution in the quantum chaotic phase. The distribution of adjacent gaps alone does not reveal the transition to quantum chaos, where the ergodic behavior is addressed as chaotic. This is addressed by the introduction of the dimensionless level spacing ratio [113].

Theory of Level Spacing Statistics

Let H be a Hamiltonian with eigenvalue equation $H|\phi_n\rangle = E_n|\phi_n\rangle$, and let the eigenvalues be ordered increasingly, i.e., $E_1 < E_2 < \dots$. Then, the level spacings s'_n are given by

$$s'_n = E_{n+1} - E_n.$$

This can be normalized to the *average level spacing*:

$$\langle s\rangle = \frac{1}{N}\sum_{n=1}^{N} s_n,$$

resulting in

$$s_n = \frac{s'_n}{\langle s\rangle}. \tag{V.3}$$

Quantum chaos can be assessed by studying the distribution of these normalized level spacings. For regular (integrable) systems, the spacings are typically predictable, while in chaotic systems, the spacings exhibit more randomness. Analyzing these distributions helps classify the system as either chaotic or integrable.

1. For integrable systems, the distribution of s_n follows a Poisson distribution, given by

$$P_{\text{Poi}}(s) \propto e^{-s}. \tag{V.4}$$

2. For chaotic systems, the distribution follows the Wigner-Dyson distribution, characteristic of random matrices:

$$P_{\text{WD}}(s) \propto s e^{-\frac{\pi}{4}s^2}. \tag{V.5}$$

When the level spacing cluster around zero than an analysis with this can fail. This issue can be addressed using the **level spacing ratio**, defined as:

$$r_n = \frac{\min(s_{n+1}, s_n)}{\max(s_{n+1}, s_n)}. \tag{V.6}$$

The corresponding distributions of the level spacing ratio are given by:

$$P_{\mathrm{Poi}}(r) \propto \frac{1}{(r+1)^2}, \qquad P_{\mathrm{WD}}(r) \propto \frac{r+r^2}{(1+r+r^2)^{5/2}}. \tag{V.7}$$

The expectation value of r_n is called the level spacing ratio $\langle r \rangle$ and is given by

$$\langle r \rangle = \frac{1}{N-2} \sum_{n=1}^{N-2} r_n. \tag{V.8}$$

It has been shown that $\langle r \rangle \approx 0.38$ indicates the system is in the localization phase, while $\langle r \rangle \approx 0.54$ corresponds to the thermodynamic (chaotic) phase.

The authors in [29] showed that the dynamical phases influence task performance in quantum reservoir computing, with the best performance observed in the transition from many-body localization ($\langle r \rangle = 0.384$) to the thermodynamic phase ($\langle r \rangle = 0.54$)[113].

Level Spacing Statistics of the Ising Model

Figure V.1 show the average level spacing ratio $\langle r \rangle$ for the transverse-field Ising model across varying disorder strengths W and external field strengths h. The figure shows results for three different system sizes: a four-site system (a), a six-site system (b), and an eight-site system (c). The color-coded heatmaps reveal how $\langle r \rangle$ transitions with changes in W and h.

Localization is quantified by a level spacing ratio of approximately $\langle r \rangle \approx 0.38$, while the ergodic phase exhibits a level spacing ratio closer to $\langle r \rangle \approx 0.53$. Intermediate values indicate potential quantum chaotic behavior.

In particular, Fig. V.1(a) shows a small region of the ergodic phase which grows as the system increases to six and eight qubits. The bottom left region for $h < 10^0$ and $W < 10^0$ corresponds to the spin glass phase while the top right region with $h > 10^0$ and $W > 3 \cdot 10^0$ represents the MBL phase. Another ergodic region is expected between $h < 10^0$ and $10^0 < W < 10^1$ but is only slightly visible. To clearly capture the expected phase transitions using level spacing statistics larger systems need to be simulated.

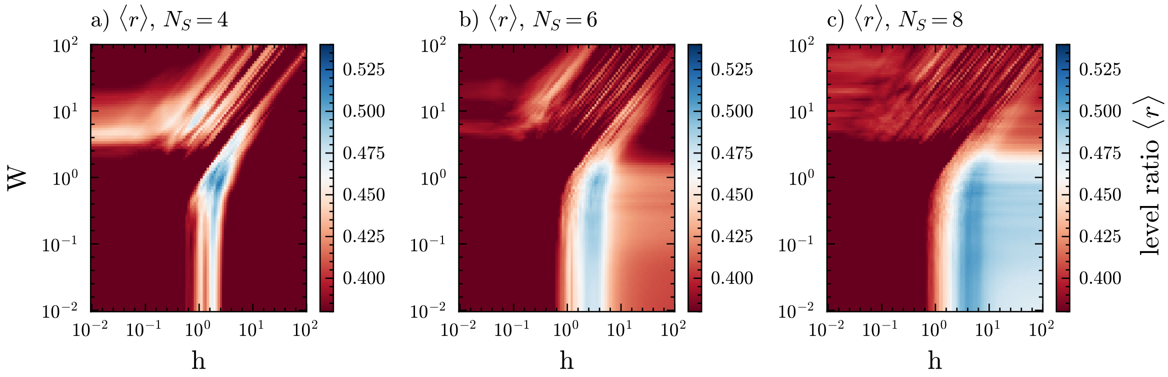


Figure V.1: Average level spacing ratio $\langle r \rangle$ for the transverse-field Ising model as a function of disorder strength W and external field strength h. The figure presents results for system sizes of four (a), six (b), and eight (c) qubits. The color-coded heatmaps illustrate the transition between localized and ergodic phases, with $\langle r \rangle \approx 0.38$ indicating localization and $\langle r \rangle \approx 0.53$ signifying ergodic behavior. The ergodic region (blue) becomes more prominent as the system size increases, suggesting an expansion of quantum chaotic behavior with larger system sizes.

V.1.3 Loschmidt Echo

A related question to quantum chaos is the problem of irreversibility in quantum mechanics. Reconciling the *arrow of time* with quantum theory is difficult because the Schrödinger and von Neumann equations describe unitary evolution. This keeps eigenvalues fixed and prevents a classical interpretation using eigenvalue analysis or Lyapunov exponents (see Section II.1). The idea of irreversibility dates back to Boltzmann [114]. The issue was raised before the discovery of quantum mechanics about 150 years ago by Loschmidt [115], with the first publication by Thomson, later known as Lord Kelvin [116]. Boltzmann, Loschmidt and Kelvin explored the question of irreversibility where Kelvin presented a modern view on the arrow of time and reversibility.

"If we allowed this equalization to proceed for a certain time, and then reversed the motions of all the molecules, we would observe a disequalization. However, if the number of molecules is very large, as it is in a gas, any slight deviation from absolute precision in the reversal will greatly shorten the time during which disequalization occurs... Furthermore, if we take account of the fact that no physical system can be completely isolated from its surroundings but is in principle interacting with all other molecules in the universe, and if we believe that the number of these latter molecules is infinite, then we may conclude that it is impossible for temperature-differences to arise spontaneously...[116]"

He argued that even a small disturbance in the reversed time evolution is sufficient to prevent the system from returning to its original state. He also emphasized that every physical system interacts with its environment; therefore, achieving perfect reversibility would require either eliminating all interactions or reversing the environment itself. Consequently, the disequalization time fundamentally depends on imperfections in the reversed evolution. This concept is closely related to the field of echo dynamics. The measure that quantifies sensitivity to reversed time evolution in quantum mechanics is known as the Loschmidt Echo or Fidelity [7, 117]. The Loschmidt Echo is defined as the overlap of a time-evolved state with the reversed evolution under a perturbed Hamiltonian.

In the case of identical forward and reversed evolution the Loschmidt Echo equals one. For different evolutions it decreases and the decay time is interpreted as the disequilibration time discussed by Loschmidt and Kelvin. Although promising for studying irreversibility the Loschmidt Echo has only recently been used for this purpose [7]. It was mainly used as a standard experimental method in nuclear magnetic resonance where the first measurements were performed [118, 119, 120, 121]. In quantum information the Loschmidt Echo is known as fidelity and is widely used to describe system quality [9, 122, 8]. The main motivation for studying the Loschmidt Echo comes from the quantum chaos community [8]. It has been shown that in the classical limit the Loschmidt Echo behaves similarly to its classical counterpart. Fidelity was first introduced by Peres [117, 123]. An important insight came from defining the Lyapunov exponent using the initial decay of the Loschmidt Echo. This will be referred to as the Loschmidt Lyapunov exponent or simply the Loschmidt exponent. A review of the main aspects of Loschmidt Echoes is given in [124].

The Loschmidt Echo $G(t)$ is defined as the absolute value of the overlap between the forward time-evolved state $|\psi(t)\rangle = e^{-iHt}|\psi_0\rangle$ and the perturbed backward-evolved state $\langle\psi'(t)| = \langle\psi_0| e^{-iH't}$, where $|\psi(0)\rangle$ is the initial state. It is given by

$$G(t) = |\langle\psi'(t)|\psi(t)\rangle| = \left|\langle\psi(0)| e^{iH't} e^{-iHt} |\psi(0)\rangle\right|$$

Here one considers a Hamiltonian H and a perturbed Hamiltonian $H' = H + \epsilon V$, where V is a potential and ϵ is a small parameter. If $G(t)$ decays exponentially fast for a small perturbation, it implies that the system is chaotic. If $G(t) \approx 1$ or remains constant, the system is considered not chaotic. The decay can be quantified by fitting an exponential function to the initial decrease of the Loschmidt Echo:

$$G(t) \propto e^{-\lambda_L t}, \tag{V.9}$$

where λ_L is interpreted as a Lyapunov exponent of the Loschmidt Echo. Larger values of λ_L for small perturbations indicate more chaotic behavior. In Appendix A more results are presented on the behavior of the Loschmidt exponent λ_L for the TFIM and the method used to compute λ_L is explained.

V.1.4 Out-of-Time-Ordered Correlators (OTOCs)

Instead of analyzing the evolution of quantum states the dynamics can be studied through the evolution of operators. In this context operator correlators were introduced and first discussed by Larkin and Ovchinnikov in 1969 to investigate the use of semiclassical methods in superconductivity [125]. In this approach two observables are considered. One evolves in the Heisenberg picture while the other remains fixed. The commutator of these operators is computed, squared, and its expectation value is taken with respect to a quantum state.

Out of time ordered correlators (OTOCs) have gained attention only recently, mainly due to their relevance in studies connecting black hole horizons to quantum chaos [10, 11, 12]. Black holes are known to show early exponential growth in OTOCs, which is a key feature of classical chaotic behavior. The Lyapunov exponent for OTOCs is similar to that of the Loschmidt Echo but involves two observables instead of comparing perturbed and unperturbed time evolution. The Sachdev-Ye-Kitaev (SYK) model has also been shown to scramble information rapidly [13]. OTOCs are used in many fields, and more details can be found in [126]. Interesting links between OTOCs and Loschmidt Echoes are discussed in [127, 128].

Growth in OTOCs can also describe the spread of quantum information, known as information scrambling [12]. The time when scrambling takes place is often called the scrambling time and marks when the subsystem reaches maximal entanglement entropy [129].

Typical OTOC behavior shows three phases. First, there is exponential growth where the OTOC Lyapunov exponent is defined. Then follows an intermediate stage and finally long time saturation [130]. Chaotic systems tend to show stable long time values, while oscillating OTOCs are often linked to non chaotic systems. The oscillations in the late time behavior can also help quantify quantum chaos [131, 132]. A full review of OTOCs is given in [133].

Theory of OTOCs

The OTOC for two operators $A(t)$ and $B(0)$ in the Heisenberg picture is given by:

$$\mathrm{OTOC}(t) = \langle [A(t), B(0)]^\dagger [A(t), B(0)] \rangle .$$

If the OTOC grows exponentially,

$$\mathrm{OTOC}(t) \propto e^{2\lambda_O t}, \tag{V.10}$$

with a Lyapunov exponent $\lambda_O > 0$, then the system is considered chaotic. If $\lambda_O \leq 0$, the system is not chaotic. Here, λ_O is interpreted as the Lyapunov exponent associated with the OTOC. In Appendix A the OTOC exponent λ_O is analyzed in detail for the TFIM and the method used to compute λ_O is explained.

V.2 Understanding Phase Transitions using Krylov Complexity

Krylov complexity is the most recent quantity used to study quantum chaos [14, 15]. In this setting the Krylov space of the Liouvillian is used to describe operator growth which leads to the definition of Krylov operator complexity. The spread of a time evolved operator in Krylov space serves as the basis for defining a complexity measure.

It has been shown that Krylov operator complexity gives useful insights into quantum chaos [61], spin chains, SYK models [16, 69, 58], two dimensional conformal field theories (2D CFTs) [74], and also in closed, driven [72] and open quantum systems [71].

In a similar way the idea of Krylov spread complexity was introduced [16]. Instead of describing operator scrambling this measure studies how a pure state evolves under the Schrödinger equation and spreads within a basis built from powers of the Hamiltonian. This method has gained interest and a full review of both Krylov operator and spread complexity is given in [134].

V.2.1 Revisiting Krylov spread Complexity

Assume an initial state $|\psi_0\rangle$ and a quantum system described by a Hamiltonian H. The time evolution is given by

$$|\psi(t)\rangle = e^{-iHt}\,|\psi_0\rangle\,. \tag{V.11}$$

This can be expanded as

$$|\psi(t)\rangle = \sum_{n=0}^{\infty} \frac{(-it)^n}{n!} H^n\,|\psi_0\rangle\,. \tag{V.12}$$

Since the coefficients $\frac{(-it)^n}{n!}$ are just constants, it follows that the time-evolved state lies in the span of successive powers of the Hamiltonian acting on the initial state, i.e.,

$$|\psi(t)\rangle \in \mathrm{Span}\Big\{ |\psi_0\rangle, H|\psi_0\rangle, H^2|\psi_0\rangle, \dots \Big\}. \tag{V.13}$$

Since H is a linear operator and the system is finite-dimensional, the Krylov subspace property implies that there exists an integer m such that

$$|\psi(t)\rangle \in \mathrm{Span}\Big\{ |\psi_0\rangle, H|\psi_0\rangle, H^2|\psi_0\rangle, \dots, H^{m-1}|\psi_0\rangle \Big\} := \mathrm{K}_m. \tag{V.14}$$

Orthonormalizing this basis using the Lanczos or Arnoldi algorithm yields

$$\mathrm{K}_m = \mathrm{Span}\Big\{ |\psi_0\rangle, H|\psi_0\rangle, \dots, H^{m-1}|\psi_0\rangle \Big\} \xrightarrow{\text{ONB}} \mathrm{Span}\Big\{ |k_0\rangle, |k_1\rangle, \dots, |k_{m-1}\rangle \Big\} = \mathrm{K}_m. \tag{V.15}$$

Any time-evolved state with initial condition $|\psi_0\rangle$ can now be represented in this basis as

$$|\psi(t)\rangle = \sum_{n=0}^{m-1} \langle k_n|\psi(t)|k_n|\psi(t)\rangle |k_n\rangle. \tag{V.16}$$

With $\alpha_n = \langle k_n|\psi(t)|k_n|\psi(t)\rangle$, the Krylov spread complexity is defined as

$$\mathcal{K}_S = \sum_{n=0}^{m-1} (n+1)|\alpha_n|^2. \tag{V.17}$$

Typically, $\mathcal{K}_S$ is studied as a function of time for integrable and chaotic systems, where chaotic systems tend to exhibit larger spreading and distinct time-dependent behavior[16].

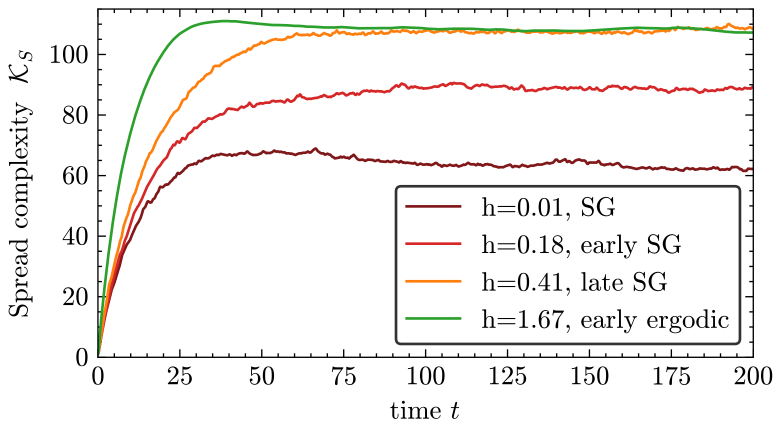
Krylov spread complexity has been analysed in some recent works, with [15] analyzing the suppression of localization in the complexity. [135] showed that the Lanczos coefficients in the computation of Krylov operator complexity scale asymptotically in the MBL phase, similarly to ergodic systems. [136] also analyses how the Lanczos coefficients evolve over time. [137] shows that the peak of Krylov spread complexity is associated with ergodic behavior, which falls off towards the MBL regime. This is similar to observations made for integrable and chaotic systems, where chaotic systems show the same peak in Krylov spread complexity at short time scales [16], while integrable systems increase and then saturate.

V.2.2 Spread Complexity at the Spin Glass to Ergodic Transition

The behavior of the spread complexity $\mathcal{K}_S$ is first analyzed across different regimes in the parameter space defined by the external field h and disorder W.

SG to ergodic transition with N_S=8 sites.

a) Localized transition



b) Ergodic transition

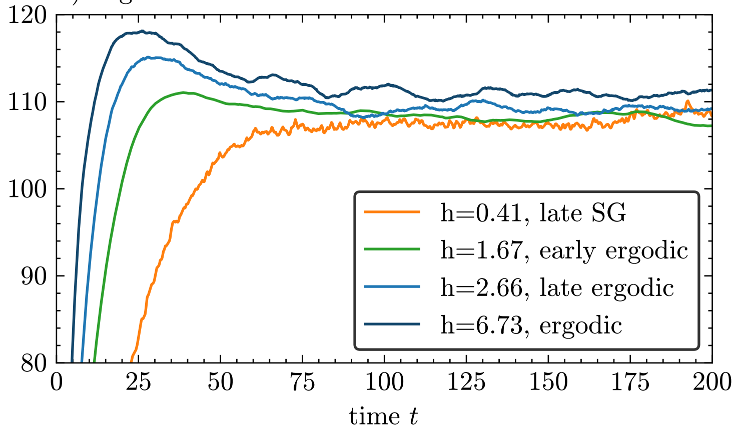


Figure V.2: Time evolution of LKrylov spread complexity $\mathcal{K}_S$ averaged for 20 Hamiltonians with $N_S = 8$ sites, where the on-site disorder was sampled from $D_i \in \mathcal{U}([-0.06, 0.06])$, and varying external fields h. (a) Shows behavior for $h = 0.01$ (SG), $h = 0.18$ (early SG), $h = 0.41$ (late SG), and $h = 1.67$ (early ergodic). (b) Shows $h = 0.41$ (late SG), $h = 1.67$ (early ergodic), $h = 2.66$ (late ergodic), and $h = 6.73$ (ergodic). A continuous increase in both late-time values and initial growth rates of $\mathcal{K}_S$ is observed across the transition from the localized to the thermal regime.

Figure V.2 shows the time evolution of the Krylov spread complexity $\mathcal{K}_S$ averaged over 20 Hamiltonians with $N_S = 8$ sites and interaction strength $W = 0.06$, as a function of the external field h. The points in parameter space are chosen to reveal the transition from the spin glass (SG) phase to the ergodic phase. In panel (a), four values of h are shown: $h = 0.01$ (SG), $h = 0.18$ (early SG), $h = 0.41$ (late SG), and $h = 1.67$ (early ergodic). In the SG regime, $\mathcal{K}_S$ grows slowly and saturates at a low value. As h increases toward the transition, the late time value of $\mathcal{K}_S$ also increases, indicating stronger spreading in the Krylov subspace (red). Closer to the transition, at $h = 0.41$ (orange), $\mathcal{K}_S$ grows and saturates at a value similar to that of the early ergodic case ($h = 1.67$, green).

Panel (b) shows $\mathcal{K}_S$ for $h = 0.41$ (late SG), $h = 1.67$ (early ergodic), $h = 2.66$ (late ergodic), and $h = 6.73$ (ergodic). As the system enters the ergodic regime, $\mathcal{K}_S$ reaches its saturation value more quickly and shows an early peak that decays toward saturation. The late SG (orange) and early ergodic (green) curves differ in their initial growth, with stronger oscillations seen in the late SG case. A peak becomes more visible as the system moves deeper into the ergodic phase. This change in early behavior—from slow growth to a peak followed by saturation—is clearest in the late ergodic case (blue). The ergodic case (dark black) shows the largest peak and the highest saturation value of Krylov spread complexity. This trend from slow growth and low saturation in the SG phase to fast growth, an early peak, and high saturation in the ergodic phase shows that Krylov spread complexity captures key features of the phase transition.

At this point it is important to note the difference in speed between quantum systems. Depending on the value of h the faster growth of Krylov spread complexity might result from larger eigenvalue differences $\omega_{ij} = \varepsilon_i - \varepsilon_j$, where ε_i are the eigenvalues of the system. For this reason it makes sense to rescale the time axis using the Heisenberg time T_H which can be calculated from the level spacings of the eigenvalues. The Heisenberg time is given by

$$T_H = \frac{2\pi}{\bar{s}} \tag{V.18}$$

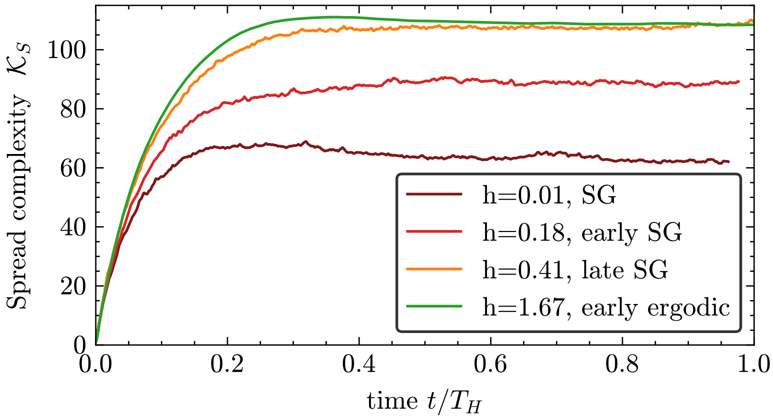
where $s_n = \varepsilon_{n+1} - \varepsilon_n$ are the level spacings and $\bar{s} = \langle s_n \rangle$ is the average spacing.

Figure V.3 shows how the rescaled Krylov complexity behaves during the transition from the SG phase to the ergodic phase. In the spin glass regime, $\mathcal{K}_S(t/t_H)$ (Figure V.3.a) shows similar behavior to $\mathcal{K}_S(t)$ in Figure V.2.a.

On the other hand, Krylov complexity changes noticeably in the early, late, and fully ergodic cases after rescaling, as shown in Figure V.3.b. In absolute time, the ergodic regime displays the fastest growth and the highest peak in Krylov complexity (Figure V.2.b). After rescaling, Figure V.3 shows that the early ergodic (b, green) and late ergodic (b, blue) cases reach their maximum values faster than the fully ergodic case (b, dark blue). These results demonstrate that Krylov spread complexity effectively captures the phase transition from the SG phase to ergodic behavior.

Rescaled SG to ergodic transition with N_S=8 sites.

a) Localized transition



b) Ergodic transition

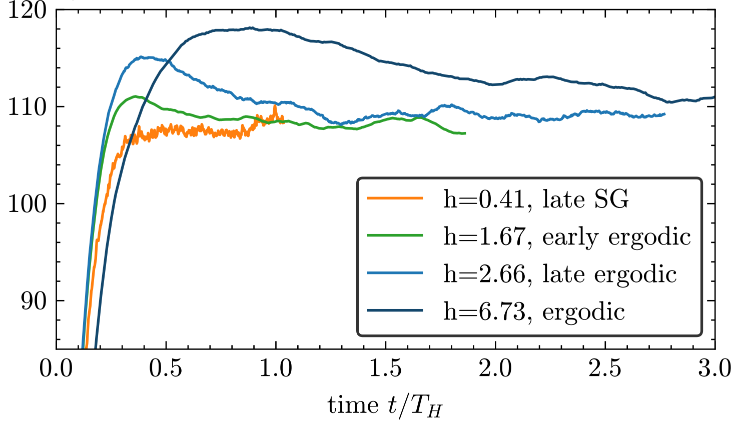


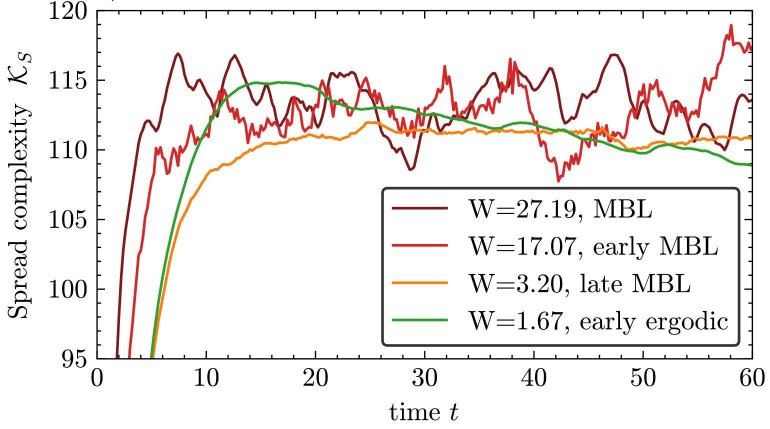
Figure V.3: Same as Figure V.2, but with rescaled time t/T_H, where T_H represents the Heisenberg time.

V.2.3 Spread Complexity at the MBL to Ergodic Transition

Figure V.4 shows the time evolution of the Krylov spread complexity $\mathcal{K}_S$ across the transition from the many body localized (MBL) to the ergodic regime. The data is averaged over 20 Hamiltonians with $N_S = 8$ sites and fixed external field $h = 6.73$, while the disorder strength W is varied to explore different regimes.

MBL to ergodic transition with N_S=8 sites.

a) Localized transition



b) Ergodic transition

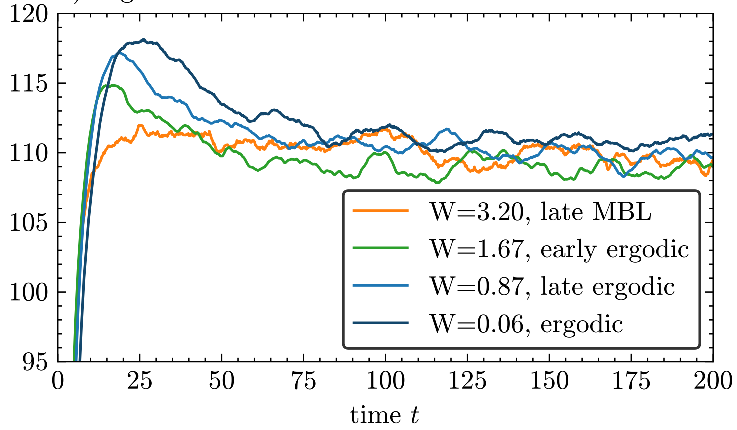


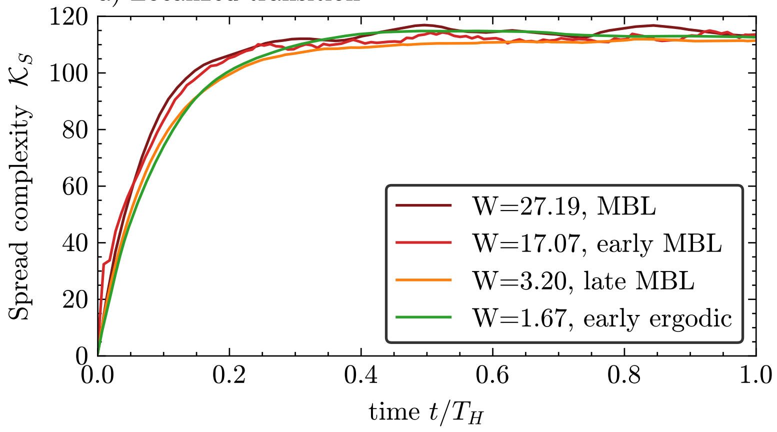
Figure V.4: Time evolution of Krylov spread complexity $\mathcal{K}_S$ averaged for 20 Hamitlonians with $N_S = 8$ sites, where the external field is fixed at $h = 6.73$ and the on-site disorder was sampled from $D_i \in \mathcal{U}([-W, W])$. (a) Shows behavior for $W = 27.19$ (SG), $W = 17.07$ (early SG), $W = 3.20$ (late SG), and $W = 1.67$ (early ergodic). (b) Shows $W = 3.20$ (late SG), $W = 1.67$ (early ergodic), $W = 0.87$ (late ergodic), and $W = 0.06$ (ergodic). A continuous increase in both late-time values and initial growth rates of $\mathcal{K}_S$ is observed across the transition from the localized to the thermal regime.

Panel (a) shows $\mathcal{K}_S(t)$ for four values of disorder strength: $W = 27.19$ (MBL), $W = 17.07$ (early MBL), $W = 3.20$ (late MBL), and $W = 1.67$ (early ergodic). The curves, especially in the MBL phase, show strong oscillations which make interpretation of the growth more difficult compared to the SG to ergodic transition where the late time average changed more clearly.

Figure V.4 (b) shows the transition into the ergodic regime with $W = 3.20$, $W = 1.67$, $W = 0.87$, and $W = 0.06$. Here $\mathcal{K}_S$ shows a trend similar to what was seen in Figure V.2. The early time peak grows larger as the system moves deeper into the ergodic phase. This similarity supports the strength of $\mathcal{K}_S$ as an indicator of ergodic dynamics.

Rescaled MBL to ergodic transition with N_S=8 sites.

a) Localized transition



b) Ergodic transition

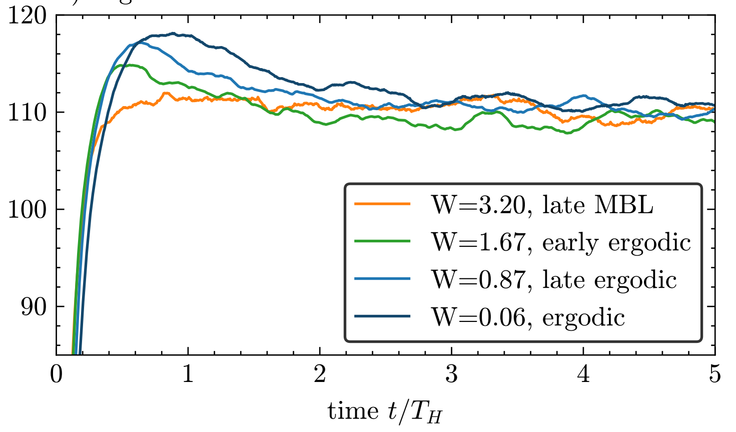


Figure V.5: Same as Figure V.4, but with rescaled time t/T_H, where T_H represents the Heisenberg time.

In Figure V.5 the time axis is rescaled to the Heisenberg time. Figure V.5.a shows that the rescaled Krylov spread complexity behaves the same for all systems in the MBL phase. This is clearly different from the behavior in the SG phase (Figure V.3.a), where the spread complexity $\mathcal{K}_S$ saturates to larger values as the system gets closer to the ergodic phase. This difference shows that the rescaled Krylov complexity $\mathcal{K}_S(t/t_H)$ can be used to tell apart the MBL and SG phases even though both are localized. Using level spacing statistics would result in both phases showing a Poisson distribution and therefore not being distinguishable. In addition, the difference with the ergodic regime, where the rescaled curves stay distinct, further supports that $\mathcal{K}_S$ captures and distinguishes the nature of quantum dynamics across localized and ergodic phase.

V.2.4 Krylov saturation time

Even though this is quantitatively analyzable for points in the parameter space, it would require examining the time evolution of $\mathcal{K}_S$ at each point in parameter space. One interesting feature observable in Figure V.2 is that during the transition, the time it takes to reach saturation appears to increase. In the SG region, $\mathcal{K}_S$ remains small, such that saturation is reached quickly, while in the ergodic case, the dynamics are fast and scramble over the whole Krylov space. In the transition region, however, it takes longer to spread over the full Krylov space.

In that regard, this work investigates the time τ_K, when the Krylov spread complexity reaches 3 dB below its saturation value, i.e.,

$$\mathcal{K}_S(\tau_K) = \frac{\bar{\mathcal{K}}_S}{\sqrt{2}}. \tag{V.19}$$

If the initial speed of spreading into the Krylov basis differs between dynamical phases, then this time is expected to become dependent on the phase. In the following, this time will be referred to as τ_K, the **critical Krylov time**.

Again, the TFIM is again considered with the same parameters as before. For each Hamiltonian, 20 random initial states $|\psi_0\rangle$ are sampled, and the spread complexity is averaged over time. For each parameter pair (h, W), 20 Hamiltonians are computed. In each case, the late-time average of the spread complexity $\bar{\mathcal{K}}_S$ is calculated by averaging over the interval $t = 60$ to $t = 120$. The time $\tau_K^{(l)}$ is then determined for each Hamiltonian $H^{(l)}(h, W)$ such that

$$\mathcal{K}_S(\tau_K^{(l)}) = \frac{\bar{\mathcal{K}}_S}{\sqrt{2}},$$

where the superscript (l) denotes the index of the Hamiltonian. The critical spreading time τ_K is then averaged over all Hamiltonians as $\tau_K = \frac{1}{20}\sum_{l=1}^{20} \tau_K^{(l)}$.

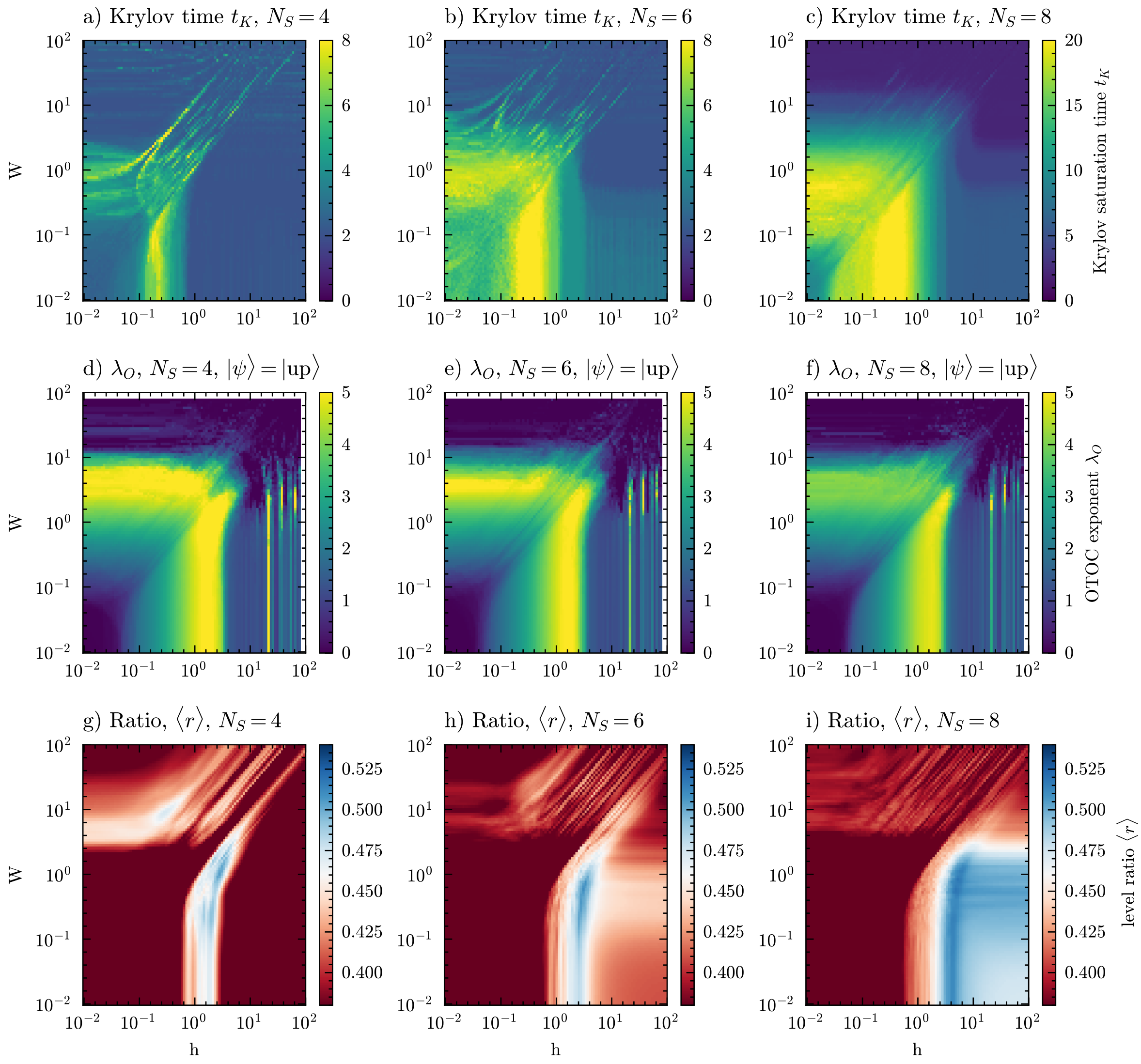


Figure V.6: **Critical Krylov time** τ_K for systems with 4, 6, and 8 sites shown in panels (a), (b), and (c), respectively. The critical Krylov time is defined as the time when the Krylov spread complexity first reaches -3 dB of its late-time value, i.e., $\mathcal{K}_S(\tau_K) = \bar{\mathcal{K}}_S/\sqrt{2}$. The second row (d–f) shows the OTOC Lyapunov exponent λ_O, computed with respect to the up state $|\mathrm{up}\rangle$. The third row (g–i) presents the level spacing ratio $\langle r\rangle$. Systems with four, six, and eight qubits are shown from left to right. Note that the peak of the critical Krylov time τ_K appears before the maximum λ_O and $\langle r\rangle$.

Figure V.6 shows the critical Krylov time τ_K, the OTOC Lyapunov exponent λ_O, and the average level spacing ratio $\langle r\rangle$ as functions of disorder strength W, for spin chains with 4, 6, and 8 qubits. The OTOC is computed using an exponential fit on the initial increase and is discussed in detail in Appendix A.

The first row (a–c) shows the critical Krylov time, defined as the time when the Krylov spread complexity first reaches $\bar{\mathcal{K}}_S/\sqrt{2}$, which corresponds to a 3 dB drop from its saturation value. The second row (d–f) shows λ_O, computed from the out of time ordered correlator, and the third row (g–i) shows the level spacing ratio $\langle r \rangle$, which is used to identify the ergodic to many body localized transition.

For all system sizes the maximum of τ_K appears at lower disorder values than the peaks of both λ_O and $\langle r \rangle$. This shows that Krylov spreading slows down before the system reaches the actual phase transition. The consistency of this behavior across different system sizes suggests that τ_K is sensitive to early dynamical changes and may be useful for detecting the approach to the SG transition even before changes appear in spectral or OTOC based measures.

V.3 Quantum Reservoir Computing and Quantum Phase Transitions

At this point, the question arises how Quantum Reservoir Computing (QRC) and phase transitions are connected. In the calculation of task performance in QRC, only four and six qubits are considered. For each point in the parameter space of h and W, the results are averaged over 20 Hamiltonians. [29] discussed dynamical phase transitions and quantum reservoir computing task performance for this Hamiltonian within the same parameter space. The authors computed the level spacing ratio while varying h and W. For task performance in QRC, they restricted their study to either fixing h and sweeping over W, or fixing W and sweeping over h, to isolate the influence of each parameter. This was most likely due to their use of 10 sites, where averaging task performance becomes significantly more difficult and resource intensive. This work therefore limits the analysis to four and six sites and samples the full parameter space of h and W. The information processing capacity is computed instead of full reservoir task performance to allow comparison of different task complexities, where higher order IPCs typically correspond to tasks that require more nonlinearity.

Figure V.7 shows the relationship between the IPC of quantum reservoir computers and quantum phase transitions in a disordered spin chain model. The localized and ergodic phases are identified using the averaged level spacing ratio $\langle r \rangle$ in Figure V.7.f. A value of $\langle r \rangle \approx 0.38$ indicates Poisson statistics and localization, while values around $\langle r \rangle \approx 0.54$ are consistent with Wigner Dyson statistics and ergodicity. The transition between these phases occurs near $W \approx 2$ and $h \approx 1$.

Panels (a) to (d) show IPC_1 through IPC_4, each as a function of the parameters W and h, averaged over 20 Hamiltonians with four qubits. IPC_1 in panel (a) represents the linear memory capacity of the QRC. It measures how well the system can reconstruct past inputs without applying transformations. Interestingly, IPC_1 is highest in the SG phase, where $\langle r \rangle$ is smallest. This suggests that QRCs are well suited for memory tasks in localized regimes.

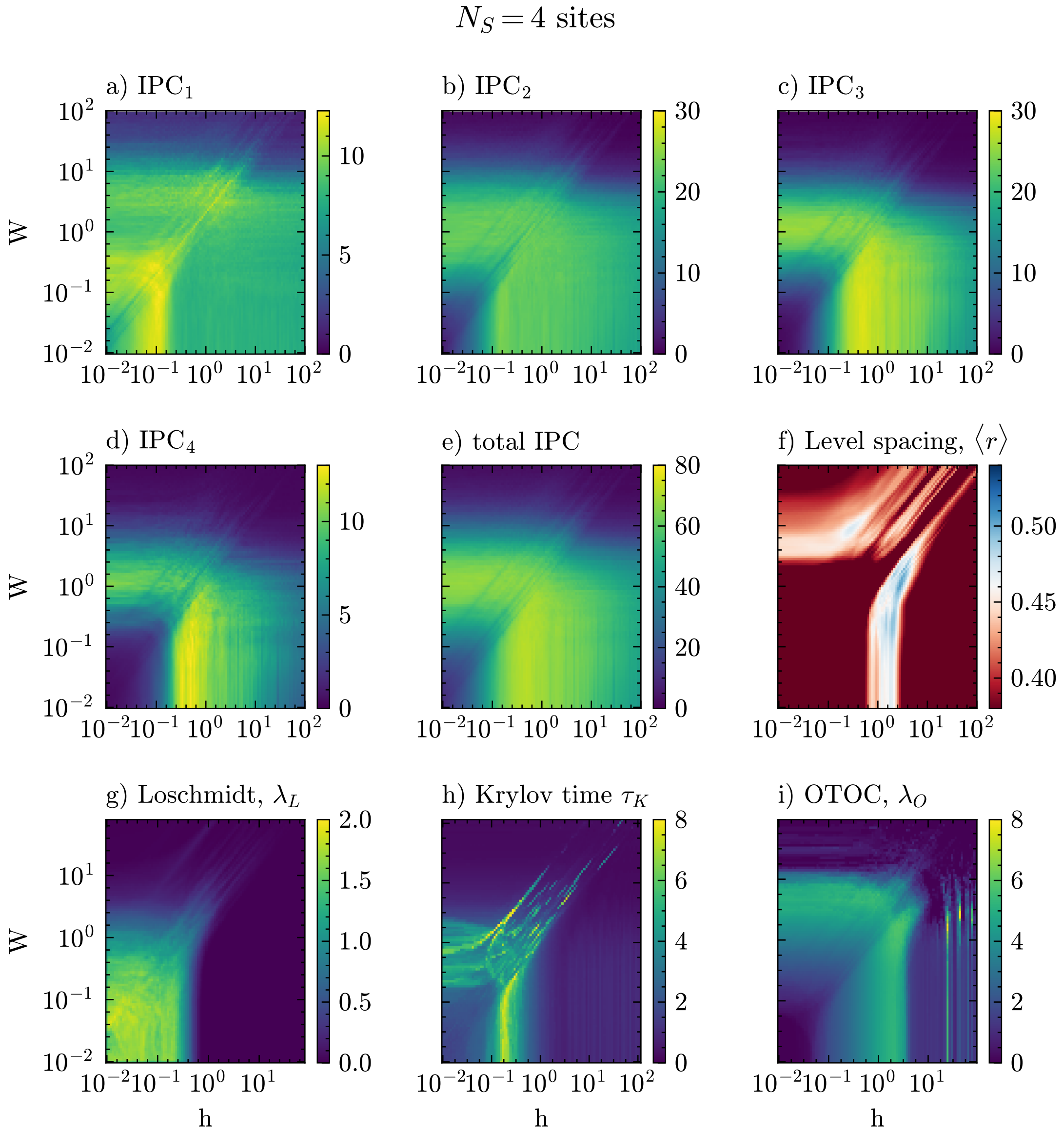


Figure V.7: First four orders of the IPC are shown in a) IPC_1, b) IPC_2, c) IPC_3, and d) IPC_4, and the total IPC in e), for different values of W and h, averaged over 20 Hamiltonians. The level spacing ratio is shown in f), the Loschmidt Lyapunov exponent λ_L in g), the critical Krylov time τ_K in h), and the OTOC Lyapunov exponent λ_O in i). It is clearly visible that the memory capacity IPC_1 is maximal in the SG phase. Higher order IPCs tend to perform better near the phase transition from SG to ergodic.

Panels (b), (c), and (d) show IPC_2, IPC_3, and IPC_4, which reflect higher order processing capacity. These values are largest near the phase transition, around $h \approx 0.5$ and $W \approx 1$. The total IPC in Figure V.7.e peaks at the SG to ergodic phase transition, indicating enhanced data expressivity.

Panel (g) shows the Loschmidt Lyapunov exponent λ_L, which is largest in the SG phase. A more detailed discussion of how this was computed is given in Appendix A. The memory capacity IPC_1 and λ_L both show large values in the bottom left region of the parameter space. In contrast, IPC_4 is largest where λ_L is small. Figure V.7.h presents the critical Krylov time τ_K, a dynamic quantity capturing the spread behavior of Krylov complexity. Among all measures, τ_K shows the best match with the peak of IPC_4. Above panel (e), the total IPC again reaches its maximum around $h \approx 0.8$, just before the phase transition. The final panel (i) shows the OTOC Lyapunov exponent λ_O, which is largest deeper in the ergodic phase.

These patterns reveal that different quantum measures track similar underlying dynamics. In the SG phase, localization limits information flow, explaining the high memory but low complexity. Since interaction between localized sites is weak, complex correlations are minimal, as shown by the small values of λ_O and higher order IPCs. The critical Krylov time τ_K appears to detect the onset of phase transitions earlier than other quantities.

To study larger reservoirs, a system with six sites is analyzed in Figure V.8. Similar behavior is observed. The linear IPC_1 performs best outside the ergodic regime. Higher order IPCs peak near the phase transition. The total IPC reaches its maximum at $h = 1$ and $W \approx 2$, while τ_K peaks slightly before. The Loschmidt Lyapunov exponent λ_L is largest in the SG phase and then decreases, while the OTOC exponent λ_O increases deeper into the ergodic regime.

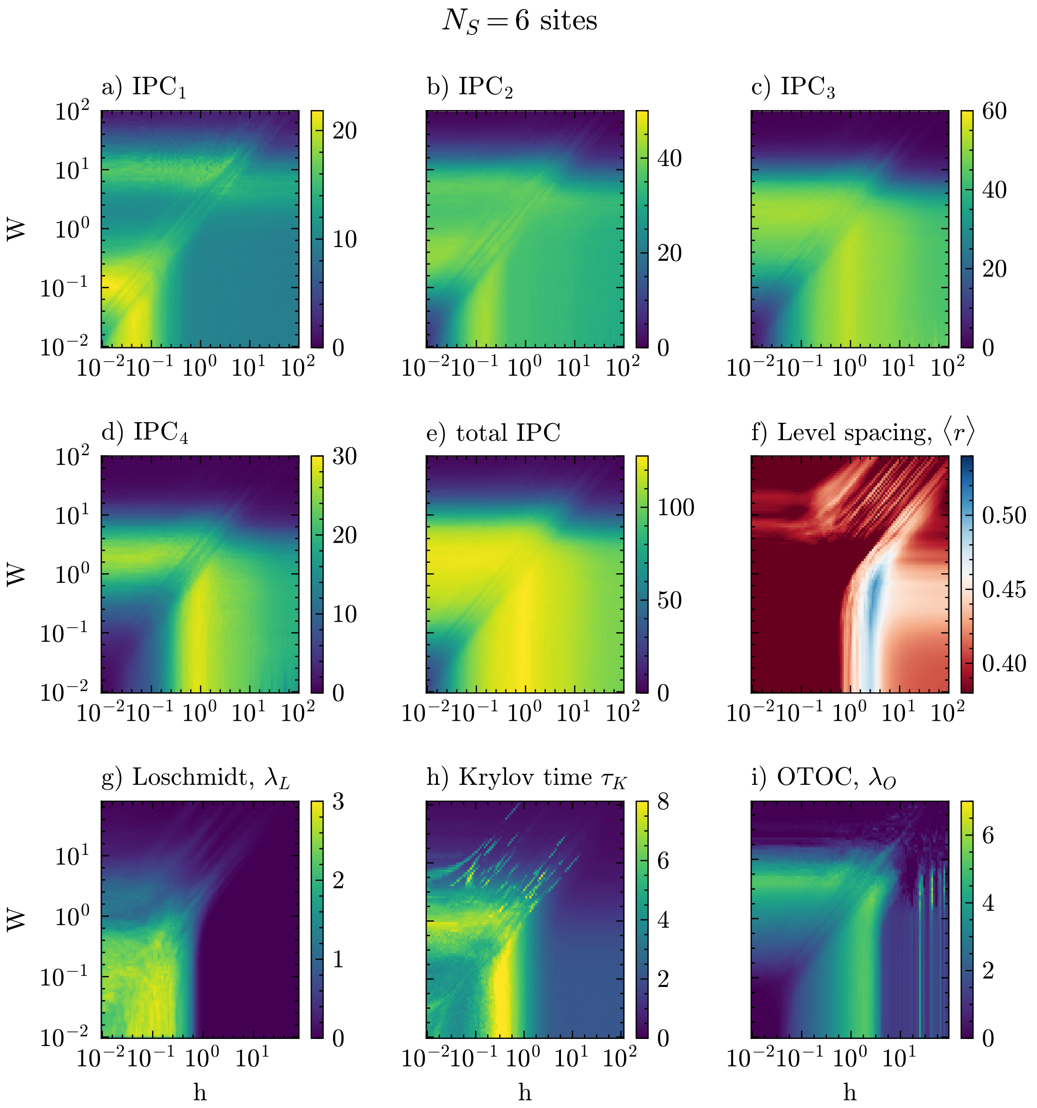


Figure V.8: Same as Figure V.7 for a system with $N_S = 6$ sites.

V.3.1 Conclusion

Lastly, it should be noted that this is not a complete analysis of quantum phase transitions. Research in this area requires careful and detailed investigation. The goal of this chapter was twofold. First, to provide an overview of how phase transitions can be probed and how concepts such as quantum chaos relate to them. Second, to connect these ideas to classical reservoir computing, where optimal task performance often appears near the edge of chaos or at the transition to chaotic behavior. It was shown that Krylov complexity is a promising candidate for identifying phase transitions. After rescaling time by the system's individual Heisenberg time, it became possible to distinguish between the spin glass and many body localized phases, something that level spacing statistics fail to do. The critical Krylov time was introduced as a feature of Krylov spread complexity that captures early dynamical changes. This chapter also revisits the results of [29]. Instead of computing task performance at select points in the parameter space of h and W, this work calculates different orders of the information processing capacity (IPC). The results suggest that memory and nonlinearity exhibit best performance in different regions of the phase diagram. Additionally, the relationship between various dynamical quantities and quantum reservoir performance was explored. The observation that rescaled Krylov complexity can distinguish between different localized regimes is particularly intriguing and calls for further investigation. Overall, this chapter is intended as an outlook for future research. It highlights strong connections between quantum phase transitions and information-theoretic measures. The role of Krylov complexity, especially the origin of its peak in the ergodic regime, remains not fully understood and presents an interesting direction for future work.

VI Chapter: Summary and Outlook

Summary

This thesis explored the interface between quantum dynamics, quantum reservoir computing (QRC), and quantum information measures, with the goal of improving the interpretability and performance of quantum reservoir computers.

A central contribution was the reformulation of Krylov complexity in terms of time-evolved quantum states and operators. In the case of states, it was shown that the Krylov space can be constructed directly through quantum measurements, without requiring explicit knowledge of the system's Hamiltonian. This finding challenged earlier claims that Krylov bases formed from Hamiltonian powers necessarily minimize spread complexity. On the basis of these constructions, two new quantities were introduced: Krylov expressivity and Krylov observability. These measures capture the effective dimension of the evolving Krylov spaces over time.

The Krylov-based measures were used to assess quantum reservoir performance. While Krylov complexity captured only the initial trends in task accuracy, Krylov observability closely reflected the full information processing capacity. This indicates that QRC effectively projects data into the Krylov basis, with observability serving as a reliable and interpretable indicator of computational power. Analysis across individual Hamiltonians further showed that low-performing reservoirs correspond to lower Krylov observability, highlighting a practical tool for system design. These results validate the utility of Krylov complexity in evaluating and optimizing quantum learning systems.

To address the time complexity problem caused by repeated quantum measurements, a memory-restricted version of QRC was proposed. By exploiting the fading memory property of reservoir computers, the approach avoids full reinitialization of the quantum state at each time step, thereby improving both computational efficiency and task performance.

In the final part of the thesis, quantum information measures were employed to investigate phase transitions in disordered systems. Krylov complexity was shown to distinguish not only between localized and ergodic regimes but also between different forms of localization, including spin glass and many-body localized phases. A critical time scale derived from Krylov complexity was found to align with optimal task performance in QRC, suggesting broader relevance for these diagnostics in the study of quantum dynamics.

Together, these results establish new links between quantum complexity, machine learning, and dynamical phase transitions. The methods developed in this work not only advance the understanding of QRC but also offer generalizable tools for exploring structure in quantum systems.

Outlook

This thesis has explored the interplay between quantum dynamics, quantum reservoir computing (QRC), and quantum information measures. The primary goal was to develop tools that provide deeper insights into how quantum systems process information, particularly in the context of machine learning tasks such as chaotic time-series prediction. Building on these results, several promising directions for future work are outlined below.

Krylov Complexity in Dynamical Systems

The Krylov-based measures introduced here were shown to capture properties of homogeneous differential equations. A natural extension is to investigate whether Krylov complexity can be generalized to arbitrary dynamical systems, including nonlinear or chaotic systems. Furthermore, Krylov complexity may serve as useful tools for probing the phase space of classical reservoir computers.

Beyond the Krylov Basis

A major theoretical insight of this work is that the standard Krylov basis, constructed from repeated Hamiltonian powers, does not necessarily minimize Krylov spread complexity. This calls for a more systematic investigation into alternative basis constructions and their implications. A goal would be to develop criteria or proofs for when a particular basis yields minimal complexity. An important point to consider is that such a basis may not preserve sensitivity to integrability or chaos, which could limit its usefulness in characterizing dynamical phases. But also the other case might very much be the case.

Applications in Quantum Machine Learning

The Krylov measures proposed here offer immediate applicability to other areas of quantum machine learning. For instance, Krylov expressivity can be used to evaluate the effectiveness of quantum data encodings, where input states $|x\rangle = U_E(x)\,|s\rangle$ are generated by parameterized circuits. Comparing expressivity across different encoding strategies could guide the design of more expressive quantum models. Krylov observability may also offer insight into the informativeness of various measurement schemes and could help identify configurations prone to barren plateaus. These tools are model-agnostic and require no additional assumptions, making them broadly useful.

Quantum Phase Transitions and Krylov Complexity

Initial results showed that Krylov complexity is sensitive to both the localized-to-ergodic transition and distinctions between different localized regimes, such as spin glass and many-body localization. Notably, this was observed even in small systems when time was rescaled by the Heisenberg time. Extending this analysis to larger systems would help establish the robustness of Krylov-based diagnostics for phase transitions. This could open a path toward more refined characterizations of disordered quantum systems, complementing traditional spectral approaches.

In summary, the methods developed in this thesis suggest that Krylov complexity and its associated measures offer a promising framework for analyzing the structure of quantum dynamics. They have the potential to impact multiple areas, including quantum information theory, machine learning, and the study of dynamical systems—quantum and classical alike.

Appendix A – Loschmidt Echo and Out of Order Time Correlators

Loschmidt Echos of the Ising Model

Figure A.1 illustrates the results for the Loschmidt echo of the Ising model. The first row (Fig. A.1a–c) presents the Loschmidt echo Lyapunov exponent λ_L for system sizes of four (a), six (b), and eight (c) qubits, while the second row (Fig. A.1d–f) shows the level spacing ratio $\langle r \rangle$ for the same system sizes.

To compute the Loschmidt Lyapunov exponent, an exponential fit is typically required, which is often numerically challenging. Therefore, the following steps are used to estimate λ_L using a linear fit:

1. Determine the saturation value of the Loschmidt Echo, denoted as G_s, such that $G(t) \to G_s$ as $t \to \infty$. If $G(t)$ remains close to one, the system cannot be considered chaotic, as the forward and backward time evolutions stay nearly indistinguishable. To account for this, introduce a threshold: if $G(t) > \alpha$, with $\alpha = 0.9$, then set $\lambda_G = 0$.

2. Identify the fitting time t_{fit}, defined by $G(t_{\text{fit}}) = G_s/\sqrt{2}$.

3. For $t \leq t_{\text{fit}}$, take the natural logarithm of the Loschmidt echo, assuming

$$G(t) \approx e^{-\lambda_L t} \Rightarrow y(t) = \ln G(t) = -\lambda_L t.$$

4. Perform a linear fit on $y(t)$ in the interval $[0, t_{\text{fit}}]$. The slope yields λ_L:

$$y(t) \approx -\lambda_L t + C.$$

This method extracts λ_L as the rate of exponential decay during the early-time dynamics, prior to saturation.

The Loschmidt echo Lyapunov exponent λ_L appears larger in the localized phase (where $\langle r \rangle \approx 0.38$), suggesting increased sensitivity to perturbations. However, λ_L changes continuously with respect to h and W, implying that only subtle changes in effective behavior are observed—except near $h \approx 10^0$, where λ_L changes drastically, aligning with the shift in level spacing statistics to $\langle r \rangle \approx 0.54$. Notably, in Fig. A.1(c),

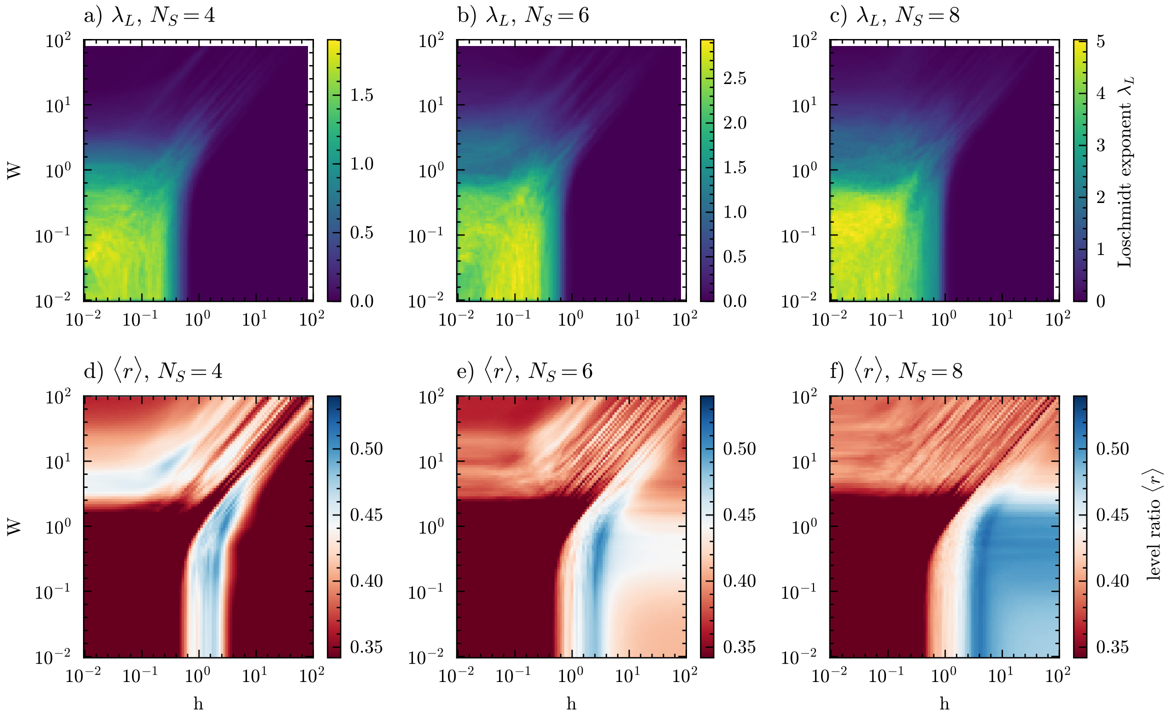


Figure A.1: Loschmidt echo Lyapunov exponent λ_L for the transverse-field Ising model as a function of disorder strength W and external field strength h. The figure presents results for system sizes of four (a), six (b), and eight (c) qubits. The level spacing ratio $\langle r \rangle$ is shown in the second row for four (d), six (e), and eight (f) qubits. At the transition from localized to ergodic phases—indicated by the change from $\langle r \rangle \approx 0.38$ to $\langle r \rangle \approx 0.54$—the Loschmidt Lyapunov exponent λ_L exhibits small value.

λ_L exhibits oscillations as a function of h, implying a highly sensitive dependence on initial conditions, indicative of quantum chaos.

Another transition is observed around $W \approx 10^0$, where λ_L decreases significantly, signaling a dynamical shift and further reinforcing the connection between spectral and dynamical properties during the localization-to-ergodicity transition.

Figure A.2 shows the behavior of the Loschmidt echo $G(t)$ across various points in phase space. In the localized phase, $G(t)$ decays quickly and remains near zero, indicating strong localization and irreversible quantum evolution. In the transition region, oscillations emerge, reflecting competition between localization and ergodicity. In the ergodic phase, $G(t)$ approaches a stable, near-constant value, demonstrating recoverable quantum information and high reversibility.

OTOCs of the Ising Model

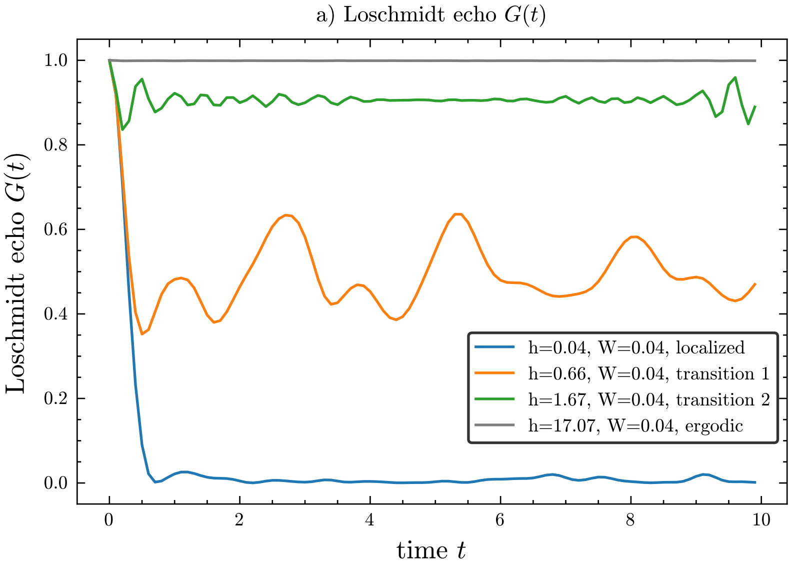


Figure A.2: The Loschmidt echo $G(t)$ is shown for different regions in phase space, illustrating the transition from localization to ergodicity. In the localized phase (blue), it decays to zero. Near the transition (orange), it oscillates around $G(t) \approx 0.5$. Deeper into the ergodic phase (green), it oscillates slightly around $G(t) \approx 0.9$, and in the fully ergodic phase (gray), it remains unchanged with $G(t) = 1$.

Figure A.3 shows the Lyapunov exponent of the out-of-time-ordered correlator (OTOC). The OTOC $C(t)$ is given by:

$$C(t) = -\langle [A(t), B(0)]^\dagger [A(t), B(0)] \rangle, \tag{A.1}$$

where time evolution is defined as:

$$O(t) = e^{iHt} O(0) e^{-iHt}. \tag{A.2}$$

The OTOC Lyapunov exponentλ_O is computed as follows

1. Determine the saturation value of the Loschmidt echo, denoted as OTOC_s, such that $\mathrm{OTOC}(t) \to \mathrm{OTOC}_s$ as $t \to \infty$. If the OTOC remains close to zero, the system cannot be considered chaotic, since the forward and backward time evolutions remain nearly identical. To account for this, introduce a threshold: if $\mathrm{OTOC}(t) < \beta$, with $\beta = 0.05$, then set $\lambda_O = 0$.

2. Identify the fitting time t_{fit}, defined by $\mathrm{OTOC}(t_{\text{fit}}) = \mathrm{OTOC}_s/\sqrt{2}$.

3. For $t \leq t_{\text{fit}}$, take the natural logarithm of the Loschmidt echo, assuming

$$\text{OTOC}(t) \propto e^{-2\lambda_L t} \quad \Rightarrow \quad y(t) = \ln\Big(\text{OTOC}(t)\Big) = 2\lambda_L t.$$

4. Perform a linear fit on $y(t)$ in the interval $[0, t_{\text{fit}}]$. The slope yields λ_L:

$$y(t) \approx 2\lambda_L t + C.$$

Similar to the computation of the Loschmidt exponent λ_L, the logarithm of $O(t)$ is taken, followed by a linear fit. Since $C(0) = 0$ is possible when the two observables commute, the fit must start at a minimum time t_{min} to ensure proper evaluation. In this analysis, the choice of observables is $A(0) = \sigma_x^{(1)}$ and $B(0) = \sigma_x^{(2)}$, allowing for the study of how local observable correlations evolve over time. The method can also be extended to global observables, such as magnetization, enabling a broader exploration of different system behaviors. The expectation value is computed with respect to the up state, defined as $|\text{up}\rangle = |00\ldots0\rangle$, resulting in the OTOC:

$$C(t) = \langle 00\ldots0| \, [\sigma_x^{(1)}(t), \sigma_x^{(2)}]^\dagger [\sigma_x^{(1)}(t), \sigma_x^{(2)}] \, |00\ldots0\rangle \,, \tag{1.64}$$

as well as with respect to the ground state $|\text{gs}\rangle$ of the corresponding Hamiltonian:

$$C(t) = \langle \text{gs}| \, [\sigma_x^{(1)}(t), \sigma_x^{(2)}]^\dagger [\sigma_x^{(1)}(t), \sigma_x^{(2)}] \, |\text{gs}\rangle \,. \tag{1.65}$$

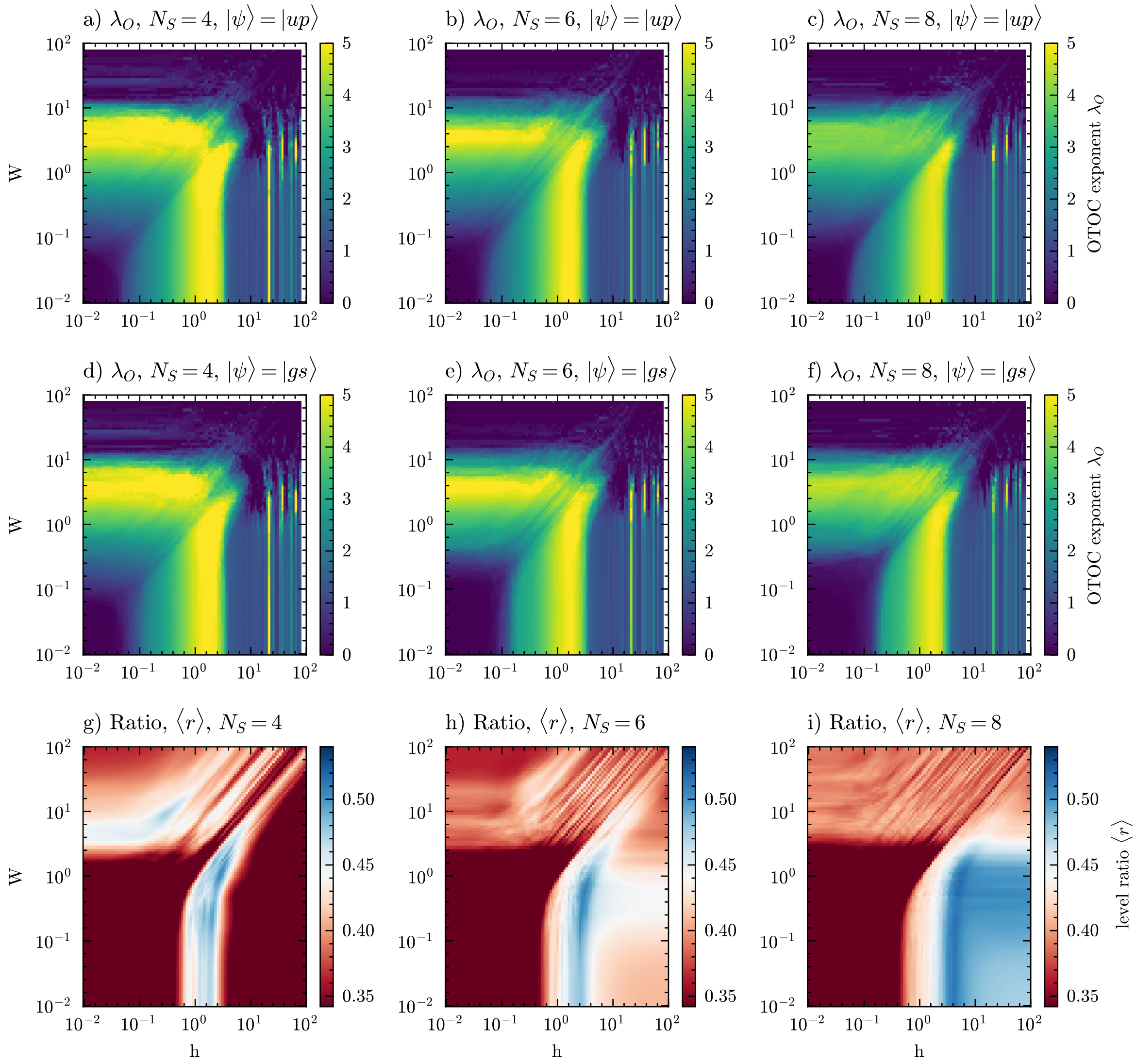


Figure A.3: OTOC Lyapunov exponent λ_O for different system sizes and initial states. The first row (a–c) shows results for the up state $|\text{up}\rangle$, the second row (d–f) for the ground state $|\text{gs}\rangle$, and the third row (g–i) shows level spacing ratios $\langle r\rangle$. Systems with four, six, and eight qubits are shown from left to right. λ_O peaks near the transition from localized to ergodic phases, indicating quantum chaos.

Another family of states that is often analyzed in relation to the OTOC is the Gibbs state, which results in the following OTOC:

$$C(t, \beta) = \text{Tr}\left(\frac{e^{-\beta H}}{\text{Tr}(e^{-\beta H})}[\sigma_x^{(1)}(t), \sigma_x^{(2)}]^\dagger [\sigma_x^{(1)}(t), \sigma_x^{(2)}]\right), \tag{1.66}$$

where $\beta = \frac{1}{k_B T}$ is the inverse temperature, with k_B being the Boltzmann constant. Although the thermal behavior of quantum systems is not analyzed in this study, its relevance is noted to illustrate the possible applications of OTOCs.

Figure A.3 presents the OTOC Lyapunov exponent λ_O for the up state $|\text{up}\rangle$ in the first row (a, b, c), for the ground state $|\text{gs}\rangle$ in the second row (d, e, f), and the level spacing ratio $\langle r \rangle$ in the third row (g, h, i), for different values of h and W. As in previous cases, the first, second, and third rows correspond to system sizes of four, six, and eight sites, respectively. The OTOC Lyapunov exponent λ_O exhibits small values in the localized phase (indicated by $\langle r \rangle \approx 0.38$). At the transition from the localized to the ergodic phase, around $h = 10^0$, λ_O reaches a maximum, similar to the behavior observed in the Loschmidt Lyapunov exponent. However, an additional region emerges where λ_O increases for increasing W, around $W \approx 10^0$ and $h \lesssim 0.5$. This suggests another phase transition, which is only weakly visible in the level spacing statistics (Figure A.3g–i). It is clearly evident that the OTOC successfully captures the phase transition and shows a maximum at the boundary between the localized and ergodic phases, which is precisely where quantum chaos is expected.

Appendix B – Engineering Quantum Reservoir Computing

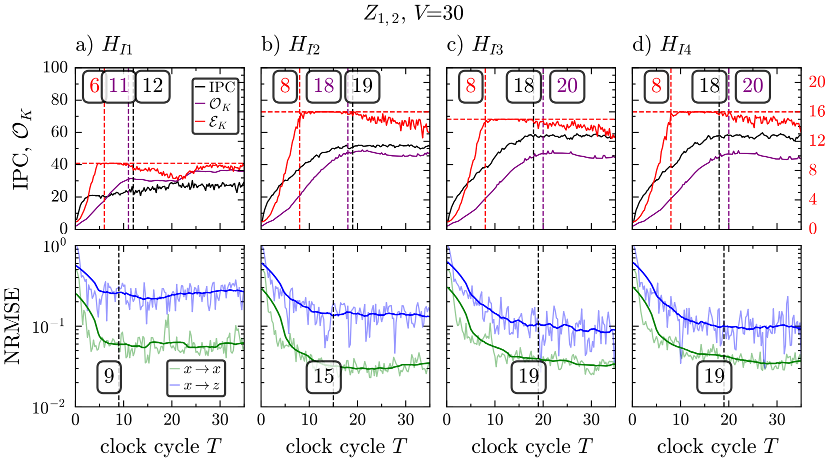


Figure B.1: Task performance of the Ising Hamiltonians with one observables measured.

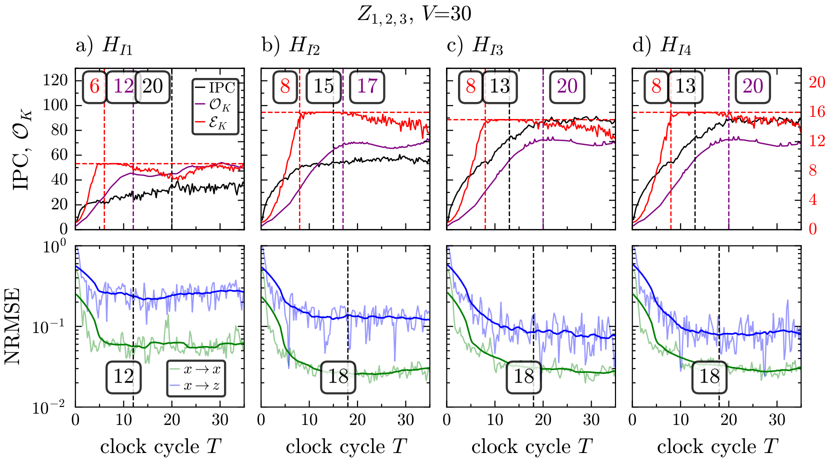


Figure B.2: Task performance of the Ising Hamiltonians with three observables measured.

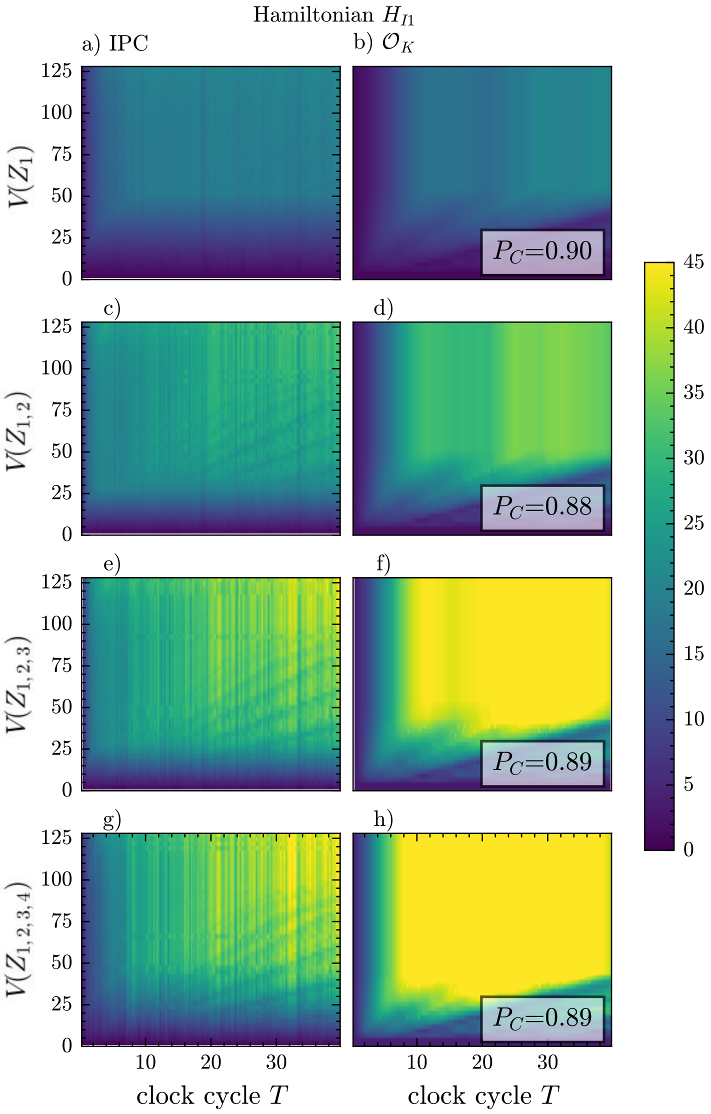


Figure B.3: Krylov observability $\mathcal{O}_K$ computed for the Ising Hamiltonian H_{I1} (Eq. (III.71)), when considering one (a), two (b), three (c) and four (d) observables in dependence of the number of measurements V and time t.

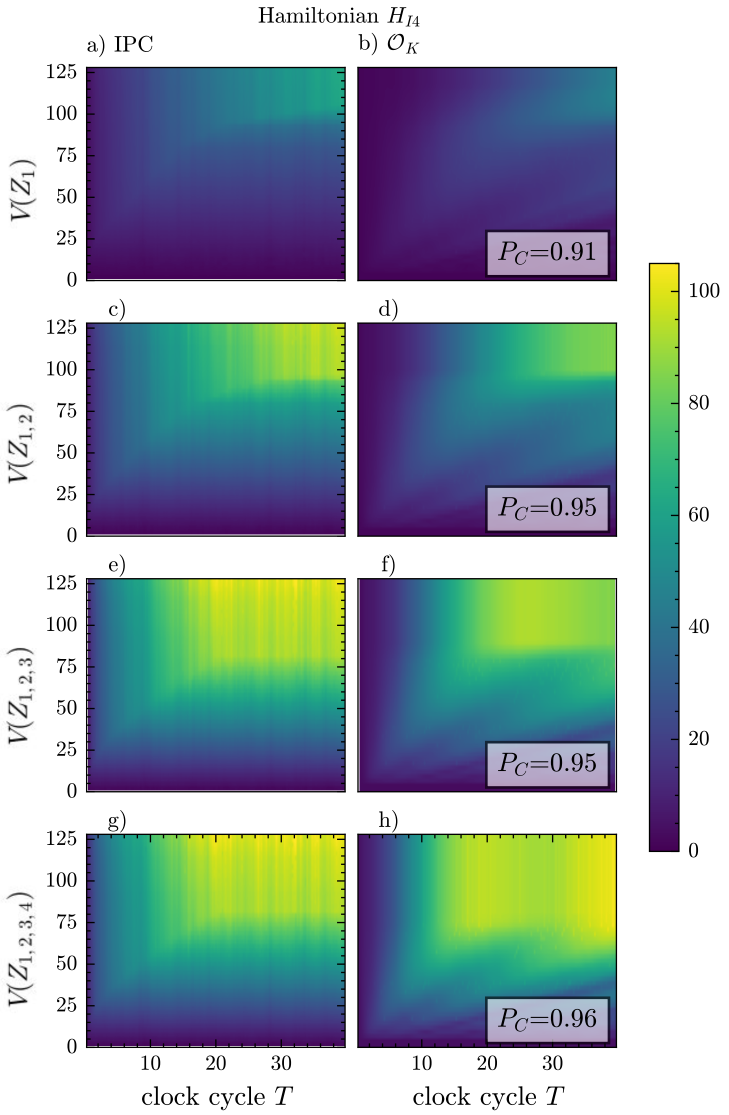


Figure B.4: Krylov observability $\mathcal{O}_K$ computed for the Ising Hamiltonian H_{I4} (Eq. (III.71)), when considering one (a), two (b), three (c) and four (d) observables in dependence of the number of measurements V and time t.

List of Figures

II.1 Simulation of three linear systems 17
II.2 Time evolution of the Lotka-Volterra system 21
II.3 Time evolution of the Lorenz system 22
II.4 Reservoir Computing Scheme 41
II.5 Quantum Reservoir Computing Performance in dependence of the number of spin sites 45
II.6 Quantum Reservoir Computing Performance in dependence of the number of measurements 45
II.7 Information processing capacity and task performance in dependence of the clock cycle 46
III.1 Spread complexity and amplitudes using Lanczos algorithm 54
III.2 Spread complexity and amplitudes using Arnoldi algorithm 55
III.3 Sketch Lemma 1 57
III.4 Sketch Theo. 3 61
III.5 Sketch Theo. 4 62
III.6 Spread complexity using measurable Krylov basis with $T_G = 10$ 63
III.7 Spread complexity using measurable Krylov basis with $T_G = 40$ 64
III.8 Spread complexity using measurable Krylov basis with $T_G = 60$ 65
III.9 Comparison of spread complexities from Krylov spaces K_m and G_m . . . 65
III.10Comparison of Krylov expressivities and spread complexities 67
III.11Operator complexity using powers of the Liouvillian 72
III.12Operator complexity via $\mathcal{F}_M$ basis with $T_G = 10$ 76
III.13Operator complexity via $\mathcal{F}_M$ basis with $T_G = 40$ 77
III.14Operator complexity via $\mathcal{F}_M$ basis with $T_G = 80$ 77
III.15Comparison of operator complexities via $\mathcal{L}^k$ and $\mathcal{F}_M$ 78
III.16Krylov observability for increasing numbers of measured observables . . 82
IV.1 Lorenz task performance, fidelity, and Krylov metrics as a function of clock cycle 92
IV.2 Average Krylov observability and IPC as a function of clock cycle 94
IV.3 Average Krylov observability and IPC as a function of the number of measurements 95

IV.4 Krylov observability 98
IV.5 Expressivity, Observability, and Lorenz Task Performance 104
IV.6 Expressivity, Observability, and Lorenz Task Performance when all sites are measured 105
IV.7 Observability 108
IV.8 Observability 109
IV.9 Information processing capacities of the Ising reservoir 116
IV.10 Influence of initial reservoir states on IPC_1 120
IV.11 Reset length optimization for Lorenz prediction tasks 121
IV.12 Performance scaling of LCQA with increasing reservoir size 122
V.1 Level spacing statistics 131
V.2 Krylov spread complexity across the Spin Glass–ergodic transition . . . 137
V.3 Rescaled Krylov spread complexity across the Spin Glass–ergodic transition 139
V.4 Krylov spread complexity across the MBL -ergodic transition 140
V.5 Rescaled Krylov spread complexity across the Spin MBL–ergodic transition 141
V.6 Krylov spreading time 143
V.7 Information Processing Capacity and Quantum Phases for four sites. . . 146
V.8 Information Processing Capacity and Quantum Phases for six sites. . . . 148
A.1 Loschmidt Echo 156
A.2 Evolution of the Loschmidt Echo 157
A.3 OTOC exponent 159
B.1 Task performance of the Ising Hamiltonians with one observables measured. 161
B.2 Task performance of the Ising Hamiltonians with three observables measured. 161
B.3 $\mathcal{O}_K$ and IPC for H_{I1} as functions of V and T 162
B.4 $\mathcal{O}_K$ and IPC for H_{I4} as functions of V and T 163

List of Tables

II.1 Selecting targets for computing the information processing capacity. . . 37
IV.1 Summary: Krylov spaces for pure state evolution 85
IV.2 Summary: Krylov spaces for operator evolution 86
IV.3 Summary: Quantum Reservoir Computing 87
IV.4 The number of pairwise distinct eigenvalues d and the inter-spin couplings J_{ij} for the different Ising reservoirs. 104
IV.5 Saturation values of NRMSE for the Lorenz task for Hamiltonians H_{I1} to H_{I4}, comparing single-site measurement Z_1 (rows one and three) to measurement of all sites (rows two and four). 107
IV.6 Initial states used for the reservoir, with corresponding definitions of the starting density matrix $\rho(0)$. 118

Bibliography

[1] O. Bohigas, M. J. Giannoni, and C. Schmit. Characterization of Chaotic Quantum Spectra and Universality of Level Fluctuation Laws. *Phys. Rev. Lett.*, 52(1):1–4, January 1984. doi: 10.1103/physrevlett.52.1. URL `https://link.aps.org/doi/10.1103/PhysRevLett.52.1`.

[2] J. M. Deutsch. Quantum statistical mechanics in a closed system. *Phys. Rev. A*, 43(4):2046–2049, February 1991. doi: 10.1103/physreva.43.2046. URL `https://link.aps.org/doi/10.1103/PhysRevA.43.2046`.

[3] M. Srednicki. The approach to thermal equilibrium in quantized chaotic systems. *J. Phys. A: Math. Gen.*, 32(7):1163, February 1999. ISSN 0305-4470. doi: 10.1088/0305-4470/32/7/007. URL `https://dx.doi.org/10.1088/0305-4470/32/7/007`.

[4] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn. Colloquium: Many-body localization, thermalization, and entanglement. *Rev. Mod. Phys.*, 91(2):021001, May 2019. doi: 10.1103/revmodphys.91.021001. URL `https://link.aps.org/doi/10.1103/RevModPhys.91.021001`.

[5] B. Žunkovič, M. Heyl, M. Knap, and A. Silva. Dynamical Quantum Phase Transitions in Spin Chains with Long-Range Interactions: Merging Different Concepts of Nonequilibrium Criticality. *Phys. Rev. Lett.*, 120(13):130601, March 2018. doi: 10.1103/physrevlett.120.130601. URL `https://link.aps.org/doi/10.1103/PhysRevLett.120.130601`.

[6] D. A. Huse, R. Nandkishore, V. Oganesyan, A. Pal, and S. L. Sondhi. Localization-protected quantum order. *Phys. Rev. B*, 88(1):014206, July 2013. doi: 10.1103/physrevb.88.014206. URL `https://link.aps.org/doi/10.1103/PhysRevB.88.014206`.

[7] R. A. Jalabert and H. M. Pastawski. Environment-Independent Decoherence Rate in Classically Chaotic Systems. *Phys. Rev. Lett.*, 86(12):2490–2493, March 2001. doi: 10.1103/physrevlett.86.2490. URL `https://link.aps.org/doi/10.1103/PhysRevLett.86.2490`.

[8] L. Viola, E. Knill, and S. Lloyd. Dynamical Decoupling of Open Quantum Systems. *Phys. Rev. Lett.*, 82(12):2417–2421, March 1999. doi: 10.1103/physrevlett.82.2417. URL https://link.aps.org/doi/10.1103/PhysRevLett.82.2417.

[9] T. R. Nielsen, J. Seebeck, M. Lorke, P. Gartner, and F. Jahnke. Microscopic theory of carrier-carrier scattering in nitride based quantum-dot systems. *Phys. Status Solidi C*, 3(6), 2006.

[10] S. H. Shenker and D. Stanford. Black holes and the butterfly effect. *J. High Energ. Phys.*, 2014(3):67, March 2014. ISSN 1029-8479. doi: 10.1007/jhep03(2014)067. URL https://doi.org/10.1007/JHEP03(2014)067.

[11] S. H. Shenker and D. Stanford. Multiple shocks. *J. High Energ. Phys.*, 2014 (12):46, December 2014. ISSN 1029-8479. doi: 10.1007/jhep12(2014)046. URL https://doi.org/10.1007/JHEP12(2014)046.

[12] S. H. Shenker and D. Stanford. Stringy effects in scrambling. *J. High Energ. Phys.*, 2015(5):132, May 2015. ISSN 1029-8479. doi: 10.1007/jhep05(2015)132. URL https://doi.org/10.1007/JHEP05(2015)132.

[13] A. Kitaev. A simple model of quantum holography, talk given at kitp program: Entanglement in strongly-correlated quantum matter, vol. 7 (usa april, 2025). URL https://online.kitp.ucsb.edu/online/entangled15/kitaev/.

[14] D. E. Parker, X. Cao, A. Avdoshkin, T. Scaffidi, and E. Altman. A universal operator growth hypothesis. *Phys. Rev. X*, 9:041017, Oct 2019. doi: 10.1103/physrevx.9.041017. URL https://link.aps.org/doi/10.1103/PhysRevX.9.041017.

[15] E. Rabinovici, A. Sánchez-Garrido, R. Shir, and J. Sonner. Krylov localization and suppression of complexity. *J. High Energy Phys.*, 2022(3):211, March 2022. ISSN 1029-8479. doi: 10.1007/jhep03(2022)211. URL https://link.springer.com/10.1007/JHEP03(2022)211.

[16] V. Balasubramanian, P. Caputa, J. M. Magan, and Q. Wu. Quantum chaos and the complexity of spread of states. *Phys. Rev. D*, 106(4):046007, August 2022. ISSN 2470-0010, 2470-0029. doi: 10.1103/physrevd.106.046007. URL https://link.aps.org/doi/10.1103/PhysRevD.106.046007.

[17] M. A. Nielsen, M. R. Dowling, M. Gu, and A. C. Doherty. Quantum Computation as Geometry. *Science*, 311(5764):1133–1135, February 2006. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1121541. URL https://www.science.org/doi/10.1126/science.1121541.

[18] P. W. Shor. *Algorithms for quantum computation: discrete logarithms and factoring*. IEEE, November 1994. ISBN 0-8186-6580-7. doi: 10.1109/sfcs.1994.365700. URL https://ieeexplore.ieee.org/document/365700.

[19] H. Jaeger. The "echo state" approach to analysing and training recurrent neural networks. GMD Report 148, GMD - German National Research Institute for Computer Science, 2001.

[20] W. Maass, T. Natschläger, and H. Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural. Comput.*, 14(11):2531, November 2002. ISSN 0899-7667. doi: 10.1162/089976602760407955.

[21] D. Verstraeten, B. Schrauwen, M. D'Haene, and D. Stroobandt. An experimental unification of reservoir computing methods. *Neural Networks*, 20(3):391–403, 2007. ISSN 0893-6080. doi: https://doi.org/10.1016/j.neunet.2007.04.003. URL https://www.sciencedirect.com/science/article/pii/S089360800700038X. Echo State Networks and Liquid State Machines.

[22] L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer. Information processing using a single dynamical node as complex system. *Nat. Commun.*, 2:468, 2011. doi: 10.1038/ncomms1476.

[23] L. Mühlnickel, J. A. Jaurigue, L. C. Jaurigue, and K. Lüdge. Reservoir computing using spin-vcsels - the influence of timescales and data injection schemes. *Commun. Phys.*, 7:370, 2024. ISSN 2399-3650. doi: 10.1038/s42005-024-01858-5. URL https://www.nature.com/articles/s42005-024-01858-5.

[24] L. C. Jaurigue and K. Lüdge. Reducing reservoir computer hyperparameter dependence by external timescale tailoring. *Neuromorph. Comput. Eng.*, 4(1):014001, 2024. doi: 10.1088/2634-4386/ad1d32. URL https://dx.doi.org/10.1088/2634-4386/ad1d32.

[25] J. A. Jaurigue, J. Robertson, A. Hurtado, L. C. Jaurigue, and K. Lüdge. Post-processing methods for delay embedding and feature scaling of reservoir computers. *Commun. Eng.*, 4:10, 2025. doi: 10.1038/s44172-024-00330-0. URL https://www.nature.com/articles/s44172-024-00330-0.

[26] K. Fujii and K. Nakajima. Harnessing Disordered-Ensemble Quantum Dynamics for Machine Learning. *Phys. Rev. Applied*, 8(2):024030, August 2017. doi:

10.1103/physrevapplied.8.024030. URL https://link.aps.org/doi/10.1103/PhysRevApplied.8.024030. Publisher: American Physical Society.

[27] K. Nakajima, K. Fujii, M. Negoro, K. Mitarai, and M. Kitagawa. Boosting Computational Power through Spatial Multiplexing in Quantum Reservoir Computing. *Phys. Rev. Applied*, 11(3):034021, March 2019. doi: 10.1103/physrevapplied.11.034021. URL https://link.aps.org/doi/10.1103/PhysRevApplied.11.034021. Publisher: American Physical Society.

[28] A. Kutvonen, K. Fujii, and T. Sagawa. Optimizing a quantum reservoir computer for time series prediction. *Sci. Rep.*, 10(1):14687, 2020. ISSN 2045-2322. doi: 10.1038/s41598-020-71673-9. URL https://doi.org/10.1038/s41598-020-71673-9.

[29] R. Martínez-Peña, G. L. Giorgi, J. Nokkala, M. C. Soriano, and R. Zambrini. Dynamical Phase Transitions in Quantum Reservoir Computing. *Phys. Rev. Lett.*, 127(10):100502, August 2021. ISSN 0031-9007, 1079-7114. doi: 10.1103/physrevlett.127.100502. URL https://link.aps.org/doi/10.1103/PhysRevLett.127.100502.

[30] W. Xia, J. Zou, X. Qiu, and X.F. Li. The reservoir learning power across quantum many-body localization transition. *Front. Phys.*, 17(3):33506, 2022. ISSN 2095-0470. doi: 10.1007/s11467-022-1158-1. URL https://doi.org/10.1007/s11467-022-1158-1.

[31] N. Götting, F. Lohof, and C. Gies. Exploring quantumness in quantum reservoir computing. *Phys. Rev. A*, 108(5):052427, November 2023. doi: 10.1103/physreva.108.052427. URL https://link.aps.org/doi/10.1103/PhysRevA.108.052427.

[32] J. Nokkala, G. L. Giorgi, and R. Zambrini. Retrieving past quantum features with deep hybrid classical-quantum reservoir computing. *ArXiv*, 2401.16961, 2024. doi: https://doi.org/10.48550/arXiv.2401.16961.

[33] A. H. Abbas, H. Abdel-Ghani, and I. S. Maksymov. Classical and Quantum Physical Reservoir Computing for Onboard Artificial Intelligence Systems: A Perspective. *Preprints*, June 2024. doi: 10.20944/preprints202406.1128.v1. URL https://www.preprints.org/manuscript/202406.1128/v1.

[34] J. Chen, H. I. Nurdin, and N. Yamamoto. Temporal information processing on noisy quantum computers. *Phys. Rev. Applied*, 14:024065, August 2020. doi: 10.1103/physrevapplied.14.024065.

[35] Y. Suzuki, Q. Gao, K. C. Pradel, K. Yasuoka, and N. Yamamoto. Natural quantum reservoir computing for temporal information processing. *Sci. Rep.*, 12 (1):1353, January 2022. ISSN 2045-2322. doi: 10.1038/s41598-022-05061-w. URL `https://www.nature.com/articles/s41598-022-05061-w`. Publisher: Nature Publishing Group.

[36] P. Pfeffer, F. Heyder, and J. Schumacher. Hybrid quantum-classical reservoir computing of thermal convection flow. *Phys. Rev. Research*, 4:033176, September 2022. doi: 10.1103/physrevresearch.4.033176.

[37] T. Yasuda, Y. Suzuki, T. Kubota, K. Nakajima, Q. Gao, W. Zhang, S. Shimono, H. I. Nurdin, and N. Yamamoto. Quantum reservoir computing with repeated measurements on superconducting devices. *arxiv*, October 2023. URL `http://arxiv.org/abs/2310.06706`.

[38] A. Sannia, R. Martínez-Peña, M. C. Soriano, G. L. Giorgi, and R. Zambrini. Dissipation as a resource for Quantum Reservoir Computing. *Quantum*, 8:1291, March 2024. doi: 10.22331/q-2024-03-20-1291. URL `https://quantum-journal.org/papers/q-2024-03-20-1291/`.

[39] D. Fry, A. Deshmukh, S. Y. Chen, V. Rastunkov, and V. Markov. Optimizing quantum noise-induced reservoir computing for nonlinear and chaotic time series prediction. *Sci. Rep.*, 13(1):19326, November 2023. ISSN 2045-2322. doi: 10.1038/s41598-023-45015-4. URL `https://www.nature.com/articles/s41598-023-45015-4`.

[40] L. Domingo, G. Carlo, and F. Borondo. Taking advantage of noise in quantum reservoir computing. *Sci. Rep.*, 13(1):8790, May 2023. ISSN 2045-2322. doi: 10.1038/s41598-023-35461-5. URL `https://www.nature.com/articles/s41598-023-35461-5`.

[41] P. Mujal, R. Martínez-Peña, G. L. Giorgi, M. C. Soriano, and R. Zambrini. Time-series quantum reservoir computing with weak and projective measurements. *npj Quantum Inf.*, 9(1):16, 2023. ISSN 2056-6387. URL `https://doi.org/10.1038/s41534-023-00682-z`.

[42] G. Franceschetto, M. Płodzień, M. Lewenstein, A. Acín, and P. Mujal. Harnessing quantum back-action for time-series processing. *ArXiv*, 2411.03979, November 2024. doi: 10.48550/arxiv.2411.03979. URL `http://arxiv.org/abs/2411.03979`.

[43] S. Čindrak, B. Donvil, K. Lüdge, and L. C. Jaurigue. Enhancing the performance of quantum reservoir computing and solving the time-complexity problem

by artificial memory restriction. *Phys. Rev. Res.*, 6:013051, Jan 2024. doi: 10.1103/physrevresearch.6.013051. URL https://link.aps.org/doi/10.1103/PhysRevResearch.6.013051. ArXiv 2306.12876.

[44] K. Kobayashi, K. Fujii, and N. Yamamoto. Feedback-Driven Quantum Reservoir Computing for Time-Series Analysis. *PRX Quantum*, 5(4):040325, November 2024. doi: 10.1103/prxquantum.5.040325. URL https://link.aps.org/doi/10.1103/PRXQuantum.5.040325.

[45] O. Ahmed, F. Tennie, and L. Magri. Prediction of chaotic dynamics and extreme events: A recurrence-free quantum reservoir computing approach. *Phys. Rev. Research*, 6(4):043082, November 2024. ISSN 2643-1564. doi: 10.1103/physrevresearch.6.043082. URL https://link.aps.org/doi/10.1103/PhysRevResearch.6.043082.

[46] C. Zhu, P. J. Ehlers, H. I. Nurdin, and D. Soh. Practical and scalable quantum reservoir computing. *ArXiv*, 2405.04799, 2024. URL https://arxiv.org/abs/2405.04799.

[47] N. Schütte, N. Götting, H. Müntinga, M. List, and C. Gies. Expressive Limits of Quantum Reservoir Computing. *arXiv*, January 2025. doi: 10.48550/arxiv.2501.15528. URL http://arxiv.org/abs/2501.15528.

[48] S. H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering.* CRC Press, Boca Raton, 2nd edition edition, 2015. ISBN 9780429492563. doi: 10.1201/9780429492563. URL https://doi.org/10.1201/9780429492563. eBook Published 23 May 2019.

[49] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20:130, 1963.

[50] A. J. Lotka. Analytical note on certain rhythmic relations in organic systems. *Proc. Natl. Acad. Sci. U.S.A.*, 6(410), 1920.

[51] J. Dambre, D. Verstraeten, B. Schrauwen, and S. Massar. Information processing capacity of dynamical systems. *Sci. Rep.*, 2:514, 2012. doi: 10.1038/srep00514.

[52] D. Verstraeten, J. Dambre, X. Dutoit, and B. Schrauwen. Memory versus non-linearity in reservoirs. In *The 2010 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2010. doi: 10.1109/ijcnn.2010.5596492.

[53] J. B. Butcher, D. Verstraeten, B. Schrauwen, C. R. Day, and P. W. Haycock. Reservoir computing and extreme learning machines for non-linear time-series

data analysis. *Neural Netw.*, 38:76–89, 2013. ISSN 0893-6080. doi: https://doi.org/10.1016/j.neunet.2012.11.011.

[54] M. Inubushi and K. Yoshimura. Reservoir computing beyond memory-nonlinearity trade-off. *Sci. Rep.*, 7:10199, 2017.

[55] K. Hashimoto, K. Murata, N. Tanahashi, and R. Watanabe. Krylov complexity and chaos in quantum mechanics. *J. High Energy Phys.*, 2023(11):40, November 2023. ISSN 1029-8479. doi: 10.1007/jhep11(2023)040. URL `https://doi.org/10.1007/JHEP11(2023)040`.

[56] S. He, P. H. C. Lau, Z. Xian, and L. Zhao. Quantum chaos, scrambling and operator growth in $$ T\overline{T} $$deformed SYK models. *J. High Energy Phys.*, 2022 (12):70, December 2022. ISSN 1029-8479. doi: 10.1007/jhep12(2022)070. URL `https://doi.org/10.1007/JHEP12(2022)070`.

[57] N. Iizuka and M. Nishida. Krylov complexity in the IP matrix model. *arXiv*, June 2023. URL `http://arxiv.org/abs/2306.04805`. arXiv:2306.04805 [hep-th, physics:quant-ph].

[58] S. Jian, B. Swingle, and Z. Xian. Complexity growth of operators in the SYK model and in JT gravity. *J. High Energy Phys.*, 2021(3):14, March 2021. ISSN 1029-8479. doi: 10.1007/jhep03(2021)014. URL `https://doi.org/10.1007/JHEP03(2021)014`.

[59] E. Rabinovici, A. Sánchez-Garrido, R. Shir, and J. Sonner. Krylov complexity from integrability to chaos. *J. High Energy Phys.*, 2022(7):151, July 2022. ISSN 1029-8479. doi: 10.1007/jhep07(2022)151. URL `https://doi.org/10.1007/JHEP07(2022)151`.

[60] J. Erdmenger, S. Jian, and Z. Xian. Universal chaotic dynamics from Krylov space. *J. High Energy Phys.*, 2023(8):176, August 2023. ISSN 1029-8479. doi: 10.1007/jhep08(2023)176. URL `https://link.springer.com/10.1007/JHEP08(2023)176`.

[61] V. Balasubramanian, J. M. Magan, and Q. Wu. Quantum chaos, integrability, and late times in the Krylov basis. *arXiv*, December 2023. doi: 10.48550/arxiv.2312.03848. URL `http://arxiv.org/abs/2312.03848`. arXiv:2312.03848 [cond-mat, physics:hep-th, physics:nlin, physics:quant-ph].

[62] H. W. Lin. The bulk Hilbert space of double scaled SYK. *J. High Energy Phys.*, 2022(11):60, November 2022. ISSN 1029-8479. doi: 10.1007/jhep11(2022)060. URL `https://doi.org/10.1007/JHEP11(2022)060`.

[63] Y. Zhou, W. Xia, L. Li, and W. Li. Diagnosing Quantum Many-body Chaos in Non-Hermitian Quantum Spin Chain via Krylov Complexity. *arXiv*, 2025. doi: 10.48550/arxiv.2501.15982. URL http://arxiv.org/abs/2501.15982.

[64] Z. Fan. Universal relation for operator complexity. *Phys. Rev. A*, 105(6):062210, June 2022. ISSN 2469-9926, 2469-9934. doi: 10.1103/physreva.105.062210. URL http://arxiv.org/abs/2202.07220. arXiv:2202.07220 [hep-th, physics:quant-ph].

[65] B. Craps, O. Evnin, and G. Pascuzzi. A relation between Krylov and Nielsen complexity. *arXiv*, November 2023. URL http://arxiv.org/abs/2311.18401. arXiv:2311.18401 [hep-th, physics:quant-ph].

[66] P. Caputa, J. M. Magan, and D. Patramanis. Geometry of Krylov complexity. *Phys. Rev. Research*, 4(1):013041, January 2022. ISSN 2643-1564. doi: 10.1103/physrevresearch.4.013041. URL https://link.aps.org/doi/10.1103/PhysRevResearch.4.013041.

[67] T. Anegawa, N. Iizuka, and M. Nishida. Krylov complexity as an order parameter for deconfinement phase transitions at large N. *J. High Energy Phys.*, 2024 (4):119, April 2024. ISSN 1029-8479. doi: 10.1007/jhep04(2024)119. URL https://doi.org/10.1007/JHEP04(2024)119.

[68] P. Caputa and S. Liu. Quantum complexity and topological phases of matter. *Phys. Rev. B*, 106(19):195125, November 2022. ISSN 2469-9950, 2469-9969. doi: 10.1103/physrevb.106.195125. URL https://link.aps.org/doi/10.1103/PhysRevB.106.195125.

[69] B. Bhattacharjee, P. Nandy, and T. Pathak. Operator dynamics in Lindbladian SYK: a Krylov complexity perspective. *J. High Energy Phys.*, 2024(1):94, January 2024. ISSN 1029-8479. doi: 10.1007/jhep01(2024)094. URL http://arxiv.org/abs/2311.00753. arXiv:2311.00753 [cond-mat, physics:hep-th, physics:quant-ph].

[70] J. Kim, J. Murugan, J. Olle, and D. Rosa. Operator delocalization in quantum networks. *Phys. Rev. A*, 105(1):L010201, January 2022. doi: 10.1103/physreva.105.l010201. URL https://link.aps.org/doi/10.1103/PhysRevA.105.L010201. Publisher: American Physical Society.

[71] C. Liu, H. Tang, and H. Zhai. Krylov complexity in open quantum systems. *Phys. Rev. Res.*, 5:033085, Aug 2023. doi: 10.1103/physrevresearch.5.033085. URL https://link.aps.org/doi/10.1103/PhysRevResearch.5.033085.

[72] A. A. Nizami and A. W. Shrestha. Krylov construction and complexity for driven quantum systems. *Phys. Rev. E*, 108(5):054222, November 2023. doi: 10.1103/physreve.108.054222. URL https://link.aps.org/doi/10.1103/PhysRevE.108.054222. Publisher: American Physical Society.

[73] H. A. Camargo, V. Jahnke, K. Kim, and M. Nishida. Krylov complexity in free and interacting scalar field theories with bounded power spectrum. *J. High Energy Phys.*, 2023(5):226, May 2023. ISSN 1029-8479. doi: 10.1007/jhep05(2023)226. URL https://link.springer.com/10.1007/JHEP05(2023)226.

[74] P. Caputa and S. Datta. Operator growth in 2d CFT. *J. High Energy Phys.*, 2021 (12):188, December 2021. ISSN 1029-8479. doi: 10.1007/jhep12(2021)188. URL http://arxiv.org/abs/2110.10519. arXiv:2110.10519 [cond-mat, physics:hep-th, physics:quant-ph].

[75] A. Dymarsky and M. Smolkin. Krylov complexity in conformal field theory. *Phys. Rev. D*, 104(8):L081702, October 2021. ISSN 2470-0010, 2470-0029. doi: 10.1103/physrevd.104.l081702. URL https://link.aps.org/doi/10.1103/PhysRevD.104.L081702.

[76] S. Guo. Operator growth in SU(2) Yang-Mills theory. *arXiv*, August 2022. URL http://arxiv.org/abs/2208.13362. arXiv:2208.13362 [hep-th].

[77] M. J. Vasli, K. B. Velni, M. R. M. Mozaffar, A. Mollabashi, and M. Alishahiha. Krylov complexity in Lifshitz-type scalar field theories. *EOJ C*, 84(3):235, March 2024. ISSN 1434-6052. doi: 10.1140/epjc/s10052-024-12609-9. URL https://doi.org/10.1140/epjc/s10052-024-12609-9.

[78] B. Bhattacharjee, S. Sur, and P. Nandy. Probing quantum scars and weak ergodicity breaking through quantum complexity. *Phys. Rev. B*, 106(20):205150, November 2022. ISSN 2469-9950, 2469-9969. doi: 10.1103/physrevb.106.205150. URL https://link.aps.org/doi/10.1103/PhysRevB.106.205150.

[79] B. Craps, O. Evnin, and G. Pascuzzi. Multiseed krylov complexity. *Phys. Rev. Lett.*, 134:050402, Feb 2025. doi: 10.1103/physrevlett.134.050402. URL https://link.aps.org/doi/10.1103/PhysRevLett.134.050402.

[80] P. Nandy, A. S. Matsoukas-Roubeas, P. Martínez-Azcona, A. Dymarsky, and A. d. Campo. Quantum Dynamics in Krylov Space: Methods and Applications. *arXiv*, 2405.09628, June 2024. doi: 10.48550/arxiv.2405.09628. URL http://arxiv.org/abs/2405.09628.

[81] S. Čindrak, A. Paschke, L. C. Jaurigue, and K. Lüdge. Measurable Krylov Spaces and Eigenenergy Count in Quantum State Dynamics. *J. High Energ. Phys.*, 2024: 83, April 2024. doi: 10.1007/JHEP10(2024)083.

[82] S. Čindrak, K. Lüdge, and L. C. Jaurigue. Quantum reservoir computing maps data onto the Krylov space. *arXiv*, January 2025. doi: 10.48550/arxiv.2502.12157. URL `http://arxiv.org/abs/2502.12157`. arXiv:2502.12157 [quant-ph].

[83] S. Čindrak, L. C. Jaurigue, and K. Lüdge. Engineering quantum reservoirs through krylov complexity, expressivity and observability. *arXiv*, February 2025. doi: 10.48550/arXiv.2409.12079. URL `https://arxiv.org/abs/2409.12079`.

[84] L. Domingo, F. Borondo, G. Scialchi, A. J. Roncaglia, G. G. Carlo, and D. A. Wisniacki. Quantum reservoir complexity by the Krylov evolution approach. *Phys. Rev. A*, 110(2):022446, August 2024. doi: 10.1103/physreva.110.022446. URL `https://link.aps.org/doi/10.1103/PhysRevA.110.022446`.

[85] M. Vetrano, G. Lo Monaco, L. Innocenti, S. Lorenzo, and G. M. Palma. State estimation with quantum extreme learning machines beyond the scrambling time. *npj Quantum Inf*, 11(1):1–8, February 2025. ISSN 2056-6387. doi: 10.1038/s41534-024-00927-5. URL `https://www.nature.com/articles/s41534-024-00927-5`.

[86] R. Martínez-Peña, J. Nokkala, G. L. Giorgi, R. Zambrini, and M. C. Soriano. Information Processing Capacity of Spin-Based Quantum Reservoir Computing Systems. *Cogn. Comput.*, 15(5):1440–1451, September 2020. ISSN 1866-9964. doi: 10.1007/s12559-020-09772-y. URL `https://doi.org/10.1007/s12559-020-09772-y`.

[87] M. Schuld and F. Petruccione. Quantum Science and Technology. Springer International Publishing, Cham, 2018. doi: 10.1007/978-3-319-96424-9. URL `https://link.springer.com/10.1007/978-3-319-96424-9`.

[88] B. Romeira, J. M. L. Figueiredo, and J. Javaloyes. Delay dynamics of neuromorphic optoelectronic nanoscale resonators: Perspectives and applications. *Chaos*, 27(11): 114323, November 2017. doi: 10.1063/1.5008888.

[89] V. Dunjko and H. J. Briegel. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Rep. Prog. Phys.*, 81(7):074001, June 2018. ISSN 0034-4885. doi: 10.1088/1361-6633/aab406. URL `https://dx.doi.org/10.1088/1361-6633/aab406`. Publisher: IOP Publishing.

[90] N. Killoran, T. R. Bromley, J. M. Arrazola, M. Schuld, N. Quesada, and S. Lloyd. Continuous-variable quantum neural networks. *Phys. Rev. Research*, 1(3):033063, October 2019. doi: 10.1103/physrevresearch.1.033063. URL `https://link.aps.org/doi/10.1103/PhysRevResearch.1.033063`. Publisher: American Physical Society.

[91] E. Schöll, J. Lehnert, A. Keane, T. Dahms, and P. Hövel. Control of desynchronization transitions in delay-coupled networks of type-I and type-II excitable systems, proc. of the international symposium, hanse institute of advanced studies, delmenhorst, 13-16 november, 2012. In A. Pelster and G. Wunner, editors, *Self-organization in Complex Systems: The Past, Present, and Future of Synergetics*, pages 25–42, Berlin, 2016. Springer. doi: 10.1007/978-3-319-27635-9_3. ISBN 978-3-319-27635-9.

[92] E. Farhi and H. Neven. Classification with Quantum Neural Networks on Near Term Processors. *arXiv*, 1802.06002, August 2018. doi: 10.48550/arxiv.1802.06002. URL `http://arxiv.org/abs/1802.06002`. arXiv:1802.06002 [quant-ph].

[93] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer. Parallel photonic information processing at gigabyte per second data rates using transient states. *Nat. Commun.*, 4:1364, January 2013. doi: 10.1038/ncomms2368.

[94] Z. Li, X. Liu, N. Xu, and J. Du. Experimental Realization of a Quantum Support Vector Machine. *Phys. Rev. Lett.*, 114(14):140504, April 2015. doi: 10.1103/physrevlett.114.140504. URL `https://link.aps.org/doi/10.1103/PhysRevLett.114.140504`. Publisher: American Physical Society.

[95] Y. Li, G. W. Holloway, S. C. Benjamin, G. A. D. Briggs, J. Baugh, and J. A. Mol. Double quantum dot memristor. *Phys. Rev. B*, 96(7):075446, August 2017. doi: 10.1103/physrevb.96.075446. URL `https://link.aps.org/doi/10.1103/PhysRevB.96.075446`. Publisher: American Physical Society.

[96] V. Saggio, B. E. Asenbeck, A. Hamann, T. Strömberg, P. Schiansky, V. Dunjko, N. Friis, N. C. Harris, M. Hochberg, D. Englund, S. Wölk, H. J. Briegel, and P. Walther. Experimental quantum speed-up in reinforcement learning agents. *Nature*, 591(7849):229–233, March 2021. ISSN 1476-4687. doi: 10.1038/s41586-021-03242-7. URL `https://www.nature.com/articles/s41586-021-03242-7`. Publisher: Nature Publishing Group.

[97] M. Schuld, R. Sweke, and J. J. Meyer. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Phys. Rev. A*, 103(3):

032430, March 2021. doi: 10.1103/physreva.103.032430. URL `https://link.aps.org/doi/10.1103/PhysRevA.103.032430`. Publisher: American Physical Society.

[98] P. Steinmüller, T. Schulz, F. Graf, and D. Herr. eXplainable AI for Quantum Machine Learning. *arXiv*, 2211.01441, November 2022. doi: 10.48550/arxiv.2211.01441. URL `http://arxiv.org/abs/2211.01441`.

[99] L. Pira and C. Ferrie. On the interpretability of quantum neural networks. *Quantum Mach. Intell.*, 6(2):52, August 2024. ISSN 2524-4914. doi: 10.1007/s42484-024-00191-y. URL `https://doi.org/10.1007/s42484-024-00191-y`.

[100] R. Lifshitz. Quantum Deep Dreaming: A Novel Approach for Quantum Circuit Design. *arXiv*, 2211.04343, November 2022. doi: 10.48550/arxiv.2211.04343. URL `http://arxiv.org/abs/2211.04343`.

[101] R. Heese, T. Gerlach, S. Mücke, S. Müller, M. Jakobs, and N. Piatkowski. Explaining Quantum Circuits with Shapley Values: Towards Explainable Quantum Machine Learning. *arXiv*, 2301.09138, March 2023. doi: 10.48550/arxiv.2301.09138. URL `http://arxiv.org/abs/2301.09138`.

[102] E. Gil-Fuster, J. R. Naujoks, G. Montavon, T. Wiegand, W. Samek, and J. Eisert. Opportunities and limitations of explaining quantum machine learning. *arXiv*, 2412.14753, December 2024. doi: 10.48550/arxiv.2412.14753. URL `http://arxiv.org/abs/2412.14753`.

[103] T. Y. Cheng, C. C. Liu, D. Y. Jhou, and C. C. Chen. Impact of coupling topology upon noise robustness of small optical reservoirs. *Sci. Rep.*, 10:14086, 2020. doi: 10.1038/s41598-020-70775-8.

[104] T. Butler, B. Tykalewicz, D. Goulding, B. Kelleher, G. Huyet, and S. P. Hegarty. Reconstruction of dynamical pulse trains via time-resolved multi-heterodyne detection. *Opt. Express*, 21(24):29109–29119, December 2013. doi: 10.1364/oe.21.029109.

[105] T. Hülser, F. Köster, L. C. Jaurigue, and K. Lüdge. Role of delay-times in delay-based photonic reservoir computing. *Opt. Mater. Express*, 12(3):1214, 2022. doi: 10.1364/ome.451016.

[106] J. Hüpfl, N. Bachelard, M. Kaczvinszki, M. Horodynski, M. Kühmayer, and S. Rotter. Optimal cooling of multiple levitated particles: Theory of far-field

wavefront shaping. *Phys. Rev. A*, 107:023112, Feb 2023. doi: 10.1103/physreva.107.023112. URL https://link.aps.org/doi/10.1103/PhysRevA.107.023112.

[107] M. Gutzwiller. Quantum chaos. *Scholarpedia*, 2(12):3146, December 2007. ISSN 1941-6016. doi: 10.4249/scholarpedia.3146. URL http://www.scholarpedia.org/article/Quantum_chaos.

[108] J. Tangpanitanon, S. Thanasilp, N. Dangniam, M. Lemonde, and D. G. Angelakis. Expressibility and trainability of parametrized analog quantum systems for machine learning applications. *Phys. Rev. Res.*, 2(4):043364, December 2020. doi: 10.1103/physrevresearch.2.043364. URL https://link.aps.org/doi/10.1103/PhysRevResearch.2.043364.

[109] B. Altshuler, H. Krovi, and J. Roland. Anderson localization makes adiabatic quantum optimization fail. *Proc. Natl. Acad. Sci. U.S.A.*, 107(28):12446–12450, July 2010. doi: 10.1073/pnas.1002116107. URL https://www.pnas.org/doi/full/10.1073/pnas.1002116107.

[110] C. Laumann, R. Moessner, A. Scardicchio, and S. Sondhi. Quantum annealing: The fastest route to quantum computation? *Eur. Phys. J. Spec. Top.*, 224(1):75–88, February 2015. ISSN 1951-6401. doi: 10.1140/epjst/e2015-02344-2. URL https://doi.org/10.1140/epjst/e2015-02344-2.

[111] J. P. Keating, N. Linden, J. C. F. Matthews, and A. Winter. Localization and its consequences for quantum walk algorithms and quantum communication. *Phys. Rev. A*, 76(1):012315, July 2007. doi: 10.1103/physreva.76.012315. URL https://link.aps.org/doi/10.1103/PhysRevA.76.012315.

[112] A. Schreiber, K. N. Cassemiro, V. Potoček, A. Gábris, I. Jex, and C. Silberhorn. Decoherence and Disorder in Quantum Walks: From Ballistic Spread to Localization. *Phys. Rev. Lett.*, 106(18):180403, May 2011. doi: 10.1103/physrevlett.106.180403. URL https://link.aps.org/doi/10.1103/PhysRevLett.106.180403.

[113] Vadim Oganesyan and David A. Huse. Localization of interacting fermions at high temperature. *Phys. Rev. B*, 75:155111, 2007. doi: 10.1103/physrevb.75.155111. URL https://link.aps.org/doi/10.1103/PhysRevB.75.155111.

[114] L. Boltzmann. *Über die Beziehung zwischen dem zweiten Hauptsatze des mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respective den Sätzen über das Wärmegleichgewicht.* K.k. Hof- und Staatsdruckerei, 1877.

[115] J. Loschmidt. *Über den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft: I [-IV].* aus der K.K. Hof - und Staatsdruckerei, 1876.

[116] W. Thomson. Kinetic Theory of the Dissipation of Energy. *Nature*, 9(232): 441–444, April 1874. ISSN 1476-4687. doi: 10.1038/009441c0. URL `https://www.nature.com/articles/009441c0`.

[117] Asher Peres. Stability of quantum motion in chaotic and regular systems. *Phys. Rev. A*, 30:1610–1615, Oct 1984. doi: 10.1103/physreva.30.1610. URL `https://link.aps.org/doi/10.1103/PhysRevA.30.1610`.

[118] E. L. Hahn. Spin Echoes. *Phys. Rev.*, 80(4):580–594, November 1950. doi: 10.1103/physrev.80.580. URL `https://link.aps.org/doi/10.1103/PhysRev.80.580`.

[119] W. Rhim, A. Pines, and J. S. Waugh. Violation of the Spin-Temperature Hypothesis. *Phys. Rev. Lett.*, 25(4):218–220, July 1970. doi: 10.1103/physrevlett.25.218. URL `https://link.aps.org/doi/10.1103/PhysRevLett.25.218`.

[120] S. Zhang, B. H. Meier, and R. R. Ernst. Polarization echoes in NMR. *Phys. Rev. Lett.*, 69(14):2149–2151, October 1992. doi: 10.1103/physrevlett.69.2149. URL `https://link.aps.org/doi/10.1103/PhysRevLett.69.2149`.

[121] P. R. Levstein, G. Usaj, and H. M. Pastawski. Attenuation of polarization echoes in nuclear magnetic resonance: A study of the emergence of dynamical irreversibility in many-body quantum systems. *J. Chem. Phys.*, 108(7):2718–2724, February 1998. ISSN 0021-9606. doi: 10.1063/1.475664. URL `https://doi.org/10.1063/1.475664`.

[122] L. Viola and E. Knill. Random Decoupling Schemes for Quantum Dynamical Control and Error Suppression. *Phys. Rev. Lett.*, 94(6):060502, February 2005. doi: 10.1103/physrevlett.94.060502. URL `https://link.aps.org/doi/10.1103/PhysRevLett.94.060502`.

[123] A. Peres. *Quantum Theory: Concepts and Methods.* Number 57 in Fundamental Theories of Physics. Imprint: Springer, Dordrecht, 1st ed. 2002 edition, 2002. ISBN 978-0-306-47120-9. doi: 10.1007/0-306-47120-5.

[124] T. Gorin, T. Prosen, T. H. Seligman, and M. Žnidarič. Dynamics of Loschmidt echoes and fidelity decay. *Phys. Rep.*, 435(2):33–156, November 2006. ISSN 0370-1573. doi: 10.1016/j.physrep.2006.09.003. URL `https://www.sciencedirect.com/science/article/pii/S0370157306003310`.

[125] Quasiclassical Method in the Theory of Superconductivity. *J. Exp. Theor. Phys.*, June .

[126] B. Swingle. Unscrambling the physics of out-of-time-order correlators. *Nature Phys.*, 14(10):988–990, October 2018. ISSN 1745-2481. doi: 10.1038/s41567-018-0295-5. URL https://www.nature.com/articles/s41567-018-0295-5.

[127] B. Yan, L. Cincio, and W. H. Zurek. Information Scrambling and Loschmidt Echo. *Phys. Rev. Lett.*, 124(16):160603, April 2020. doi: 10.1103/physrevlett.124.160603. URL https://link.aps.org/doi/10.1103/PhysRevLett.124.160603.

[128] A. Bhattacharyya, W. Chemissany, S. S. Haque, and B. Yan. Towards the web of quantum chaos diagnostics. *Eur. Phys. J. C*, 82(1):87, January 2022. ISSN 1434-6052. doi: 10.1140/epjc/s10052-022-10035-3. URL https://doi.org/10.1140/epjc/s10052-022-10035-3.

[129] Y. Sekino and L. Susskind. Fast scramblers. *J. High Energy Phys.*, 2008(10):065, October 2008. ISSN 1126-6708. doi: 10.1088/1126-6708/2008/10/065. URL https://dx.doi.org/10.1088/1126-6708/2008/10/065.

[130] R. A. Jalabert, I. García-Mata, and D. A. Wisniacki. Semiclassical theory of out-of-time-order correlators for low-dimensional classically chaotic systems. *Phys. Rev. E*, 98(6):062218, December 2018. doi: 10.1103/physreve.98.062218. URL https://link.aps.org/doi/10.1103/PhysRevE.98.062218.

[131] E. M. Fortes, I. García-Mata, R. A. Jalabert, and D. A. Wisniacki. Gauging classical and quantum integrability through out-of-time-ordered correlators. *Phys. Rev. E*, 100(4):042201, October 2019. doi: 10.1103/physreve.100.042201. URL https://link.aps.org/doi/10.1103/PhysRevE.100.042201.

[132] R. He and Z. Lu. Characterizing many-body localization by out-of-time-ordered correlation. *Phys. Rev. B*, 95(5):054201, February 2017. doi: 10.1103/physrevb.95.054201. URL https://link.aps.org/doi/10.1103/PhysRevB.95.054201.

[133] J. Garcia-Beni, G. L. Giorgi, M. C. Soriano, and R. Zambrini. Scalable photonic platform for real-time quantum reservoir computing. *Phys. Rev. Applied*, 20:014051, 2023. doi: 10.1103/PhysRevApplied.20.014051.

[134] S. Nandy, B. Mukherjee, A. Bhattacharyya, and A. Banerjee. Quantum state complexity meets many-body scars. *J. Phys. Condens. Matter*, 36(15):155601,

April 2024. ISSN 0953-8984, 1361-648X. doi: 10.1088/1361-648x/ad1a7b. URL https://iopscience.iop.org/article/10.1088/1361-648X/ad1a7b.

[135] F. B. Trigueros and C. Lin. Krylov complexity of many-body localization: Operator localization in Krylov basis. *SciPost Phys.*, 13(2), August 2022. ISSN 2542-4653. doi: 10.21468/scipostphys.13.2.037. URL http://arxiv.org/abs/2112.04722. arXiv:2112.04722 [cond-mat].

[136] H. Georg Menzler and R. Jha. Krylov Delocalization/Localization across Ergodicity Breaking. *Phys. Rev. B*, 110(12), September 2024. ISSN 2469-9950, 2469-9969. doi: 10.1103/physrevb.110.125137. URL http://arxiv.org/abs/2403.14384. arXiv:2403.14384 [quant-ph].

[137] M. Ganguli and A. Jana. State Dependent Spread Complexity Dynamics in Many-Body Localization Transition, September 2024. URL http://arxiv.org/abs/2409.02186. arXiv:2409.02186 [cond-mat].