Nonequilibrium dynamics of closed and open quantum systems

Kaelan Donatella

Université Paris Cité

École doctorale 564 : Physique en Île-de-France

Laboratoire Matériaux et Phénomènes Quantiques (MPQ, CNRS-7162)

Thèse de doctorat en physique

Nonequilibrium dynamics of closed and open quantum systems

présentée par

Kaelan Donatella

sous la direction de

Cristiano Ciuti

Jury

PR	Cristiano Ciuti	Université Paris Cité	Directeur de thèse
DR	Markus Holzmann	CNRS (LPMMC)	Rapporteur
CR	Pérola Milman	CNRS (MPQ)	Examinatrice
DR	Roberta Zambrini	IFISC	Rapporteuse

Présentée et soutenue publiquement à Paris le 12 décembre 2022

Résumé

Cette thèse est consacrée à l'étude de la dynamique hors-équilibre dans les systèmes quantiques fermés et ouverts. La question de l'obtention d'équations dynamiques numériquement efficaces pour simuler de tels systèmes, qui souffrent d'une *malédiction de dimensionnalité* lorsqu'ils sont de grande taille est au centre de la thèse, ainsi que les divers nouveaux résultats physiques qui peuvent être obtenus grâce à ces équations efficaces. La question de l'obtention de représentations efficaces des états quantiques est également centrale, à travers la construction d'*Ansatz*.

En particulier, nous étudions théoriquement la propagation de l'intrication dans les systèmes photoniques dissipatifs à forte interaction, motivés par la réalisation expérimentale récente de telles plateformes. Nos résultats montrent qu'en dépit du caractère dissipatif du système étudié, la propagation de l'intrication est ballistique avec une vitesse en accord avec l'image de quasiparticules impliquées dans la dynamique, des *doublons* et des *holons* photoniques. Notre analyse révèle que la dissipation est fortement asymétrique dans les deux cas étudiés de l'injection d'un doublon ou de son extraction, avec une influence plus forte dans le cas de l'extraction.

Ensuite, nous présentons la méthode du sous-espace dynamique, une nouvelle méthode numérique pour la simulation de l'évolution temporelle de systèmes quantiques ouverts avec une entropie modérée. En représentant la matrice densité par un sous-espace constitué de M états purs, la complexité de l'intégration de l'équation maîtresse de Lindblad est considérablement réduite, et l'algorithme obtenu est entièrement déterministe. Nous appliquons ensuite cette méthode pour modéliser des algorithmes quantiques bruités, en simulant un algorithme quantique bruité de transformation de Fourier qui est au centre de l'algorithme de Shor pour la factorisation de nombre premiers. Nous trouvons une loi d'échelle pour l'erreur en fonction du taux de dissipation, et le comportement de l'infidélité par rapport au circuit idéal est modélisé en fonction des états initiaux qui sont fournis en entrée à l'algorithme.

Enfin, l'application des méthodes d'apprentissage automatique à la mécanique quantique est présentée, à travers la classe d'Ansatz des états quantiques neuronaux (NQS). La construction générale des NQS est introduite, ainsi que sa relation avec d'autres Ansatz tels que les réseaux tensoriels. Ensuite, le problème de la simulation de la dynamique des systèmes fermés avec NQS est étudié, ainsi que les problèmes courants liés à la géométrie de l'espace des paramètres qui peuvent survenir. Un schéma alternatif est introduit et appliqué à des dynamiques de trempes dépendantes du temps dans des systèmes critiques, où le mécanisme de Kibble-Zurek peut être étudié. Enfin, différentes constructions d'Ansatz pour les systèmes ouverts sont présentées, notamment la combinaison de la méthode du sous-espace dynamique avec un Ansatz neuronal.

Keywords : physique quantique, systèmes quantiques ouverts, équation maîtresse de Lindblad, réseaux de tenseurs, circuits quantiques bruités, apprentissage automatique, états quantiques neuronaux, mécanisme de Kibble-Zurek, dynamique de l'intrication, photons fortement corrélés.

Summary

This thesis is devoted to the study of nonequilibrium dynamics in both closed and open quantum systems. The question of obtaining effective dynamical equations for such systems, that are plagued with a curse of dimensionality when scaled up is at the center of the thesis, as well as the various new physical results that can be obtained thanks to these effective equations. The question of obtaining efficient representations of quantum states is also central, through the construction of educated guesses or *ansätze*.

In particular, we theoretically study the propagation of entanglement in dissipative strongly-interacting photonic systems, motivated by the recent experimental realization of such platforms. Our findings show that in spite of particle losses the quantum entanglement propagation exhibits a ballistic character with propagation speeds related to the different quasiparticles that are involved in the dynamics, namely photonic doublons and holons respectively. Our analysis reveals that photon dissipation has a strikingly asymmetric behavior in the two configurations with a much more dramatic role on the holon entanglement propagation than for the doublon case.

Then, we present the dynamical corner-space method, a novel numerical method for the simulation of the dynamical evolution of open quantum systems with moderate entropy. By representing the density matrix by a sub-space of M corner states, the complexity of the integration of the Lindblad master equation is dramatically reduced, and the obtained algorithm is fully deterministic. We then apply this method to model noisy quantum algorithms, by simulating a noisy quantum Fourier transform algorithm that is at the center of the Shor algorithm for factoring. We derive a scaling law for the error as a function of the dissipation rate, and the behaviour of the infidelity with respect to the ideal circuit is modeled as a function of the initial states that are fed to the algorithm.
Finally, the application of machine learning methods to quantum mechanics is presented, through the neural quantum state (NQS) ansatz class. The general construction of NQS is introduced, as well as its relation to other ansätze such as tensor networks. Then, the problem of simulating dynamics of closed systems with NQS is investigated, as well as common problems related to the geometry of parameter space that may occur. An alternative scheme is introduced and applied to time-dependent quenches in critical systems, where the Kibble-Zurek mechanism can be investigated.

Keywords: quantum physics, open quantum systems, Lindblad master equation, tensor networks, quantum circuits, machine learning, neural quantum states, Kibble-Zurek mechanism, entanglement dynamics, strongly correlated photons.

Acknowledgements

Love and work are the cornerstones of our humanness.
— Sigmund Freud, *apparently misattributed*, ca. 1910

I would first like to thank members of the jury for taking the time to read this manuscript, correcting it and participating to my defense.

I would first like to thank Cristiano for his absolute trust during the whole thesis, from the moment we first met more than four years ago. Thank you for encouraging me to persist, try things out and tinker on, however poor my ideas may have seemed at first. Then, I would like to thank Alexandre, my unofficial co-directeur de thèse. Thank you for assisting me in all my projects and for providing rigorous guidance, which was very needed at many occasions.
A special mention goes to Zakari, who has been my intellectual brother during this thesis, and who, as Alexandre, helped me put rigor into my work, something I am not particularly endowed with unlike them. I think you can allow me some hippyish holism, only this time, in saying that sometimes the sum is more than the combined individual parts. You have brought so much to me, from French cinema to strange Spaniard customs, which makes me eternally grateful to you. Thanks again for all the moments shared.
Now to the theory team: Thank you to Zejian, for tolerating me in the office, and for having such a collective mindset. It has been a pleasure. Thank you to Valentin, for not crushing me at rugby, for all the shared pints and food, and for impressing me with your sheer indefatigability regarding mathematical details. Thank you to Geva for all the political discussions as well as explanations as to why the seniority of a researcher is anticorrelated with the truth value of their papers. Thank you to Phuong, who I can only admire for the fact that his computer invariably displays work-related windows. Thank you to Dalin for proving that you don't mess with Albanians, Federico for embodying the Torinese, and Luca for his (soon to be published?) complete theory of pasta. Thank you to Yuncheng for funny stories about China, recounted with the upmost sense of detail.
Thank you to Alberto, for assisting me at the beginning of this journey and for showing me that playing around and having fun pays dividends. Thank you to Filippo for all the nice moments around Europe.
A word for everyone I met at the lab, which made the experience even more interesting and/or enjoyable: Jérémie for being a fellow normal person in this strange world, Samuel for his cynical laugh and all the others who knew the Caminito and the Loma before they died. Now I will thank everyone outside the lab who, in their own way, contributed to this

work. Jurgen and Richard, thank you so much. Jurgen, for being “mon pote de droite" (sorry), and presenting things with such an interesting and original opinion in each of our discussions. Richard, for your wisdom, and to our deep mutual incomprehension when we were young and to our deep understanding and empathizing nowadays. Alexis, thanks for being one of my soulmates for all these years, I hope in the future we will continue to try teaching each other things. Julia, thanks for being my BFF (best female friend). Your point of view was needed in many occasions and you helped me grow a lot. Thanks to Gabi for the online art lessons during covid and for giving me back my morality book. Merci à Tim et Clémentine, qui m'ont beaucoup fait grandir. Merci aussi à la team INSA et leurs pièces rapportées pour tous ces rires et moments agréables. Ben, for accompanying me on so many adventures involving drum'n'bass and/or Britons. Elodie, merci de continuer à m'inviter à tes fêtes malgré l'expression fréquente de certains points de vue inaudibles. Kelly, thank you for being a mentor for so long. Ever since we met you gave me so much. Your voice remains in my head to this day. May the future be bright and travels be fertile. Romain, thank you for shaking me up in M1 and making me realise my immaturity, and for all the fun we had at parties. I would also like to thank everyone I met in M1 at Jussieu, Miguel, Martin, Mourad, John, Pascal, Jules and others. It is there that the physics journey really began and my memories from this time are still vivid.

Finally, I would like to thank my brothers, for being who they are each in their own way. Baveux, nerd and wannabe millionaire make quite a tribe. Thanking people for being themselves is a formula due to my mother, to whom I would like to express my love. Thank you for the continuous support and for being much more than “good enough". Merci papa, toi qui m'a aussi soutenu de manière inconditionnelle et sans qui le barbecue et le jazz manouche n'auraient peut-être jamais fait partie de ma vie. Je vous aime tous.

Publications

The present manuscript is structured mostly around the results reported in [1–3]. Results presented in [4] are not discussed here.

[1] K. Donatella, A. Biella, A. Le Boité and C. Ciuti,
"Entanglement dynamics in dissipative photonic Mott insulators",
Physical Review Research **2**, 043232 (2020).

[2] K. Donatella, Z. Denis, A. Le Boité and C. Ciuti,
"Continuous-time dynamics and error scaling of noisy highly-entangling quantum circuits",
Physical Review A **104**, 062407 (2021).

[3] K. Donatella, Z. Denis, A. Le Boité and C. Ciuti,
"Dynamics with autoregressive neural quantum states: application to critical quench dynamics",
arXiv:2209.03241 (2022).

[4] Z. Li, V. Heyraud, K. Donatella, Z. Denis and C. Ciuti,
"Machine learning via relativity-inspired quantum dynamics",
Physical Review A **106**, 032413 (2022).

Contents

General Introduction

Since the scientific revolution, physics has mainly consisted in understanding dynamical laws that constituents of reality obey. The origin of this paradigm can be traced back to Galileo [5], who wrote about describing *matter in motion* to explain physical phenomena. From there, more and more elaborate theories of matter in motion appeared, from Newton's second law to the Schrödinger wave equation. Physics mainly consisted in finding the most elaborate dynamical laws, that would explain everything: this would be known as a theory of everything.

However, this quest, impersonated so well by the minds of Einstein, Feynman, and "the particle physicist" has terribly failed. The search for a theory of everything, and the ultimate quest for reduction, can be considered to be lost. Let us say that we find a theory of everything, a complete description of a dynamical law that all particles and fields obey, which can be beautifully summed up inside a lengthy Lagrangian; what will we have learned? This is not to say that reduction brings *nothing*, but that reduction does not bring *everything*. There seems to be a fundamental truth to the nature of reality, that is that scales decouple; from some laws at the microscopic level, others can emerge at the macroscopic level. An instance of this general idea can be found in Conway's *Game of life* [6], in which a few deterministic rules suffice to create complex, self-replicating structures at much larger scales. This game is a perfect example of why reductionism cannot be enough if we wish to understand physical phenomena in the broadest sense. There are situations for which we know the rules, the laws of motion, but where we are fundamentally more interested in the *features* of the game that appear to be completely disconnected from these laws. For instance, we have known the laws of classical mechanics for at least two centuries, however many features of the theory are yet to be discovered and discussion about many fundamental topics hidden in these laws is ongoing.

From these considerations a new paradigm arose: the paradigm of complexity. Complexity arises when many particles interact, which leads to behaviour that cannot be computed, whether it may be with pen and paper or computers. Such considerations are beautifully explained in the famous article by Anderson, *More is different* [7]. From this, physics has largely become the science of effective models, where complete descriptions of physical systems are abandoned to transition to a more computable and scale-adapted effective behavior. Nowadays, it seems all fields of physics are filled with effective descriptions, from quantum chromodynamics [8] to fluid mechanics [9].

This thesis is devoted to the quest of finding better effective descriptions of closed and open quantum systems, and in particular of the complex dynamics that occur in many-body quantum systems. The interest of developing such methods is twofold: firstly, a new understanding of the problem at hand can be obtained by unveiling new aspects of

its mathematical structure, as has been the case with tensor network methods [10]. The development of such methods has led to the entanglement *area law* [11] for local unidimensional gapped systems, which has since then proved invaluable. Secondly, they enable researchers to study new problems, as they are more efficient to treat different regimes that were previously unattainable. A significant part of the thesis is dedicated to new physical results that have been obtained thanks to these effective descriptions.
Such models for quantum mechanical systems have for now mostly been applied to equilibrium and ground state properties. Density functional theory [12], a theoretically straightforward effective model useful for complex molecules, is perhaps the best example of the success of effective models for equilibrium properties, as it is nowadays ubiquitously used in research institutions and the private sector. Dynamical properties, as we will see in more detail in Chapter 1, are harder to pinpoint and involve a more complex mathematical structure than ground state properties in the general case. In particular, open quantum systems, a class of systems that describe quantum systems coupled to an external environment, hence being nonequilibrium almost by definition, are particularly elusive in that very few general results are known about them. Recently, general properties related to dissipative phase transitions [13] have been put forward, but the exposition of dynamical properties for open systems remain rare for now.

In Chapter 1, we present the theoretical framework surrounding the study of nonequilibrium quantum systems. We start by presenting the problems that arise for closed quantum systems, where nonequilibrium dynamics involve superpositions of eigenstates of a many-body Hamiltonian, thus resulting in a highly complex evolution. We present results related to correlation propagation in such systems [14], as well as the Kibble-Zurek mechanism, that has been experimentally verified in the last decade [15, 16]. We also present various numerical methods, from mean-field methods to tensor networks, that have yielded many important results in previous years. Then, we present the framework of open quantum systems. When a quantum system is coupled to an external environment, one can generally not compute the dynamics of the full composite system. Therefore, an efficient way to treat this problem is to treat the environment as a bath that is weakly coupled to the system, that is, remains weakly entangled with the system. By also supposing that the timescale related to the dynamics of the bath is much faster than that of the system, one can treat the system with a Lindblad master equation approach [17]. This equation closely resembles the Schrödinger equation, and has additional terms accounting for the coupling of the system with the environment. We present various numerical methods that were developed for such systems, as well as superconducting circuits [18], being a prominent platform for the study of such systems.
In Chapter 2, entanglement and its evolution in a strongly-correlated open quantum systems is studied. As mentioned, the study of correlation propagation as for now mostly been concentrated on closed systems [14, 19], and the impact of the interplay of strong interactions and dissipation on entanglement spreading has remained relatively untouched [20]. We investigate the nonequilibrium dynamics in a photonic Mott insulator induced by adding or removing a particle from the center of the chain. As such, we observe a ballistic entanglement propagation, with a speed that is predicted by an approximate solution for the closed system in the strongly interacting regime. We also observe a drastic

change in the behaviour of the system as a function of whether we start by removing or adding a particle, related to the form of the states involved in the dynamics. These results pave the way to more general results about entanglement dynamics in open systems.

In Chapter 3, the *dynamical corner-space* method is introduced. This method is based on the idea that there exists a class of systems whose von Neumann entropy is low enough that the density matrix describing the state of the open quantum system may be described as a collection of a limited number of pure states, thus rendering the integration of the Lindblad master equation more efficient. This idea is at the core of another method developed a few years ago, the *corner-space renormalization method*. The application of this method to the quantum Fourier transform will be presented. This is a key sub-routine in Shor's algorithm, and when subject to dissipation and/or decoherence, undergoes open system dynamics where a sequence of quantum gates are applied, entangling the qubits that form the register of the algorithm. Thanks to this method, new scaling behaviour for this algorithm as a function of dissipation and decoherence rates for this was possible to unveil. The impact of the initial state to the computation is also investigated, and a highly biased performance is found as a function of the loss channels.

Chapter 4 is devoted to using machine learning techniques to study nonequilibrium dynamics. Having shown immense success to represent high-dimensional data, such as images, text and audio, the latest advances in machine learning are currently strongly impacting the quantum physics community, where such advances have allowed to make progress on a large number of problems [21, 22], from ground state search, to quantum tomography or the simulation of quantum circuits. Such methods will be presented in detail, and we will then focus on their use for the simulation of nonequilibrium dynamics. This endeavor has been plagued by several problems, to which we propose a solution by reformulating the problem at hand onto an optimization problem to be solved at each time step of the dynamics. We will also see how machine learning methods can be used to study open quantum systems dynamics, and that representing a density matrix with neural networks presents several problems. As such, we envision an extension of the dynamical corner-space method to the use of neural network states, which we describe towards the end of the chapter.

1 Nonequilibrium quantum systems

I Introduction

As mentioned in the introduction, it quickly became clear that the Schrödinger equation could not be solved for more than a few particles in the presence of interactions [23]. This means that one cannot know the exact dynamics of most physical systems. However, this does not mean that many features of these dynamics cannot be accessed. Luckily for us, one can, through various approximation methods and educated guesses on the form of the wavefunction, referred to as *ansätze* in this manuscript, make great progress towards this goal.

In this section, we present the two main problems that are studied throughout the thesis. The first problem is understanding the nonequilibrium dynamics of ideal quantum systems, as described by the Schrödinger equation. Great progress has been made concerning quantum mechanics at equilibrium, however nonequilibrium properties are only starting to be properly investigated thanks to the rapid improvement of numerical methods and experimental platforms. Indeed, there is a much greater scarcity in the tools that can be used to study such systems. Many long standing problems have not been solved, such as the eigenstate thermalization hypothesis [24], and the validity of the Kibble-Zurek mechanism [25, 26], to name a few. In ideal quantum systems, nonequilibrium dynamics occur when the initial state of a dynamical process is not an eigenstate of the applied Hamiltonian [27]. This can be done by locally adding a particle to the system, or to apply a quench, that can be either time-independent or time-dependent. In the first part of this chapter, the physical and numerical ramifications of this problem will be presented, as well as the Kibble-Zurek mechanism and the experimental works that have been devoted to studying these issues in recent years. The main numerical methods used to study these issues such as mean-field and tensor networks will be presented in section II.3.

The second problem is understanding the dynamical properties of open quantum systems. Such systems are inherently out of equilibrium, as they can be viewed as ideal quantum systems coupled to an external environment. As explained in the introduction, studying open quantum systems is crucial as it enables one to understand quantum noise and its effect in quantum computational processes [28] as well as thermodynamical tasks [29]. In this context, dynamical properties can be studied by preparing the system in a certain state, and evolving the open quantum system in time. Most questions around this problem aim to understand how results for ideal quantum systems hold when an external environ-

ment is present. For example, the spreading of correlations in open quantum systems is an active field of research [20], that will be discussed in more detail in Chapter 2. On the quantum information side, understanding how noise processes affect quantum algorithms is of crucial importance if we wish to use them one day [30]. The theoretical framework for treating open quantum systems will be presented, as well as various experimental platforms which can be described by this framework. In this context, systems that interact with a physical environment are considered, nontrivial dynamical phenomena will arise, as dissipation will act as a new player.

The many-body problem

Let us now be more specific about the many-body problem, in both closed and open quantum systems. According to the axioms of quantum mechanics, the state of an ideal quantum system is fully specified by its wavefunction $|\psi\rangle$, that belongs to the Hilbert space associated to the system $\mathcal{H}$. $|\psi\rangle$ can be expanded in a given basis of this Hilbert space. A natural choice for this basis is a composite basis of local quantum numbers s_i (for spin systems, these local quantum numbers are the local spins), since most quantum systems we wish to describe are composite: they consist of L interacting degrees of freedom. In this part, and in most of the manuscript, we suppose that the equations to be solved to describe the time evolution of a physical system cannot be solved. Therefore, the only way to go forward is to perform a numerical integration of this equation, which involves storing in memory the state of the system at each step, and incrementally calculating the evolution of the system. The basis expansion of $|\psi\rangle$ reads:

$$|\psi\rangle = \sum_{s_1,s_2,\ldots,s_L} c_{s_1,s_2,\ldots,s_L} |s_1, s_2, \ldots, s_L\rangle \tag{1.1}$$

with $c_{s_1,s_2,\ldots,s_L}$ the amplitudes (complex numbers) corresponding to $|s_1, s_2, \ldots, s_L\rangle$, the basis elements of the Hilbert space (configurations) of the system. The Born rule tells us that by taking the modulus square of such amplitudes, we obtain the probabilities for finding the quantum state in such configuration after performing a measurement, i.e. $p(s_1, s_2, \ldots, s_L) = |c_{s_1,s_2,\ldots,s_L}|^2$. A combinatorial explosion can be seen directly here, as there are l^L amplitudes one must store (either on paper or on a computer) if one wishes to have a complete description of the system, with l the dimension of the local Hilbert space (which corresponds to the number of values the local quantum numbers s_i can take, assuming they each can take the same). More concretely, to describe a system of just 50 two-level systems, $2^{50} \sim 10^{15}$ complex numbers have to be described, or stored in memory. This is impossible even with modern day computers.
Regarding an open quantum system, the problem is similar but worse. As will be explained in this chapter, the state of an open quantum system can be fully described by a density matrix $\hat{\rho}$, this time belonging to the bounded operator space $\mathcal{B}(\mathcal{H})$ associated to $\mathcal{H}$. $\hat{\rho}$ can also be expanded over projectors, of which there will be l^L:

$$\hat{\rho} = \sum_{\substack{s_1,s_2,\ldots,s_L \\ s'_1,s'_2,\ldots,s'_L}} c^{s'_1,s'_2,\ldots,s'_L}_{s_1,s_2,\ldots,s_L} |s_1, s_2, \ldots, s_L\rangle \langle s'_1, s'_2, \ldots, s_L| . \tag{1.2}$$

It appears from this expansion that the complete description consists in l^{2L} complex numbers to be stored. Computationally, this means that a naive approach would enable

to describe open quantum systems of half the size of the largest closed systems we can simulate. Conveniently, many approaches exist that are tailored for open quantum systems to describe the properties of open quantum systems without storing the full density matrix.

II Dynamics of closed systems

II.1 Sudden quenches

As mentioned, the problem of quench dynamics involves solving the time-dependent Schrödinger equation, starting from a certain state $|\psi_0\rangle$. Throughout the manuscript, we fix $\hbar = 1$. The Schrödinger equation reads:

$$\frac{d\,|\psi(t)\rangle}{dt} = -i\hat{H}\,|\psi(t)\rangle\,. \tag{1.3}$$

If $|\psi(t=0)\rangle = |\psi_0\rangle$ is an eigenstate of $\hat{H}$, then the evolution as trivial: the phase of the state oscillates with frequency ω_0 corresponding to the energy of $|\psi_0\rangle$ and the solution is:

$$|\psi(t)\rangle = \exp(-i\omega_0 t)\,|\psi_0\rangle\,. \tag{1.4}$$

However, if the initial state is not an eigenstate of $\hat{H}$, the time evolution is nontrivial, and involves oscillations given by the frequencies of all the eigenstates of $\hat{H}$. To see this, one can expand $|\psi(t)\rangle$ as a superposition over eigenstates of $\hat{H}$, denoted $|\epsilon_i\rangle$:

$$|\psi(t)\rangle = \sum_i c_i(t)\,|\epsilon_i\rangle\,.$$

where the coefficients $c_i(t)$ depend on time. Hence the right hand-side of Eq. (1.3) reads:

$$-i\hat{H}\,|\psi(t)\rangle = -i\sum_i c_i(t)\epsilon_i\,|\epsilon_i\rangle = \sum_i \dot{c}_i(t)\,|\epsilon_i\rangle\,.$$

It is clear that if one wishes to integrate this equation and find the coefficients $c_i(t)$, one must know what the eigenvalues ϵ_i of the Hamiltonian are. Computationally speaking, this requires to diagonalize $\hat{H}$ and obtain its full spectrum, in contrast to many problems in quantum physics and chemistry that require the ground state of the system and the first excited states only. Thus in general, if one wants to obtain a high-precision integration of Eq. (1.3), one requires to be able to describe all the excited states of the system, which can be of varying complexity. For instance, for local gapped (i.e., outside of the critical region) one-dimensional systems, the complexity of the ground state is significantly weaker than that of excited states. This fact will be explained and described in more detail when discussing tensor network methods, whose limits reside in such considerations.
Note that a quench protocol instantly drives the state away from its previous equilibrium, thus cannot be treated perturbatively as is routinely done in the context of linear response theory, in which important results such as the fluctuation-dissipation have been found [31]. Since this protocol involves strong interactions, few general results have been obtained for it.

Correlation propagation

A set of results that has been obtained for sudden quench protocols in isolated quantum systems concerns correlation propagation. In a seminal work by Lieb and Robinson [32], it was shown that in generic quantum systems with finite-range interactions and a finite local Hilbert state there exists a speed limit on information propagation. This is known as the *Lieb-Robinson bound*, which reads[1]:

$$||[\hat{A}(t), \hat{B}]|| \leq c \exp\{-a[d(X,Y) - v|t|]\} \tag{1.5}$$

with $\hat{A}$ and $\hat{B}$ two observables with finite supports X and Y (for lattice systems, the sites associated to the Hilbert spaces on which $\hat{A}$ and $\hat{B}$ act), and a, c, v positive constants. As can be seen from the expression inside the exponential, v corresponds to the velocity of correlations[2]: the commutator between two operators with distinct support will decay exponentially the further one goes out of the light cone $d(X,Y) - v|t|$. From there an analogy with the light cone from special relativity was drawn, and such a limit was experimentally verified for various physical systems [14, 33].

An intuitive explanation of this phenomenon was first provided by Calabrese and Cardy [34]. When suddenly quenching some Hamiltonian parameters, excitations, also known as quasiparticles, will be created, since the equiblibrium quantum state now consists of a superposition of excitations in the new eigenbasis of the quenched Hamiltonian, as explained in section II.1. Only quasiparticles in close regions will be initially entangled. From there, quasiparticles spread baillistically in different directions with some group velocity, hence correlations between distant sites will appear as initially entangled quasiparticles reach them, as shown in Fig. 1.1. In this picture the Lieb-Robinson velocity is simply the maximal group velocity of quasiparticles. Such considerations will be explored further in Chapter 2, where the entanglement dynamics of dissipative quantum systems will be investigated, as well as the validity of the quasiparticle picture to explain such dynamics when coupling to an external environment is taken into account.

II.2 Time-dependent quenches and the Kibble-Zurek mechanism

Another equally interesting protocol consists in performing time-dependent quenches on critical systems that exhibit continuous phase transitions. In this context, a quantum state $|\psi_0\rangle$ is prepared as the ground state of some time-dependent Hamiltonian $\hat{H}[\lambda(t)]$ at time $t = 0$, which depends on some control parameter $\lambda(t)$. Depending on how the control parameter is tuned, various scenarios can take place. If the control parameter is tuned abruptly, we recover the case of sudden quenches. On the other hand, if the control parameter is tuned more slowly than the inverse of the gap of the system, the system remains in its ground state, which is known as adiabatic tuning. For physical systems that do not exhibit phase transitions, the system follows an adiabatic evolution and remains in the ground state of the Hamiltonian $\hat{H}[\lambda(t)]$ at all times (but may acquire a phase).

[1] Note that there exists a whole class of Lieb-Robinson bounds, this is just the most generic.

[2] Correlation functions appear in the expansion of the norm of the commutator $[\hat{A}(t), \hat{B}]$.

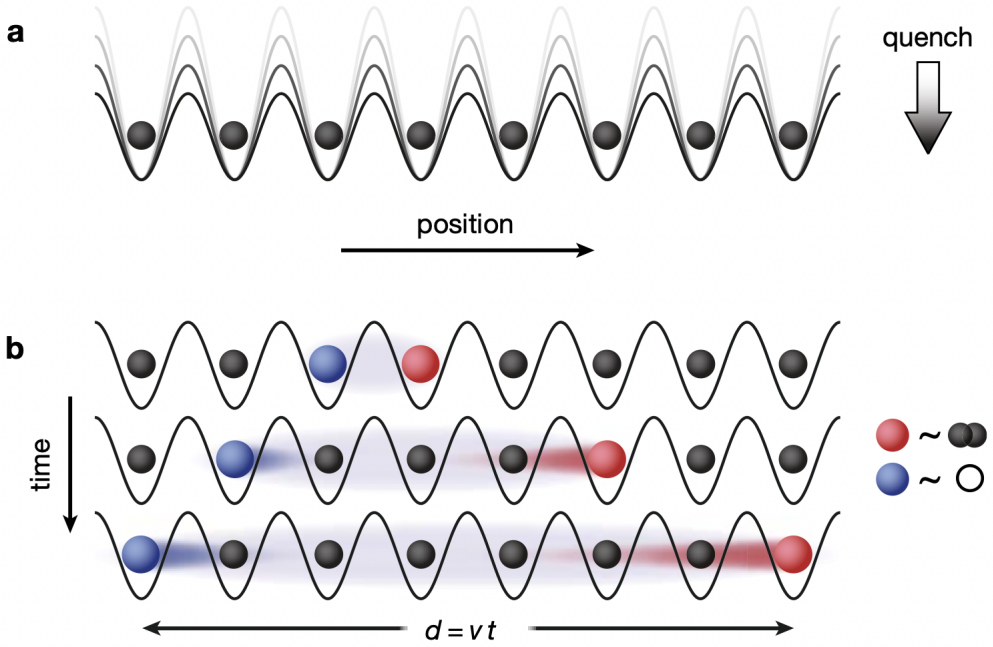

Figure 1.1: Sketch of the quasiparticle picture, in this case for bosons trapped in an optical lattice, described by the Bose-Hubbard model [14]. By quenching the lattice depth (that is related to the interaction strength of such models), quasiparticle excitations are created and propagate in the lattice, which leads to correlation propagation that obeys the Lieb-Robinson bound. From [19].

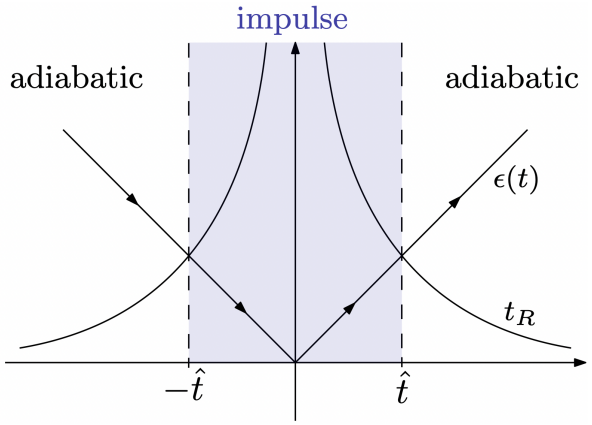

Figure 1.2: Scheme of the Kibble-Zurek mechanism according to the adiabatic-impulse approximation. At times $\pm\hat{t}$, the relaxation time of the system t_R crosses the time scale of the quench, which defines the limit between the adiabatic region, where the system essentially stays in the ground state of the system at that time, and the impulse regime, where nonequilibrium dynamics take place.

As an example, let us again consider the transverse-field Ising model, which will serve in the rest of the manuscript. The Hamiltonian is given by:

$$\hat{H}_{\mathrm{TFI}} = -J \sum_{\langle m,n \rangle} \hat{\sigma}_m^z \hat{\sigma}_n^z - g \sum_m \hat{\sigma}_m^x, \tag{1.6}$$

with a nearest-neighbor coupling strength J and a magnetic field strength g. It exhibits a second-order quantum phase transition that separates a ferromagnetic (for $J > 0$) or antiferromagnetic ($J < 0$) phase for $g < J$ from a paramagnetic phase for $g > J$. It faithfully describes many physical systems constituted of coupled two-level systems immersed in a magnetic field, such as neutral atom systems or magnetic nanomaterials. In one-dimension, the model is exactly solvable by using a Jordan-Wigner transformation. However, for higher dimensions, much less is known about the physics of the system. In 2D, the critical point is situated around $g/J \approx 3.044$ as given by quantum Monte-Carlo simulations [35], which are approximate methods.
The Kibble-Zurek mechanism [25, 26] predicts the formation of topological defects quenched systems undergoing a second-order phase transition as the parameters are linearly swept across a critical point. This universal behavior stems from the fact that there exists a time, denoted $\hat{t}$, at which the characteristic evolution time of the system (the relaxation time t_R, related to the inverse of the energy gap) becomes larger than the characteristic quench time τ_q, as schematically shown in Fig. 1.2. Before $\hat{t}$, the dynamics remains quasi-adiabatic, as the Hamiltonian parameters are tuned slowly with respect to the time scale of the system. After this time, in contrast, genuinely non-equilibrium dynamics takes place as critical slowing down sets in and the system dynamics becomes slower than the Hamiltonian parameters sweep. Ignoring specific details of this dynamics, one can derive a scaling law for the density of created defects [3] at the end of the linear quench [26]. This

[3] with respect to the ground state at the end of the quench. In a 1D classical spin system, this will be proportional to the number of domains with spins flipped up if all spins point down in the ground state.

scaling law is derived quite simply. Consider that for second-order phase transitions, the correlation length ξ and relaxation time τ scale as:

$$\xi \sim g^{-\nu}, \tau \sim g^{-z\nu} \tag{1.7}$$

where we have taken the magnetic field g to be an external control parameter. The freeze-out time $\hat{t}$ is a solution of the equation

$$\tau(\hat{t}) = g(\hat{t})^{-z\nu} \tag{1.8}$$

since we suppose these timescales cross at time $\hat{t}$. With $g(t) = g_0(1 - t/\tau_q)$, corresponding to a linear quench:

$$\hat{t} = (1/\tau_q)^{-z\nu/(z\nu+1)} \tag{1.9}$$

$$\hat{\xi} = (1/\tau_q)^{-\nu/(z\nu+1)} \tag{1.10}$$

with $\hat{\xi}$ the correlation length of the domains induced by the nonequilibrium dynamics. The final defect density is inversely related to the this average correlation length:

$$\langle \hat{n}_d(\tau_q) \rangle \sim \hat{\xi}^{-d}, \tag{1.11}$$

and the final scaling law therefore reads

$$\langle \hat{n}_d(\tau_q) \rangle \sim \tau_q^{-d\nu/(z\nu+1)}. \tag{1.12}$$

The scaling law for the density of defects has a physical explanation in terms of the change in the different timescales during the quench. This is shown in Fig. 1.2. Far away from the critical point, the gap of the system is large, therefore the characteristic timescale of the system (related to the inverse of the energy gap) remains small; almost any parameter tuning will be adiabatic (corresponding to the far left and right of the figure). As one approaches the critical point, precisely at the freeze-out time $-\hat{t}$, $t_R = \epsilon(t)$: the characteristic time scales of the quench and of the system cross. From there, the evolution stops being adiabatic, entering the impulse regime. Correlations can therefore develop within domains of a limited size, related to the freeze-out time, which is time the system stays in this regime. Then, at $+\hat{t}$, the evolution resumes into the adiabatic regime, and domains remain frozen.

II.3 Numerical methods

As we have seen in the previous sections, numerical simulations can be an important tool to study the dynamics of nonequilibrium quantum systems, both in sudden and time-dependent quench protocols. In this section we will present a few of the important methods used to study such protocols.

Mean-field approaches

A class of methods that enables one to treat large quantum systems in arbitrary dimensions is mean-field methods, also known as Gutzwiller ansätze [36]. These methods have previously mostly been used for electronic systems [37–39] and variants of it have been used for the study of bosonic gases [40]. The key idea is to write the wavefunction as a product of local wavefunctions:

$$|\psi\rangle = \bigotimes_i |\phi_i\rangle \,. \tag{1.13}$$

As such, this enables one to separate the Schrödinger equation into L local equations. One can also view this as a drastic reduction of the number of free parameters required to describe the many-body wavefunction. Indeed, the single-site wavefunctions are described by l parameters, with l the local Hilbert space dimension. They are subject to the orthogonality condition $\langle\phi_i|\phi_j\rangle = \delta_{ij}$, with δ_{ij} denoting the Kronecker delta. In the case of spin systems, this results in $|\phi_i\rangle$ having only two coefficients corresponding to the probability amplitudes of the spin being up or down. This results in lL complex parameters for an $L-$body system of local Hilbert space dimension l, which is a drastic improvement upon the l^L coefficients needed for the exact wavefunction. The coefficients of the many-body wavefunction are given by:

$$c_{s_1,s_2,\ldots,s_L} = \Pi_{s_i=s_1}^{s_L} a_{s_i} \tag{1.14}$$

with a_{s_i} being vectors of dimension l. However, this approach will only work for weakly correlated quantum systems, since non-separable states cannot be described by such approaches.

An extension of this idea exists as *cluster mean-field* approaches [41]. The key idea is to consider more than one site per factorized wave function. The state is written as:

$$|\psi\rangle = \bigotimes_i^{N_c} |\phi_i\rangle \,. \tag{1.15}$$

with N_c the number of clusters. The size of a cluster is given by L/N_c. This ansatz is able to capture states with an entanglement between sites belonging to the same cluster, thus having strictly more capacity to represent entanglement than the ansatz in Eq. (1.13).

Tensor networks and matrix product states

Beyond mean-field ansätze, powerful methods that have proved very effective in the study of one-dimensional systems are based on tensor networks [10]. The key idea behind such methods is to approximate the coefficients of the many-body wave function defined in Eq. (1.1) as a product of tensors:

$$c_{s_1,s_2,\ldots,s_L} = \sum_{i_1,i_2,\ldots,i_L} A_{s_1}^{i_1,i_2} A_{s_2}^{i_2,i_3} \ldots A_{s_L}^{i_L,i_1}. \tag{1.16}$$

This is known as the *matrix product state* (MPS) ansatz. The tensors $A_{\sigma_i}^{i,j}$ are of rank 3, with one dimension corresponding to the physical degrees of freedom σ_i, and two

dimensions given by M^2 coefficients that can introduce correlations between sites. The number of parameters contained in an MPS ansatz is therefore lM^2, which is no longer exponential in the system size.
To see how the MPS ansatz is constructed, let us define the Schmidt decomposition of a quantum state $|\psi\rangle$ belonging to a composite Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ with $\dim(\mathcal{H}_A) = m, \dim(\mathcal{H}_B) = n, m \leq n$. There exists vectors $\{|u_1\rangle, \dots |u_m\rangle\} \in \mathcal{H}_A$ and $\{|v_1\rangle, \dots |v_n\rangle\} \in \mathcal{H}_B$ which are pairwise orthogonal $\langle u_i|u_j\rangle = \delta_{ij}$ and $\langle v_i|v_j\rangle = \delta_{ij}$ and a real numbers $\Lambda_1 \geq \Lambda_2 \geq \dots \Lambda_m \geq 0$ such that $|\psi\rangle$ can be written as a Schmidt decomposition:

$$|\psi\rangle = \sum_{i=1}^{m} \Lambda_i |u_i\rangle \otimes |v_i\rangle \tag{1.17}$$

with Λ_i the Schmidt coefficients, and m the Schmidt rank of the state. Note that for a separable state we have $m = 1$. Defining the reduced density matrix $\hat{\rho}_A = \mathrm{tr}_B|\psi\rangle\langle\psi|$, it can be shown that the Schmidt coefficients are related to the entanglement entropy

$$S_A = -\mathrm{tr}\{\hat{\rho}_A \log \hat{\rho}_A\} = -\sum_{i=1}^{m} |\Lambda_i|^2 \log |\Lambda_i|^2. \tag{1.18}$$

For a one-dimensional system of L sites with a local Hilbert space dimension l, the coefficients $c_{s_1,s_2,\dots,s_L}$ can be expressed as a function of the Schmidt coefficients, by successive Schmidt decompositions:

$$c_{s_1,s_2,\dots,s_L} = \sum_{\alpha_1=1}^{l} \sum_{\alpha_2=1}^{l^2} \cdots \sum_{\alpha_n=1}^{\min(l^n, l^{L-n})} \cdots \sum_{\alpha_L=1}^{l} = \Gamma^{[1]s_1}_{1,\alpha_1} \Lambda^{[1]}_{\alpha_1} \Gamma^{[2]s_2}_{\alpha_1,\alpha_2} \Lambda^{[2]}_{\alpha_2} \cdots \Lambda^{[L-1]}_{\alpha_{L-1}} \Gamma^{[L]s_L}_{\alpha_L,1}. \tag{1.19}$$

The $\Gamma^{[i]}$ matrices correspond to a basis transformation between the full basis and the Schmidt basis, and the $\Lambda^{[i]}$ vectors correspond to the Schmidt coefficients after tracing out the $(L-i)$th site. This encoding seems particularly inefficient, as the $\Gamma^{[L/2]}$ contains by itself l^L elements. However, the beauty and central point of MPS ansätze is that the entanglement entropy of local, gapped many-body systems is limited. This is known as the entanglement *area law* [42]. Formally, for such systems,

$$S_{A(L)} \sim \partial L \underset{1D}{\sim} \text{constant} \tag{1.20}$$

with $A(L)$ a subsystem of size L, and ∂L the size of the boundary region between L and the rest of the system. This means that the number m of Schmidt coefficients Λ_i is limited, and they in fact decay exponentially [10]. This enables one to safely truncate the size of the matrices in the previous equation to a certain bond dimension D. This leads to the construction presented in Eq.(1.16), with a number of parameters scaling as lLD^2. These methods are powerful, as they enable to describe states with correlations beyond that of the mean-field ansatz while keeping the number of parameters polynomial in the system size. However, for more complex geometries, and in particular in 2D, constructing such an ansatz presents various problems, and in addition, even if the entanglement of a 2D system follows an area law, the area ∂L in 2D is not constant anymore, which means the bond dimension D will scale less favorably than in 1D.

II.4 Experimental systems

Sudden quenches

One of the main reasons to study the nonequilibrium dynamics of quantum systems is the recent progress of experimental platforms that enable researchers to perform controlled experiments, thus being able to directly test predictions. These platforms are known as quantum simulators [43], and consist in implementing a Hamiltonian with a physical analogue to that Hamiltonian. These allow one to tailor a specific Hamiltonian and the degree of control is such that quenches can be performed, particles can be added, and many different operations can be performed to study effects presented in the previous section. This is reminiscent of Feynman's vision of simulating quantum mechanics by using an analogous quantum mechanical system [44].
As an example, let us consider the quantum simulation of a spin Hamiltonian, as presented in ref. [15]. An important platform that has been perfected in the last decades has been optical lattices, where atoms can be trapped by optical tweezers and one can use the strong nonlinearity of the energy spectrum of atoms to effectively realize two-level systems. The atoms are often put in highly excited states, such as Rydberg states [45] that correspond to a high principal quantum number n. This enables a larger degree of control and coupling to electromagnetic fields. Such a platform can be faithfully described by the following Hamiltonian, where the two eigenstates of the $\hat{\sigma}^z$ operator correspond to the ground $|g\rangle$ and excited states $|e\rangle$ of the Rydberg atoms:

$$\hat{H} = -\sum_i \Omega\hat{\sigma}_i^x - \sum_i (\mathrm{I}_i - \Delta)\hat{\sigma}_i^z - \sum_{i\neq j} \frac{V_{ij}}{2}\hat{\sigma}_i^z\hat{\sigma}_j^z \tag{1.21}$$

where $V_{ij} = C_6/|\mathbf{r}_i - \mathbf{r}_j|^6$ is an atomic potential (C_6 is a constant given by experimental details), Ω is the Rabi frequency of an applied transverse-field that drives transitions between $|e\rangle$ and $|g\rangle$, and $(\mathrm{I}_i - \Delta)$ determines the frequency of the longitudinal field. Δ is the laser detuning, and $\mathrm{I}_i = \sum_{j,i\neq j} V_{ij}/2$ represents a site-independent detuning. Here V_{ij} is taken to be positive, corresponding to attractive interactions, and in the absence of transverse and longitudinal fields the ground state of this Hamiltonian is an antiferromagnet[4]. Note that by taking a homogeneous spacing between atoms, one can take V_{ij} to be constant, and simply consider $V_{ij} \approx J$. Therefore, one obtains a slightly modified antiferromagnetic Ising Hamiltonian, and quench dynamics can be induced by suddenly tuning the transverse and longitudinal field.
In Fig. 1.4, results of the correlation functions $C(i,r) = 4\left(\langle\hat{\sigma}_i^z\hat{\sigma}_{i+r}^z\rangle - \langle\hat{\sigma}_i^z\rangle\langle\hat{\sigma}_{i+r}^z\rangle\right)$ are shown for a quench from a paramagnetic phase, with $\Delta \gg J \gg \Omega$, to an antiferromagnetic phase with $J \gg \Delta, \Omega$. In the figure, $C(i,i)$ is related to the local magnetization as $C(i,i) = 1 - 4\langle\hat{\sigma}_i^z\rangle^2$. One can clearly see an inversion of the sign of $C(0,1)$ and $C(1,0)$ as a function of the detuning, transitioning from ferromagnetic ($\Delta < 0$) to antiferromagnetic ($\Delta > 0$) correlations. Note that for the upper panels, corresponding to sudden quenches, the correlations will remain short-ranged, thus the results could be compared to data from exact diagionalization for a 4×4 lattice, which are a good fit to the experiment. Another

[4]We have added a minus sign to the Hamiltonian in ref. [15] for the sake of simplicity.

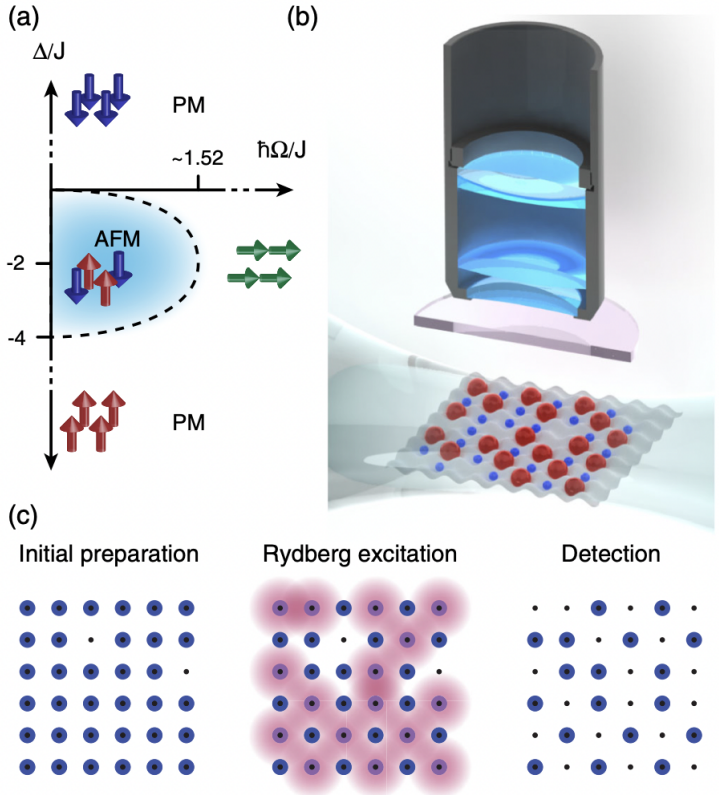

Figure 1.3: (a): Phase diagram of the model described by the Hamiltonian in Eq.(1.21). One is mostly interested in the antiferromagnetic region (AFM), whose eigenstates can be probed by suddenly quenching the detuning from deep in the paramagnetic phase, with $\Delta \ll J$. (b) Scheme of the experimental setup: a 2D array of atoms trapped in an optical lattice whose states $|e\rangle$ (large red spheres) $|g\rangle$ (small blue spheres) can be coupled with an infrared laser. The top part of the scheme represents the high-resolution objective that can resolve individual sites, enabling precise analyses of experimental data. (c): Typical protocol. The system is prepared in a trivial phase, with a small number of defects (left), and is then quenched into an excited configuration (center). Then, increasing the lattice depth causes the atoms in the excited state to be lost (right), hence only the atoms in the ground states remain to simplify detection. Reproduced from [15].

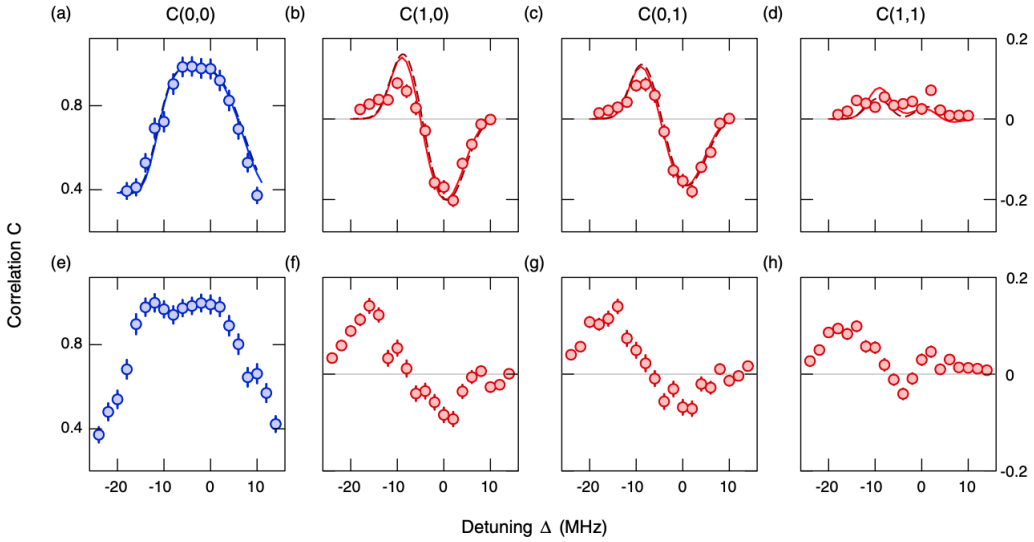

Figure 1.4: Correlation functions after a sudden quench in the two-dimensional antiferromagnetic Ising model as a function of the detuning Δ, for different values of the quench time T. Dashed lines correspond to linked-cluster calculations, while solid lines correspond to exact diagonalization data. Upper panels (a), (b), (c) and (d) correspond to time $\Omega T = 0.5\pi$, which is fast enough to be considered a sudden quench, while lower panels (e, (f), (g) and (h) correspond to $\Omega T = 2.97\pi$, where long-range correlations develop. Reproduced from [15].

group has performed a similar study the same year [46], and more results on more complex models, for which the phase diagram is unknown, are expected in the coming years.

Time-dependent quenches

Time-dependent quenches can also be implemented in quantum simulators, and in recent years many works have investigated this, using Rydberg atoms [16], in a setup close to what was presented previously, trapped ions [47], Bose-Einstein condensates [48], and even digital quantum simulators such as IBM's machines [49]. To implement a time-dependent quench on Rydberg atom platforms, the transverse and longitudinal fields have to be tuned linearly in time.

Here we will focus on the implementation of a linear quench to examine the validity of the Kibble-Zurek mechanism on Rybderg simulator, close to what was presented previously. In Fig. 1.5, experimental results are shown for a linear quench performed on a system that undergoes an Ising-like quantum phase transition, breaking $\mathbb{Z}_2$ symmetry. In the two-level basis spanned by the ground state $|g\rangle$ and the excited state $|e\rangle$, the Hamiltonian describing the system is given by:

$$\hat{H} = \frac{\Omega}{2}\sum_i \hat{\sigma}_i^x - \Delta\sum_i \hat{P}_i^{|g\rangle} + V_{ij}\sum_{\langle i,j\rangle} \hat{P}_i^{|g\rangle}\hat{P}_j^{|g\rangle} \tag{1.22}$$

where Ω is the Rabi coupling frequency between states $|g\rangle$ and $|e\rangle$, Δ is the laser detuning, V_{ij} is the interaction strength and $\hat{P}_i^{|g\rangle} = \frac{1}{2}(\hat{\mathbb{1}}_i + \hat{\sigma}_i^z)$ is the projector onto state $|g\rangle$ on site i.

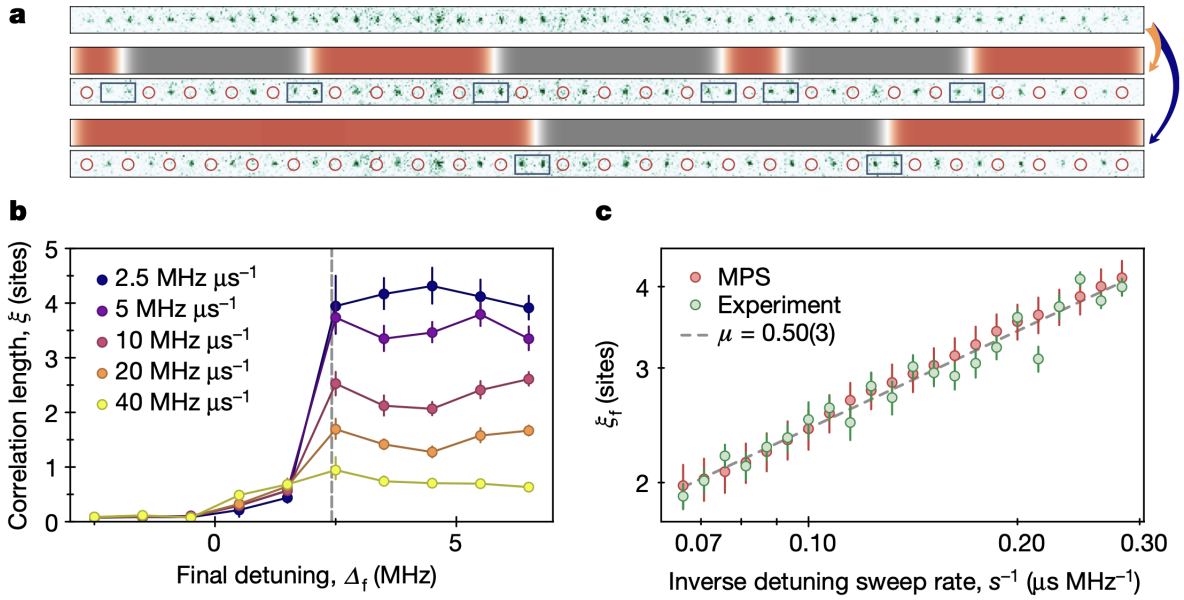

Figure 1.5: Experimental results for a time-dependent linear quench on a Rydberg quantum simulator. (a): Single-shot images of the system before the quench (top row), after a fast quench (medium row) and after a slow quench (bottom row). The correlated domains are larger in the bottom row, as more time is spent in the impulse region hence domains have more time to develop. A higher density of defects is observed for fast quenches. (b) Correlation length as a function of the final detuning. The grey dashed line indicates the critical detuning: if the quench is stopped before this, the correlation length remains small, as the dynamics remain quasi-adiabatic. (c) Scaling of the correlation length as a function of the sweep rate, which is the inverse of the quench time τ_q, compared with results from an MPS numerical integration. Reproduced from [16].

The system is initialized in a homogeneous state with all atoms in the $|g\rangle$ state, and after the quench, some domains with atoms all in state $|e\rangle$ appear, whose density is predicted by the Kibble-Zurek mechanism. As such, the critical detuning can be estimated, and the Kibble-Zurek mechanism can be verified. In the same work [16], authors investigate other phases, as their machine is programmable, and other critical phases of the system can be investigated. Once again, it is really the degree of control of such platforms that make them ideal to study quench dynamics.

III Open quantum systems

III.1 General setting

Many other interesting cases of nonequilibrium quantum phenomena can be found within the framework of open quantum systems. To describe this class of systems, let us consider a quantum system S, described by a density matrix $\hat{\rho}$. After performing a few experiments, one can quickly notice that S interacts with its environment E: for example, if S describes a photonic system, photons can leak out in the environment. A straightforward way to understand how the leakage of photons takes place would be to describe the composite system $S + E$, made of the system S and the environment E. The composite system is described by a Hamiltonian $\hat{H}_{S+E}$ that describes the degrees of freedom of both the system and the environment. It can be written in a general form

$$\hat{H}_{S+E} = \hat{H}\left(\{\hat{S}_k\}\right) + \hat{H}_E\left(\{\hat{\Gamma}_k\}\right), \tag{1.23}$$

with $\hat{H}$ the system Hamiltonian, $\hat{H}_E$ the environment Hamiltonian, $\{\hat{S}_k\}$ the set of system operators and $\{\hat{\Gamma}_k\}$ the set of environment (or bath) operators. It is clear that the many-body problem appears once again, only worse, as realistic environments can be comprised of many bosonic and fermionic degrees of freedom that describe the electromagnetic field and the molecules composing the air and the experimental apparatus. Neglecting these last contributions and supposing that the system interacts with only a few bosonic modes is still not enough for us to compute the dynamics of most composite systems, its Hilbert space still being too large.
A way around this is to search for an effective description of $\hat{\rho}_S$ and obtain a time-evolution equation for it. Note that if the reduced system interacts with its environment, the wave function of the composite system will become entangled (non-separable):

$$|\Psi\rangle_{SE} \neq |\psi\rangle_S \otimes |\chi\rangle_E . \tag{1.24}$$

This means that if one wants to obtain a description of S, by performing a partial trace on $|\Psi\rangle_{SE}$ one will obtain a density matrix $\hat{\rho}_S$, and cannot use the Schrödinger equation anymore. Hence one needs to change the formalism, and see how a time-evolution equation can be obtained for the density matrix when taking into account the interaction between system and environment. Let us first properly define the objects we will need to use to obtain an effective dynamical equation on the system.

III.2 Density matrices

A density matrix $\hat{\rho}$ is an operator in the bounded operator space $\mathcal{B}(\mathcal{H})$ associated to a Hilbert space $\mathcal{H}$ and should satisfy the following properties to describe a physical system [17]:

1. Trace one: $\mathrm{Tr}\hat{\rho} = 1$,
2. Positive-semidefinite: $p_i \geq 0$, with p_i its eigenvalues,
3. Hermiticity: $\hat{\rho}^\dagger = \hat{\rho}$.

It also admits infinitely many decompositions of the form

$$\hat{\rho} = \sum_i p_i \left|\phi_i\right\rangle \left\langle\phi_i\right|, \tag{1.25}$$

with p_i its eigenvalues and $|\phi_i\rangle$ the corresponding eigenstates. The physical interpretation for such a decomposition is that the density matrix represents a collection of states that each have a probability p_i. According to this view, the density matrix tells us which states will be measured with which frequency if we perform a large number of experiments. One can also stay agnostic about the physical interpretation of this picture and consider that a density matrix describes a reduced quantum state, with uncertainty coming from its entropy that stems from entanglement with the environment.

III.3 Quantum operations

Let us now see how the a density matrix can change, meaning how it evolves under quantum operations. A quantum operation is a completely-positive, trace-preserving (CPTP) map, that one applies to a density matrix. The Kraus theorem [50] states that this map can be decomposed as the application of N^2 so-called Kraus operators:

$$\hat{\chi} = \mathcal{K}[\hat{\rho}] = \sum_i \hat{K}_i \hat{\rho} \hat{K}_i^\dagger \tag{1.26}$$

with $\hat{\chi}$ the transformed density matrix. A quantum operation corresponds physically to any transformation that has happened to the physical system, such as time evolution, a measurement or a certain experimental protocol. The Kraus operators must satisfy the condition $\sum_i \hat{K}_i^\dagger \hat{K}_i = \mathbb{1}$ for Eq. (1.26) to be valid. $\mathcal{K}$ must also be linear

$$\mathcal{K}[\alpha\hat{\rho}_1 + \beta\hat{\rho}_2] = \alpha\mathcal{K}[\hat{\rho}_1] + \beta\mathcal{K}[\hat{\rho}_2] \tag{1.27}$$

and conserve the Hermiticity of $\hat{\rho}$. All these conditions ensure that the transformed quantum state $\hat{\chi}$ conserves the properties of $\hat{\rho}$. $\mathcal{K}$ is a superoperator, meaning that it can be compactly represented by a $N^2 \times N^2$ matrix acting on operators, much like operators act on vectors. One can already see the inherent complexity to numericaly study open quantum systems, as objects of size $N^2 \times N^2$ must be manipulated.
As an example of a quantum operation, let us consider a single qubit, whose levels are written $|\uparrow\rangle$ and $|\downarrow\rangle$. We assume that the state $|\uparrow\rangle$ has been prepared, and encodes some

kind of information. The simplest model to describe energy relaxation is through the *amplitude damping* channel $\mathcal{N}_p$, that is a quantum operation with Kraus operators given by:

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \; K_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}. \tag{1.28}$$

The strength of the channel is given by p. Applying this channel on the state $|\uparrow\rangle$, we have:

$$\mathcal{N}_p[|\uparrow\rangle\langle\uparrow|] = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix} = p\,|\downarrow\rangle\langle\downarrow| + (1-p)\,|\uparrow\rangle\langle\uparrow| \tag{1.29}$$

which is a statistical mixture of states $|\downarrow\rangle$ and $|\uparrow\rangle$ with probabilities p and $1-p$, respectively. Through this simple example, we also see how entropy has been created through the process, since we go from a pure state to a mixed state. We have kept things very simple here, with only a single qubit, therefore no coupling between qubits, and have supposed no Hamiltonian inducing some time-evolution. In the general case, these ingredients come into play and the dynamics is much more complex.

III.4 The Lindblad master equation

Having described what a general quantum operation looks like, recall that we are interested in obtaining an equation for the time evolution of the quantum system represented by $\hat{\rho}$. Specifically, we are interested in finding a general form for the map $\mathcal{M}$ defined by

$$\hat{\rho}(t+\delta t) = \mathcal{M}\hat{\rho}(t) \tag{1.30}$$

for vanishing δt. By expanding up to second-order, one obtains:

$$\hat{\rho}(t+\delta t) = \sum_i \hat{M}_i(t)\hat{\rho}(t)\hat{M}_i(t) = \hat{\rho}(t) + \delta t \partial_t \hat{\rho}(t) + O(\delta t^2) \tag{1.31}$$

By assuming that the environment of the system is much larger than the system itself, one can consider time-independent Kraus operators $\hat{M}_i$. Indeed, in this case, the environment itself acts as a bath and remains at equilibrium at all times, hence the action of the environment on the system is the same at all times. This is known as the Born approximation.
Now, some care must be taken with respect to the time increment δt, by considering the physical timescales at hand. δt must be small with respect to the system timescale Δt_S (that can be taken to be the smallest characteristic evolution time) for the limit to make sense. However, δt must be much larger than Δt_E, the environment timescale, so that no coherent effects remain in the system-environment interaction. This happens when the spectrum of the environment is flat, and the environment operators are then delta-correlated:

$$\langle \hat{\Gamma}_k(t)\hat{\Gamma}_k(t')\rangle = \delta(t-t'). \tag{1.32}$$

This is known as the Markov approximation. We can now take the appropriate limit $\delta t \to 0$, and obtain a general form for the Kraus operators $\hat{M}_i$. We take one of them to

be of the order of unity, and others of the order of $\sqrt{\delta t}$ in order to obtain a second-order update on the density matrix:

$$\begin{cases} \hat{M}_0 & = \mathbb{1} - i\delta t \hat{K} \\ \hat{M}_{i>0} & = \sqrt{\delta t}\hat{J}_i. \end{cases} \tag{1.33}$$

One can split the operator $\hat{K}$ into a Hermitian part and an anti-Hermitian part, given respectively by $\hat{I} = (\hat{K} + \hat{K}^\dagger)/2$ and $\hat{G} = i(\hat{K} - \hat{K})/2$, with $\hat{K} = \hat{I} - i\hat{G}$. One has

$$\hat{M}_0^\dagger \hat{\rho}(t) \hat{M}_0 = \hat{\rho}(t) - i\delta t[\hat{I}, \hat{\rho}(t)] - \delta t\{\hat{G}, \hat{\rho}(t)\} + O(\delta t^2). \tag{1.34}$$

Using the normalization condition on the Kraus operators, we have:

$$\mathbb{1} = \sum_i \hat{M}_i^\dagger \hat{M}_i = \hat{M}_0^\dagger \hat{M}_0 + \delta t \sum_{i>0} \hat{J}_i^\dagger \hat{J}_i + O(\delta t^2) = \mathbb{1} - 2\delta t \hat{G} + \delta t \sum_{i>0} \hat{J}_i^\dagger \hat{J}_i + O(\delta t^2). \tag{1.35}$$

By rearranging terms, we obtain $\hat{G} = \sum_{i>0} \hat{J}_i^\dagger \hat{J}_i/2$. The dynamics of the density matrix $\hat{\rho}(t)$ is therefore given by an equation of the form

$$\hat{\rho}(t+\delta t) = \hat{\rho}(t) - i\delta t[\hat{I}, \hat{\rho}(t)] + \delta t \sum_{i>0} \left(\hat{J}_i \hat{\rho}(t) \hat{J}_i^\dagger - \frac{1}{2}\{\hat{J}_i^\dagger \hat{J}_i, \hat{\rho}(t)\} \right) + O(\delta t^2). \tag{1.36}$$

Recall that the evolution of a closed quantum system is given by the von Neumann equation:

$$\hat{\rho}(t+\delta t) = \hat{\rho}(t) - i\delta t \left[\hat{H}, \hat{\rho}\right] + O(\delta t^2) \tag{1.37}$$

we identify the Hamiltonian in Eq. (1.36), and we finally obtain the Lindblad master equation:

$$\partial_t \hat{\rho} = -i \left[\hat{H}, \hat{\rho}\right] + \frac{1}{2} \sum_{i>0} \left(2\hat{J}_i \hat{\rho} \hat{J}_i^\dagger - \left\{ \hat{J}_i^\dagger \hat{J}_i, \hat{\rho} \right\} \right) \tag{1.38}$$

where we have dropped the time dependence for readability. By comparing the von Neumann equation and the Lindblad master equation, one can see that the operators $\hat{J}_i$ must vanish when the evolution is that of a closed system: therefore we see that they must describe an effect of the environment. The Lindblad master equation can be derived in many other ways more formal ways [17, 51]. Let us see how the operators $\hat{J}_i$ can be physically interpreted as jump operators, that induce a collapse of the system wavefunction due to a continuous weak measurement process.

III.5 The stochastic Schrodinger equation and quantum trajectories

Another way to derive the Lindblad master equation is by considering a physical system under continuous measurement. By doing this, we obtain what is known as an unraveling of the master equation in terms of a stochastic process. The picture here is that of a physical system that stochastically undergoes jumps from one state to another, with a collapse of the wavefunction occuring at random times. For example, consider an atomic system at room temperature. At each instant, there is a nonzero probability

that excitations are lost to the thermal environment. This can be seen as a stochastic process with two outcomes: either the excitation is lost or it is not.

In the general case, this picture is well described by the generalized measurement picture: the quantum state after a generalized measurement is given by:

$$\hat{\rho}(t) \;\rightarrow\; \hat{\rho}(t+\delta t) = \sum_{\mu}^{n_\mu} \frac{\hat{M}_\mu \hat{\rho}(t) \hat{M}_\mu^\dagger}{p_\mu(t)}, \tag{1.39}$$

with $\hat{M}_\mu$ the measurement operators. That is, by performing a measurement of the system described by $\hat{\rho}(t)$, the state has collapsed to a certain state, but we are uncertain about which state it is, of which there are n_μ possible outcomes. $p_\mu(t) = \langle \hat{M}_\mu^\dagger \hat{M}_\mu \rangle(t)$ is therefore the probability of measuring the μth outcome if a projective measurement is carried out. Let us consider the following choice of measurement operators

$$\begin{cases} \hat{M}_0 = \hat{\mathbb{1}} - (i\hat{H} + \hat{J}^\dagger \hat{J}/2)\delta t, \\ \hat{M}_1 = \hat{J}\sqrt{\delta t}, \end{cases} \tag{1.40}$$

where we have supposed that there is only one jump operator $\hat{J}$. This closely corresponds to what was given in Eq. (1.33): the density matrix gets updated in time as under the action of the Lindblad master equation. The dynamics of an open system can thus be interpreted as the result of a continuous weak measurement of the jump operator $\hat{J}$ by the environment. The time at which a jump happens is non-deterministic in this view. and its probability is given by $p_1 = \langle \hat{M}_1^\dagger \hat{M}_1 \rangle(t) = \langle \hat{J}^\dagger \hat{J} \rangle(t) \delta t$. We refer to a specific realisation of the stochastic process as a *quantum trajectory*. In this picture, the density matrix is interpreted statistically, and represents our ignorance of the particular realisation of the outcome record $r(t)$.
When $\hat{M}_1$ is measured ($r(t) = 1$), the state vector becomes

$$|\psi_1(t+\delta t)\rangle = \frac{\hat{M}_1 |\psi(t)\rangle}{\sqrt{p_1(t)}} = \frac{\hat{J}}{\sqrt{\langle \hat{J}^\dagger \hat{J} \rangle(t)}} |\psi(t)\rangle \,. \tag{1.41}$$

This corresponds to the collapse of the wave function upon the environment having successfully measured the outcome of $\hat{L}$. If instead $\hat{M}_0$ is measured ($r(t) = 0$), corresponding to no jump having occured with associated probability $p_0 = 1 - p_1$, the system's state vector becomes

$$|\psi_0(t+\delta t)\rangle = \frac{\hat{M}_0 |\psi(t)\rangle}{\sqrt{p_0}} = \Big\{ \hat{\mathbb{1}} - \mathrm{d}t \Big[i\hat{H} - \frac{1}{2}\hat{J}^\dagger \hat{J} + \frac{1}{2}\langle \hat{J}^\dagger \hat{J} \rangle(t) \Big] \Big\} |\psi(t)\rangle \,. \tag{1.42}$$

Interestingly, in this picture, the coupling of the system to its environment, even in the event of no jump, influences the time-evolution of the system. By combining Eqs. (1.41) and 1.42, and considering infinitesimal time intervals dt and outcome records $dr(t)$[5], the evolution of the wave function is given by the following stochastic differential equation [52]:

$$\begin{aligned} d|\psi(t)\rangle &= (1 - dr(t)) |\psi_0(t+dt)\rangle + dr(t) |\psi_1(t+dt)\rangle - |\psi(t)\rangle \\ &= dt \Big(\frac{\langle \hat{J}^\dagger \hat{J} \rangle(t) - \hat{J}^\dagger \hat{J}}{2} - i\hat{H} \Big) |\psi(t)\rangle + dr(t) \Big(\hat{J} / \sqrt{\langle \hat{J}^\dagger \hat{J} \rangle(t)} - 1 \Big) |\psi(t)\rangle \end{aligned} \tag{1.43}$$

[5] $dr(t)$ is an infinitesimal stochastic variable, also known as a Wiener process.

which is known as the stochastic Schrödinger equation. This leads to the following picture: the wavefunction evolves in time according to a non-Hermitian Hamiltonian, with sudden quantum jumps occuring at random times. Particular realisations of this process correspond to conditional single trajectories $\{|\psi^{(i)}\rangle\}_i$. It can be shown that one recovers the density matrix $\hat{\rho}(t)$ at a given time by averaging over infinitely many trajectories

$$\hat{\rho}(t) = \lim_{N_{\text{traj}}\to+\infty} \frac{1}{N_{\text{traj}}} \sum_{i=1}^{N_{\text{traj}}} |\psi^{(i)}(t)\rangle\langle\psi^{(i)}(t)| \,. \tag{1.44}$$

thus the map given in Eq. (1.43) is equivalent to that of the Lindblad master equation, Eq.(1.38). It follows that any observable can be obtained analogously:

$$\text{Tr}[\hat{\rho}(t)\hat{O}] = \lim_{N_{\text{traj}}\to+\infty} \frac{1}{N_{\text{traj}}} \sum_{i=1}^{N_{\text{traj}}} \langle\psi^{(i)}(t)|\hat{O}|\psi^{(i)}(t)\rangle \,. \tag{1.45}$$

In the numerical methods section, we will see how this translates to a straightforward numerical method to reduce the complexity of integrating the Lindblad master equation in time.

III.6 Numerical methods

Now that the main theoretical framework to treat open quantum systems has been presented, let us examine in more detail some of the numerical methods that have been developed over the years to study such systems. There are two main objects of study with open quantum systems: (i) finding the steady-state and/or the Liouvillian gap, which is the distance between the two smallest eigenvalues of the Liouvillian, important in the study of dissipative phase transitions, and (ii): simulating the time-evolution to investigate transient regimes. This section and most results in this manuscript mostly focus on (ii).

Monte-Carlo wave function (MCWF)

A widely used numerical method that naturally appears from the stochastic Schrödinger equation is the Monte-Carlo wave function (MCWF) method [53]. This method consists in evolving a quantum state according to a non-Hermitian Hamiltonian, and applying quantum jumps by randomly choosing from the set of jump operators $\hat{J}_i$ describing the system-environment interaction, according to some probability. The algorithm to simulate the dynamics of an open quantum system can be decomposed in the following steps:

1. Draw a random number a uniformly between 0 and 1.

2. Time-evolve the wavefunction $|\psi(t)\rangle$ according to the non-Hermitian Hamiltonian $\tilde{H}$ until time t', defined by $\langle\psi(t')|\psi(t')\rangle = a$.

3. Compute the jump probabilities p_i associated to each jump operator $\hat{J}_i$, given by

$$p_i = \langle\psi(t')|\hat{J}_i^\dagger \hat{J}_i|\psi(t')\rangle = \langle\hat{J}_i^\dagger \hat{J}_i\rangle(t') \tag{1.46}$$

4. Select the ith jump operator with probability p_i and apply it to $|\psi\rangle(t')$ along with renormalizing it as:
$$\left|\tilde{\psi}(t')\right\rangle = \frac{\hat{J}_i}{\sqrt{\langle \hat{J}_i^\dagger \hat{J}_i \rangle(t')}} |\psi(t')\rangle . \tag{1.47}$$

5. Repeat until the desired time is reached by restarting step 1 with the state $\left|\tilde{\psi}(t')\right\rangle$.

This algorithm has many variants, but this is computationally the most efficient [54] with respect to other implementations. By repeating this N_{traj} times, one obtains dynamics that are closer to the exact solution. However, how many trajectories are needed in practice highly depends on the problem, and no general procedure exists to determine N_{traj} [55]. The complexity of this method is given by $O(N_{\text{traj}} l^L)$, an exponential advantage with respect to a full integration of the Lindblad master equation. This is particularly useful for highly dissipative systems and when the number of dissipative channels is high, as we will see in more detail in chapter 3.

Mean-field and cluster approaches

The ideas presented in section II.3 for mean-field ansätze can be naturally extended to open quantum systems. Rather than taking the wavefunction to be a product of local wavefunction, we take the *density matrix* to be a product of local density matrices:

$$\hat{\rho} = \bigotimes_{i=1}^{N_c} \hat{\rho}_i, \tag{1.48}$$

with N_c ($N_c = L$ in the Gutzwiller mean-field approach) the number of clusters and $\hat{\rho}_i \in \mathcal{B}(\mathcal{H}_i)$. As for closed systems, this approach limits quantum correlations to be inside the cluster, meaning it will work best for weakly and/or locally interacting systems. This has the same value as for closed systems, reducing the number of free parameters required to describe the system density matrix $\hat{\rho}$ from l^L to $N_c l^{N_c/N}$, with N_c/N the size of the cluster. This leads to N_c/N equations to solve to integrate the master equation in time, leading to a more tractable solution. This approach has been quite successful in capturing a certain number of features of dissipative phase transitions [56] and nonequilibrium properties of strongly-correlated open quantum systems [57].

Matrix Product Operators

Tensor networks can also be applied to open quantum systems. The problem amounts to finding an expression for the many-body density matrix coefficients in terms of local tensors, much like Eq. (1.16). An additional issue when dealing with density matrices is that positivity must be enforced, without ever constructing the full density matrix[6]. A simple ansatz can be constructed by using the Choi isomorphism [58] that associates a density matrix to a pure state in an enlarged Hilbert space of dimension l^{2L}. The then

[6]Naively, enforcing positivity can be done by diagonalizing the full matrix, but the method then stops to be computationally efficient.

vectorised density matrix can be described like an MPS in this Hilbert space, and the density matrix coefficients defined in Eq.(1.2) can be written as a tensor product:

$$c^{s'_1,s'_2,\ldots,s'_L}_{s_1,s_2,\ldots,s_L} = A^{i_1,i_2}_{s_1,s'_1} A^{i_2,i_3}_{s_2,s'_2} \cdots A^{i_N,i_1}_{s_L,s'_L} \tag{1.49}$$

which is equivalent to the MPS ansatz with an additional dimension for each tensor that corresponds to the set of spin indices $\{s_i\}$. This is known as the matrix product operator (MPO) ansatz. While it has proved effective and yielded interesting results for 1D systems, one of the main issues of such a construction is that the positivity of the density matrix is not ensured when performing time evolution of the ansatz. Note that this problem has been proved to be computationally hard [59], hence no approach can be a perfect solution[7].

A partial solution to this issue is to built what is known as locally-purified density operators (LPDO) [60]. This is done by constructing a locally enlarged Hilbert space, built as $\mathcal{H} \otimes \mathcal{A} = \mathcal{H}_1 \otimes \mathcal{A}_1 \otimes \mathcal{H}_2 \otimes \mathcal{A}_2 \otimes \ldots \mathcal{H}_L \otimes \mathcal{A}_L$ with $\mathcal{A}$ the ancilla space. A wavefunction is then written as an MPS in this basis, whose coefficients are given by:

$$\langle s_1, a_1 s_2, a_2, \ldots, s_L, a_L|\psi\rangle = c_{s_1,a_1 s_2,a_2,\ldots,s_L,a_L} = A^{i_1,i_2}_{s_1,a_1} A^{i_2,i_3}_{s_2,a_2} \cdots A^{i_N,i_1}_{s_L,a_L}. \tag{1.50}$$

The density matrix coefficients are then obtained by tracing out the ancilla sites on $|\psi\rangle\langle\psi|$:

$$c^{s'_1,s'_2,\ldots,s'_L}_{s_1,s_2,\ldots,s_L} = B^{i_1,i_2}_{s'_1,s_1} B^{i_2,i_3}_{s'_2,s_2} \cdots B^{i_N,i_1}_{s'_L,s_L}. \tag{1.51}$$

Note that this expression is not equivalent to the MPO ansatz, as the B tensors are obtained by doing a partial trace on the wave function $|\psi\rangle$, and the MPO simply corresponds to a vectorized density matrix. The LPDO ansatz ensures the positivity of the density matrix, as it is obtained by a partial trace on a quantum state. However, some works have shown that the required bond dimension to obtain a similar precision is higher with LPDOs than MPOs [61, 62]. This hints to a lower capacity of LPDO to represent highly entangled states. Finally, we note that most works employ the MPO ansatz without worrying about the positivity of the resulting density matrix while obtaining highly accurate results [63, 64].

Corner-space renormalization

A method that is of particular interest for strongly-correlated systems is the corner-space renormalization method [65]. This method is rather different from mean-field or MPO methods, as it is constructed from the idea that there exists a class of physical systems, *low-entropy* quantum systems, for which the density matrix is accurately described by only a restricted set of M states. Mathematically, this is written as

$$\hat{\rho} = \sum_k^N p_k \, |\phi_k\rangle\langle\phi_k| = \sum_k^M p_k \, |\phi_k\rangle\langle\phi_k| + O(\epsilon_M) \tag{1.52}$$

with $\epsilon_M = 1 - \sum_k^M p_k \ll 1$. This means that one can truncate the eigenvalues to keep only M states, and the truncated density matrix will still faithfully describe the physical system at hand.

[7] unless P=NP.

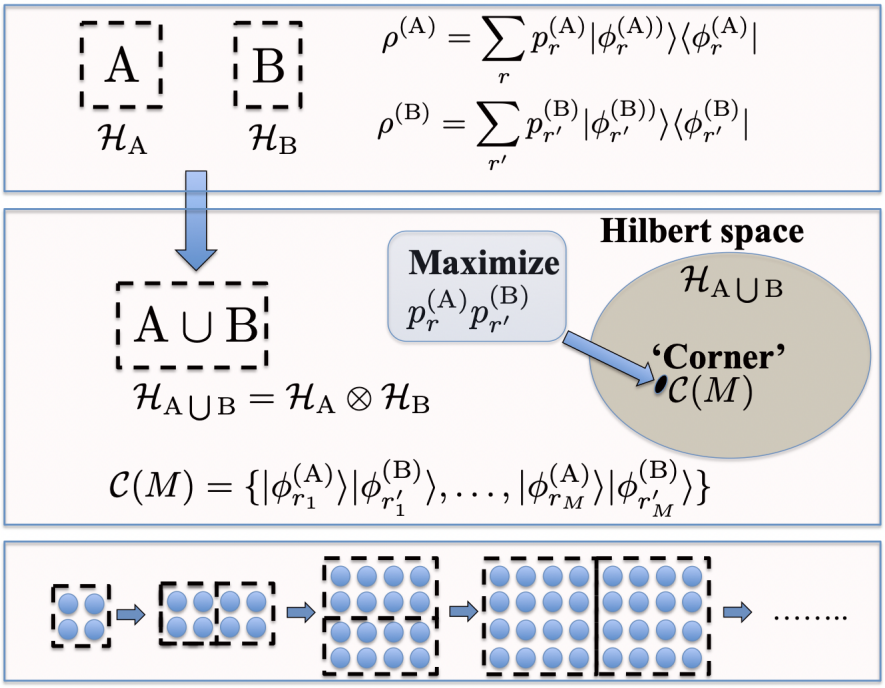

Figure 1.6: Sketch of the corner-space renormalization method. Taken from [65].

The corner-space renormalization method aims to solve the steady-state $\hat{\rho} \in \mathcal{H}$ of an open quantum system, described by Eq.(1.38). The basic idea behind this method is shown in Fig.1.6. It can be decomposed in three steps, as pictured by the three light rectangles forming the figure:

1. Take two-subsystems A and B whose steady-state can be found exactly, and obtain the steady-state by brute-force integration. Diagonalize the two corresponding steady-states $\hat{\rho}^{(A)}, \hat{\rho}^{(B)}$ to obtain the form given in Eq.(1.52).

2. Merge the two sub-systems into a larger system $A \cup B$. Construct a basis $\mathcal{C}(M)$, spanned by product states of the form $\left|\phi_r^{(A)}\right\rangle \left|\phi_{r'}^{(B)}\right\rangle$, and retain only the M most probable ones by keeping only the largest joint probabilities $p_r^{(A)} p_{r'}^{(B)}$. Solve the Lindblad master equation in this subspace of dimension M.

3. Repeat until the size of the full system is reached.

Note that this method allows for the states in the solution to be entangled between subsystems A and B: product states are constructed as the basis for the solution of the steady-state, hence this steady-state can comprise any superpositions of these product states. In this sense, this method is more powerful than mean-field or tensor network methods. Its limitation mainly resides in the cut-off of the M states, which will be limiting for high-temperature states with a large entropy. Note that the complexity of the algorithm for this method resides in the diagonalization of the density matrix at the end of the first step. This complexity of this step is $O(\max[M^3, l^{3N_{A,B}}])$, since the first

step is done with exact diagonalization of subsystems A, B with Hilbert space dimension $l^{N_{A,B}}$, and subsequent steps concern density matrices of size $M \times M$.

A method to simulate the dynamics of open quantum systems, based on the same principle that only M states may suffice to describe low-entropy systems, will be introduced in chapter 3.

III.7 Experimental systems

As mentioned, any quantum system that couples to an external environment, provided the Born-Markov approximation is satisfied, can be depicted as an open quantum system.

Superconducting circuits

A class of experimental setups that can be described as open systems are superconducting circuits. These circuits have proved immensely useful for the development of quantum information in the past two decades [66], as well as for exploring many-body physics [67]. One of the building blocks of such systems are electromagnetic LC resonators, composed of a capacitor C and an inductor L. One can describe a linear resonator with the Hamiltonian

$$\hat{H} = \frac{1}{2C}\hat{Q}^2 + \frac{1}{2L}\hat{\phi}^2 \tag{1.53}$$

where $\hat{Q}$ is the quantized charge of the capacitor, and $\hat{\phi}$ is the quantized magnetic flux flowing through the inductor. These operators are described by the same commutation rules as for position and momentum, i.e. $[\hat{Q}, \hat{\phi}] = -i$, hence the Hamiltonian can be diagonalized in exactly the same way as for a harmonic oscillator, by introducing the bosonic annihilation (and creation) operator, defined as:

$$\hat{a} = \frac{1}{\sqrt{\omega}}\left(\frac{\hat{\phi}}{\sqrt{2L}} - i\frac{\hat{Q}}{\sqrt{2C}}\right) \tag{1.54}$$

and the Hamiltonian is written as

$$\hat{H} = \omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right). \tag{1.55}$$

Such a Hamiltonian has a linear spectrum of n levels corresponding to n bosonic excitations, and its energy levels are given by

$$E_n = \omega\left(n + \frac{1}{2}\right), \tag{1.56}$$

with ω the frequency of the resonator.
What one really wants to investigate issues related to nonequilbrium dynamics is nonlinear terms in the Hamiltonian, so that the states involved in the dynamics are strongly correlated. The key ingredient to creating a nonlinear superconducting circuit is the Josephson junction [68], that adds nonlinearities to resonators. These are hybrid structures, composed of two superconductors separated by an insulating barrier, that behave as nonlinear inductors. The Hamiltonian term describing this element is given by

$$\hat{H}_J = E_J \cos\hat{\phi}_J \tag{1.57}$$

with $E_J = I_0\phi_0/2\pi$ (I_0 is the current passing through the junction, ϕ_0 is the flux quantum) and $\hat{\phi}_J$ is the flux flowing through the junction, that is again quantized. The cosine term is at the origin of the nonlinearity. By building a circuit made of a capacitor and a Josephson junction, one has the following Hamiltonian, that can be Taylor-expanded by supposing that the system is in a weakly nonlinear regime:

$$\hat{H} = \frac{1}{2C}\hat{Q}^2 + E_J \cos\hat{\phi}_J \tag{1.58}$$

$$\hat{H} = \frac{1}{2C}\hat{Q}^2 + E_J\left(1 - \frac{1}{2}\hat{\phi}_J^2 + \frac{1}{24}\hat{\phi}_J^4\right) + O\left(\hat{\phi}_J^6\right). \tag{1.59}$$

The spectrum is now slightly modified, and the energy separation between levels is not homogeneous. This leads to the following Hamiltonian, in terms of creation and annihilation operators $\hat{b}$, $\hat{b}^\dagger$:

$$\hat{H} = \omega\left(\hat{b}^\dagger\hat{b} + \frac{1}{2}\right) - E_J - E_C\left(\hat{b}^\dagger + \hat{b}\right)^4 \tag{1.60}$$

with $\omega = \sqrt{8E_CE_J}$, $E_C = e^2/2C$, e being the charge of the electron [18]. This leads to being able to realize two-level systems known as *transmon qubits*, crucial for quantum information provided the nonlinear terms are strong enough, such that a resonant drive at energy ω only couples the first two levels.
Alternatively, one can study nonlinear lattice models with such platforms, such as the Bose-Hubbard model. This model is described by the following Hamiltonian:

$$\hat{H} = \sum_{i=1}^{L}\left(\omega_c\hat{b}_i^\dagger\hat{b}_i + \frac{U}{2}\hat{b}_i^\dagger\hat{b}_i^\dagger\hat{b}_i\hat{b}_i\right) - J\sum_{i=1}^{L-1}\left(\hat{b}_i^\dagger\hat{b}_{i+1} + H.c.\right), \tag{1.61}$$

for a 1D lattice with open boundary conditions, whose features will be studied in more detail in Chapter 2. ω_c is the cavity frequency, J the nearest-neighbour coupling, and U the boson-boson interaction strength. The coupling J can be engineered by capacitive couplings, that are highly tunable elements [66]. Such systems exhibit natural loss channels, due to their coupling to the external environment. A simple model for such loss channels is that of local single-particle losses, that are described by jump operators

$$\hat{J}_i = \sqrt{\Gamma_1}\hat{b}_i. \tag{1.62}$$

Another ingredient that can be added is an external driving, that can be coherent (entering the Hamiltonian) or incoherent (entering the dissipative part of the Liouvillian). For a coherent drive in the rotating wave approximation, the following terms may be added to the Hamiltonian

$$\hat{H}_D = \sum_i F_i(\hat{b}_i^\dagger + \hat{b}_i). \tag{1.63}$$

This leads to richer physics, as the system consists of driven-dissipative coupled nonlinear elements. Many interesting phases can be engineered, and the physics of such systems has only begun to be explored.
In Fig. 1.7, a driven-dissipative superconducting platform is depicted. The full system is about 10 mm^2, and contains eight coupled transmon resonators, whose observables can be measured thanks to readout resonators. The chain of transmons is driven by a coherent

driving on an external site, and another site is coupled to the system, acting as an energy dump. This enabled authors of ref. [69] to stabilize strongly correlated phases, coined a *Mott insulator of photons*, that was never accessed before.

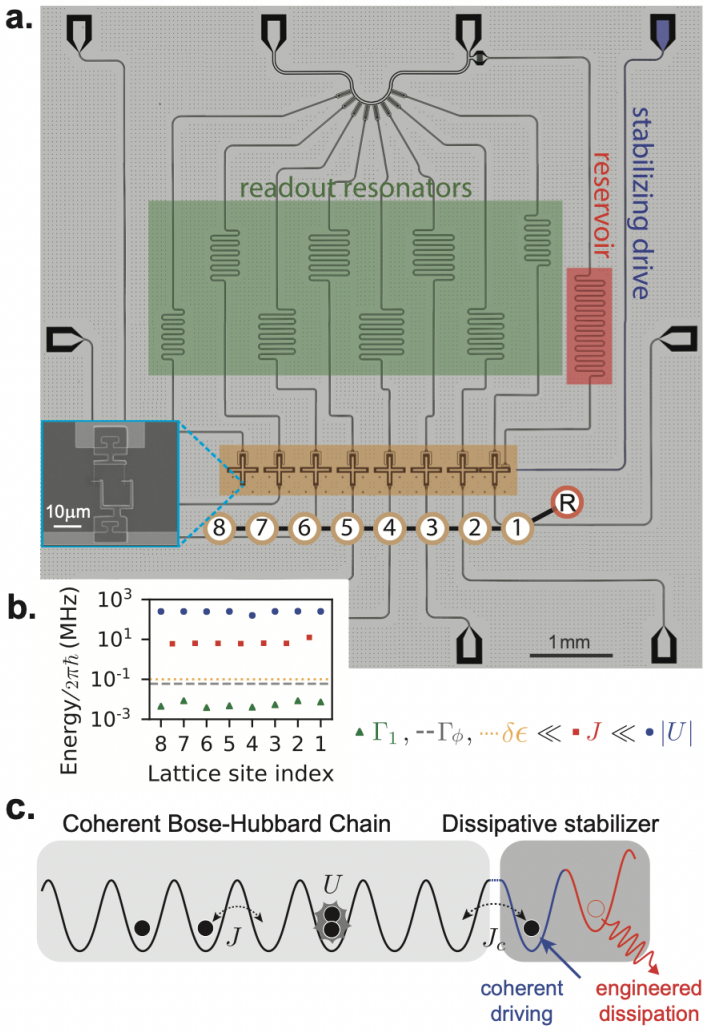

Figure 1.7: (a): Optical image of a driven-dissipative superconducting circuit where a chain of transmons, labeled from 1 to 8, are coupled to a reservoir and to a stabilizing drive. Observables can be measured by the readout resonators. The inset to the figure is a scan-electron microsopy image of a single transmon qubit. (b) Energy scales in such a system, where the interaction dominates all other energy scales. (c) Sketch of the effective model obtained, where a strongly-interacting Bose-Hubbard model is realized, coupled to two sites that act as a stabilizer for Mott insulator states. Reproduced from [69].

2 Entanglement dynamics in strongly-correlated open quantum systems

I Introduction

After being a subject of early intense debate at the dawn of quantum mechanics [70, 71], entanglement is now recognized as a key feature of quantum physics [72]. The efforts towards building a complete mathematical description of this notion were instrumental in the development of quantum information. In this context, the core of the theory is centered around three main tasks: detecting [73], quantifying [74] and manipulating entanglement [75]. The progress made on these three fronts would allow to outperform classical methods in the fields of metrology [76], cryptography [77] and computation [28].

In addition to providing sound foundations to the field of quantum information, entanglement theory has also paved the way to new discoveries in other areas of physics. As anticipated at the beginning of the millenium [78], quantities such as the entanglement entropy have proved to be very valuable tools for characterizing the ground-state wave function of many-body quantum systems [79–81]. The study of entanglement in many-body systems has not been restricted to their ground state properties: entanglement dynamics and its propagation in space in quantum systems has also been the subject of intense research activities for spin chains [82–84], fermionic [20] and bosonic systems [33, 85–87]. These works were important in inspecting the validity and limits of predictions about Lieb-Robinson bounds in lattice systems, as well as providing information about the properties and excitations of complex many-body systems.

Extending these investigations to open quantum systems represents a timely frontier of research [88] that is of fundamental importance because much less is known with respect to the state-of-the-art in isolated quantum many-body systems at thermal equilibrium or exhibiting unitary Hamiltonian dynamics. Whereas in general experimentalists try to protect their system from interacting with its environment, other approaches based on the general concept of "reservoir engineering" try to exploit the openness of a system and take advantage of judiciously designed dissipation to reach non-trivial quantum states in the transient regime [89] or in the steady state [69]. In recent years, experimental progress in tailoring effective photon-photon interactions in cavity and circuit quantum electrodynamics (QED) devices has lead to the emergence of controllable quantum optical many-body systems [90–92]. Unlike most condensed matter setups where the system is close to thermal equilibrium, this new class of systems are open quantum platforms in which intrinsic losses, due to the photon finite lifetime, have to be compensated by an

external coherent or incoherent driving.

In the past years, several works have been devoted to transport properties of strongly-correlated photonic platforms [93–96]. However, only a few exist regarding entanglement and correlation propagation in driven-dissipative systems, which mainly focus on free fermion systems [20, 97]. The recent experimental demonstration of dissipatively stabilized photonic Mott insulators [69] in chains of superconducting microwave resonators paves the way to the exploration of such an exciting frontier.

In this chapter, after presenting the physics of strongly-interacting bosonic systems, we will explore entanglement propagation in photonic Mott insulators, showing genuine physical effects associated to the openness of such systems. In contrast to most works about correlation propagation in interacting bosonic systems, here we do not consider global quenches of the system that typically consist in abruptly changing the value of the interaction strength in all the lattice [33, 87]. Instead, we consider two configurations where one photon is injected or removed from one cavity in the middle of a chain and investigate the propagation of entanglement that is produced between distant sites as a function of time and of their spatial separation. Such a study is achieved by monitoring the negativity of two-site reduced density matrices, that witnesses entanglement. We show a strinking different role of photon dissipation in the two configurations.

II The Bose-Hubbard model

Let us now present a few features of the paradigmatic Bose-Hubbard model. This model consists of L interacting bosonic modes, described by the following Hamiltonian:

$$\hat{H} = \sum_{i=1}^{L} \left(\omega_c b_i^\dagger b_i + \frac{U}{2} b_i^\dagger b_i^\dagger b_i b_i \right) - J \sum_{i=1}^{L-1} (b_i^\dagger b_{i+1} + h.c.), \tag{2.1}$$

with ω_c the cavity mode frequency, U the photon-photon (Kerr) on-site interaction, J the nearest-neighbor photon hopping coupling, and b_i ($b_i^\dagger$) the annihilation (creation) photon operators for each site. It presents a Mott phase in the strongly interacting ($U \gg J$) limit and a superfluid phase in the weakly interacting ($U \ll J$) regime. In this section we restrict the discussion to one-dimensional systems for simplicity. The physical systems described by the Bose-Hubbard Hamiltonian include, but are not limited to, lattices of microwave resonators in circuit QED platforms as described in Chapter 1 [98–103], semiconductor microcavities [99, 104] and ultracold gases in optical lattices [105, 106]. These platforms, as most realistic many-body systems, exhibit dissipation and dephasing due to the coupling to the environment. In cold atom systems dephasing is dominant [87] while for microwave photons in circuit QED platforms particle loss is typically the most important channel [107, 108].

II.1 Strongly correlated regime

One can start by considering the limit $U/J \to \infty$. In this limit, one can exactly perform a fermionization procedure in 1D, which enables one to solve the model [109]. In such a

limit, bosons become impenetrable, and the Hamiltonian reads:

$$\mathcal{H} = \omega_c \sum_{i=1}^{L} b_i^\dagger b_i - J \sum_{i=1}^{L-1} (b_i^\dagger b_{i+1} + h.c.), \tag{2.2}$$

which corresponds to a tight-biding model with a local energy ω_c. By using a Jordan-Wigner transform, this model is exactly solvable [109, 110].

In the following we will focus on a more interesting and practical case, where $U \gg J$ (but not so much larger). This does not completely wash out interactions, hence there will be some interplay between interactions and dissipative processes. In such a regime, in order to describe the physics of a photonic Mott insulator with one photon per site on average, we can safely truncate the local Hilbert space to a maximum of two photons per site by retaining only the $|0\rangle$, $|1\rangle$, and $|2\rangle$ Fock number states. The validity of this assumption was carefully tested numerically by increasing the local Hilbert space cutoff and checking that the results were unchanged. A Mott insulator phase corresponding to one photon per site for $U \gg J$ is approximately described by the factorized state

$$|\Psi_{\text{Mott}}\rangle = |1\rangle_1 \otimes |1\rangle_2 \otimes ... \otimes |1\rangle_L = |11...1\rangle \,. \tag{2.3}$$

In the regime of strong interactions, the Hamiltonian can be diagonalized by using generalized Jordan-Wigner and Bogoliubov transformations [33], via a mapping to a spin-1 model. By considering L coupled three level systems, one can rewrite the Hamiltonian as interacting doublon and holon excitations. These fermionic quasiparticles are described by local creation operators $d_i^\dagger$ and $h_i^\dagger$ for doublons and holons respectively, such that $d_i^\dagger |1\rangle_i = |2\rangle_i$, $h_i^\dagger |1\rangle_i = |0\rangle_i$. We therefore have:

$$\hat{H} = \sum_{j=1}^{L} \hat{P} \left\{ -2J\hat{d}_j^\dagger \hat{d}_{j+1} - J\hat{h}_j^\dagger \hat{h}_{j+1} - \sqrt{2}J \left(\hat{d}_j^\dagger \hat{h}_{j+1}^\dagger - \hat{h}_j \hat{d}_{j+1} \right) + \text{H.c.} + \frac{U}{2} \left(\hat{n}_{d,j} + \hat{n}_{h,j} \right) \right\} \hat{P} \tag{2.4}$$

with $\hat{P} = \Pi_j \left(\mathbb{1} - \hat{n}_{d,j} \hat{n}_{h,j} \right)$. One sees that the complexity of the model resided in the projector $\hat{P}$, which contains non-quadratic terms. However, one can reasonably approximate $\hat{P} \approx \mathbb{1}$, since the probability of having both a doublon and a holon on the same site $|\langle \hat{n}_{d,j} \hat{n}_{h,j} \rangle|^2$ should be 0. By doing this, one obtains a quadratic Hamiltonian that we can diagonalize using a Bogoliubov transform. The eigenmodes of the system are given by:

$$\gamma_{d,k}^\dagger = u(k) d_k^\dagger + v(k) h_{-k}^\dagger \tag{2.5}$$

$$\gamma_{h,-k}^\dagger = u(k) h_{-k}^\dagger - v(k) h_k^\dagger \tag{2.6}$$

with

$$u(k) = \cos(\theta_k/2), v(k) = i\sin(\theta_k/2), \tag{2.7}$$

$$\theta_k = \arctan\left(\frac{\sqrt{32}J\sin k}{U - 6J\cos k}\right) \tag{2.8}$$

$$\hat{d}_k = \sqrt{\frac{2}{L+1}}\sum_i \sin(ki)\hat{d}_i \tag{2.9}$$

$$\hat{h}_{-k} = -\sqrt{\frac{2}{L+1}}\sum_i \sin(ki)\hat{h}_i \tag{2.10}$$

$$k = \frac{\pi n}{L+1}, n \in \mathbf{N} \text{ for open boundary conditions.} \tag{2.11}$$

The Hamiltonian can now be written in a diagonal form:

$$\hat{H} = \sum_k \epsilon_d(k)\hat{\gamma}^\dagger_{d,k}\hat{\gamma}_{d,k} + \epsilon_d(-k)\hat{\gamma}^\dagger_{h,-k}\hat{\gamma}_{h,-k} \tag{2.12}$$

with

$$\epsilon_d(k) = -J\cos k + \frac{1}{2}\sqrt{(U - 6J\cos k)^2 + 32J^2\sin^2 k} \tag{2.13}$$

$$\epsilon_h(-k) = J\cos k + \frac{1}{2}\sqrt{(U - 6J\cos k)^2 + 32J^2\sin^2 k}. \tag{2.14}$$

For a closed system, we can extract the velocity for the corresponding doublons and holons as the derivative of the eigenenergies with respect to k:

$$v_{d,h}(k) = \partial_k\epsilon_{d,h}(k) = \pm J\sin k + \frac{1}{4}\frac{12JU\sin k - 8J^2\sin k\cos k}{\sqrt{(U - 6J\cos k)^2 + 32J^2\sin^2 k}} \tag{2.15}$$

with $\pm$ respectively corresponding to doublons and holons. This enables one to obtain bounds on the quasiparticle speed, which for doublons reads $v^D_{max} = 4J\left[1 - \frac{4J^2}{U^2}\right] + \mathcal{O}\left(\frac{J^3}{U^4}\right)$ and $v^H_{max} = 2J\left[1 + \frac{17J^2}{2U^2}\right] + \mathcal{O}\left(\frac{J^3}{U^4}\right)$ for holons. These results will serve as a basis to understand what happens in terms of entanglement propagation and the effect of dissipation and decoherence in the case of an open system.

III Entanglement propagation in dissipative photonic Mott insulators

We now consider the system described by the Hamiltonian of Eq. (B.1) coupled to an external environment. Within an open quantum systems approach, the time evolution of the system density matrix ρ can be described by the following Lindblad master equation:

$$\frac{d\rho}{dt} = -\mathrm{i}[\hat{H}, \rho] + \frac{1}{2}\sum_{i=1}^{L}\sum_{\mathcal{C}} 2J_i^{(\mathcal{C})}\rho J_i^{(\mathcal{C})\dagger} - \{J_i^{(\mathcal{C})\dagger}J_i^{(\mathcal{C})}, \rho\}, \tag{2.16}$$

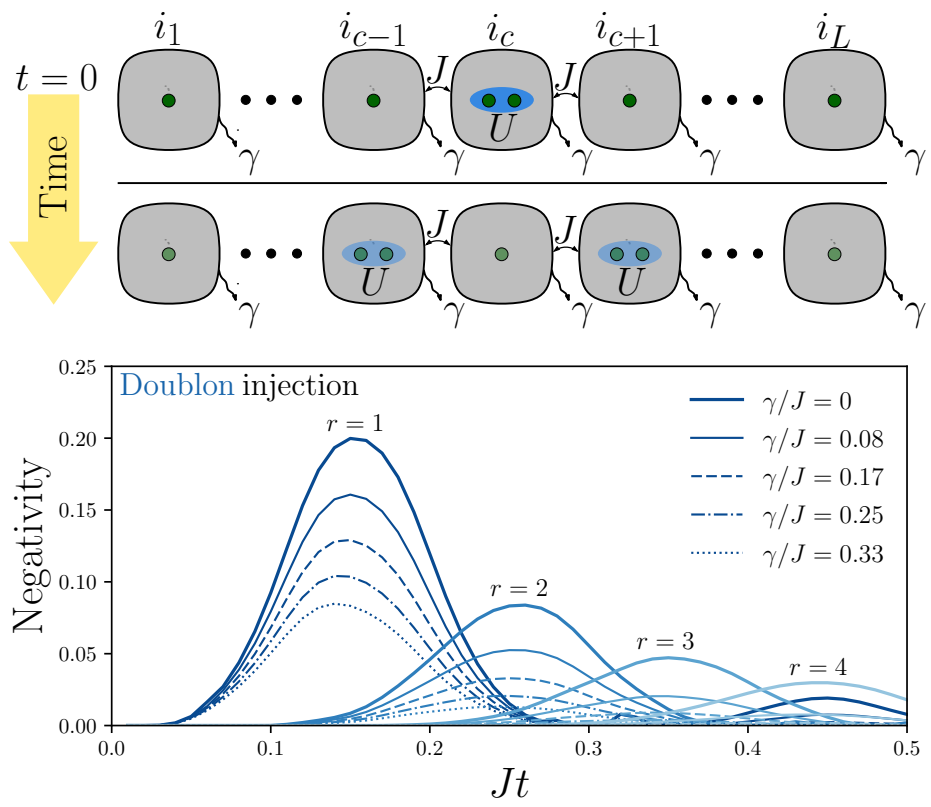

Figure 2.1: Upper panel: sketch of the considered system, a chain of coupled photonic resonators with on-site photon-photon interaction U and nearest-neighbor hopping coupling J. The top chain depicts the initial time configuration with a Mott insulator of photons (one photon per cavity) where a double occupation (doublon) has been created in the central site $i = i_c$. The bottom chain depicts the configuration at a later observation time t, with entanglement existing between distant sites. The photonic modes are subject to losses and dephasing. Lower panel: entanglement negativity $\mathcal{N}_r(t)$ between sites i_{c-r} and i_{c+r} as a function of time t for different values of the spatial separation $r = \{1, 2, 3, 4\}$ from the central site i_c. The shade of the lines gradually decreases going from $r = 1$ to $r = 4$. Calculations were performed with an MPO ansatz with bond link dimension $\chi = 200$ on a chain of $L = 20$ cavity sites. For each value of r, results for different values of the photon loss rate γ in units of the hopping J are shown. The initial state at $t = 0$ is $|\Psi_D\rangle$ (see the text) corresponding to a doublon excitation localized in the central site. In this figure, the pure dephasing rate Γ_d is 0. The on-site interaction for all the cavities is $U/J = 33.3$.

with $J_i^{(\mathcal{C})}$ the jump operator for the i-th site and the dissipation channel $\mathcal{C}$. When the temperature is low enough and the thermal photon occupancy is negligible, the jump operator for the particle loss channel ($\mathcal{C} = l$) due to the finite photon lifetime reads $J_i^{(l)} = \sqrt{\gamma} b_i$, where γ is the photon loss rate. The pure dephasing channel ($\mathcal{C} = d$) due to fluctuations in the environment is described by the jump operator $J_i^{(d)} = \sqrt{2\Gamma_d} b_i^\dagger b_i$, with Γ_d the pure dephasing rate. The factor $\sqrt{2}$ here is put to ensure that the norm of the

jump operators to dephasing and dissipation are the same, to obtain fair comparisons.

III.1 Entanglement generation protocol

Since $|\Psi_{\rm Mott}\rangle$ is a factorized state, an interesting question is how to perturb such a photonic Mott insulator in order to create entanglement in a simple way and study its propagation in a direct fashion. In the following we will show that this is possible by injecting (or removing) one photon from an occupied site. As shown in the upper panels of Figs. 2.1 and 2.2, we will consider such manipulation on the central site of a linear chain of resonators. In the case of a photonic insulator with a large U, this can be achieved simply by applying a coherent π-pulse drive on the central site that induces a Rabi rotation from the $|1\rangle$ to the $|2\rangle$ (or to the $|0\rangle$) Fock number state in the considered site. We have explicitly verified that such operation can be performed with fidelity close to 1 thanks to the strong anharmonicity produced by the large on-site interaction U. This way, it is possible to prepare the state $|\Psi_D\rangle = |1...2...1\rangle$ ($|\Psi_H\rangle = |1...0...1\rangle$) where D (H) stands for doublon (holon), corresponding to the injection of a single localized excitation on top of the quasiparticle vacuum. In the doublon case, the excess photon in the central site can hop to the right nearest-neighbor site or, with the same probability, to the left site. Due to the symmetry of the chain with respect to the central site and the lack of which-path information, such propagation creates an entangled state that can propagate along the chain. For circuit QED platforms, a Mott insulator can be prepared and maintained through an active stabilization process [69]. In the following, we will consider the dissipative dynamics of the system after the creation of the localized doublon (holon) in the absence of stabilization.

III.2 Entanglement detection

In order to witness bipartite entanglement between two partitions A and B of the system, we have considered the negativity function [111], defined as $\mathcal{N}(\rho_{AB}) = \sum_{\lambda<0} |\lambda|$, where the sum is taken over the negative eigenvalues λ of $\rho_{AB}^{T_A}$, which is the partial transpose with respect to subsystem A of the joint density matrix ρ_{AB}. In the following, A and B will be two resonators at a distance r from the central cavity ($i = i_c$). The time-dependent negativity of the reduced density matrix for these two sites at the positions i_{c-r} and i_{c+r} will be denoted by $\mathcal{N}_r(t)$. For systems with a relatively moderate Hilbert space dimension, we can compute the time evolution of the full density matrix of the system via an exact integration of the master equation. Once we get the full density matrix, we can trace out with respect to the degrees of freedom of all the sites except the two sites under study.

In the regime where we can consider only the $|0\rangle$, $|1\rangle$, and $|2\rangle$ states as local basis for a given site, we can reconstruct a two-site reduced density matrix by exploiting the fact that any Hermitian operator can be decomposed over the generators of the group associated to its Hilbert space [112]. As we truncated the local Hilbert space to states with up to 2 photons, the generators of the SU(3) $\otimes$ SU(3) group allows us to reconstruct the reduced density matrix as:

$$\rho_3^{(2)} = \frac{1}{9} \sum_{i_1,i_2=0}^{8} r_{i_1 i_2} \Lambda^{(i_1)} \otimes \Lambda^{(i_2)}, \tag{2.17}$$

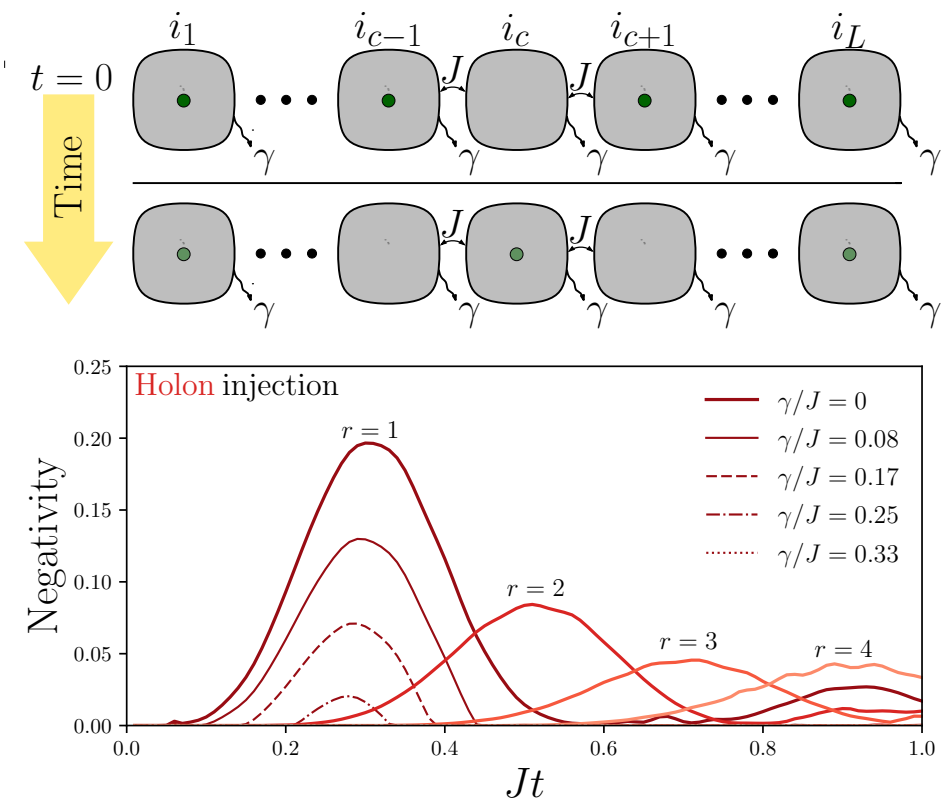

Figure 2.2: Upper panel: sketch like in Fig. 2.1, but where the initial state has an empty central site (holon). Lower panel: temporal dynamics of the negativity with same parameters as in Fig. 2.1, but with an initial state $|\Psi_H\rangle$ corresponding to a holon localized in the central site.

with $\Lambda^{(i)}$ the generators [1] of the SU(3) group and

$$r_{i_1 i_2} = \frac{9\langle \Lambda^{(i_1)} \otimes \Lambda^{(i_2)} \rangle}{\mathrm{tr}((\Lambda^{(i_1)} \otimes \Lambda^{(i_1)})^2)}.$$

Using this method we can reconstruct the two-site density matrix for all times and obtain the entanglement negativity. Since the Hermitian operators $\Lambda^{(i)}$ can also be expressed as a function of the bosonic creation and annihilation operators up to their third-order power [2], this tomographic method can be used in experiments to measure the two-site entanglement negativity. Indeed, the measurement of expectation values of moments of the photon fields has become a rather standard procedure in circuit QED platforms (see,

[1] The nonzero matrix elements of the Hermitian $\Lambda^{(i)}$ matrices, such that $\Lambda^{(i)}_{r,s} = (\Lambda^{(i)}_{s,r})^\star$, are the following ones: $\Lambda^{(0)}_{11} = \Lambda^{(0)}_{22} = \Lambda^{(0)}_{33} = 1$, $\Lambda^{(1)}_{12} = 1$, $\Lambda^{(2)}_{12} = -i$, $\Lambda^{(3)}_{11} = -\Lambda^{(3)}_{22} = 1$, $\Lambda^{(4)}_{13} = 1$,$\Lambda^{(5)}_{13} = -i$, $\Lambda^{(6)}_{23} = 1$, $\Lambda^{(7)}_{23} = -i$, $\Lambda^{(8)}_{22} = 1/\sqrt{3}$ and $\Lambda^{(8)}_{33} = -2/\sqrt{3}$.

[2] The eight generators $\Lambda^{(i)}$ can be expressed in terms of the bosonic annihilation and creation operators. Namely: $\Lambda^{(1)} = \frac{1}{2}(bb^{\dagger 2} + b^2 b^\dagger)$, $\Lambda^{(2)} = \frac{i}{2}(b^2 b^\dagger - bb^{\dagger 2})$, $\Lambda^{(3)} = \frac{3}{4}b^2 b^{\dagger 2} - \frac{1}{2}bb^\dagger$, $\Lambda^{(4)} = \frac{1}{\sqrt{2}}(b^{\dagger 2} + b^2)$, $\Lambda^{(5)} = \frac{i}{\sqrt{2}}(b^2 - b^{\dagger 2})$, $\Lambda^{(6)} = \frac{1}{\sqrt{2}}(b^{\dagger 2} b + b^\dagger b^2)$, $\Lambda^{(7)} = \frac{i}{\sqrt{2}}(b^\dagger b^2 - b^{\dagger 2} b)$, and $\Lambda^{(8)} = \frac{1}{\sqrt{3}}(bb^\dagger - b^\dagger b)$.

e.g. [113]). Note that we have conveniently developed and used this approach for numerical simulations based on the Matrix Product Operator (MPO) technique [93, 114, 115]. Indeed, MPO simulations are effective to simulate longer chains of cavities but do not allow for a direct access to the full system density matrix, an issue that we bypassed with the procedure described above.

IV Results and discussion

In Fig. 2.1, we report results for the negativity $\mathcal{N}_r(t)$ for different values of r and of the photon loss rate γ (here no pure dephasing is considered, $\Gamma_d = 0$). The negativity shows a well resolved peak for most values of the spatial separation r and γ/J: increasing r delays the negativity peak, showing a clear entanglement propagation. A revival peak of entanglement is visible in the $r = 1$ curve at longer times. Previous studies of two-qubit systems with non-Markovian environments have revealed entanglement revival effects [116–118]. In our system, the two-site dynamics is non-unitary and non-Markovian even for $\gamma = 0$ since the other sites of the chain have been traced out for the calculation of the negativity. The value of the negativity peaks decreases with increasing dissipation γ. However, it is remarkable that the entanglement propagation speed is negligibly influenced by dissipation and remains essentially ballistic. In Fig. 2.2, we report the analogous dynamics of negativity for the other configuration where the holon state $|\Psi_H\rangle$ is prepared. With the same parameters as in Fig. 2.1, in the holon case not only the propagation speed is slower, but the role of dissipation is more dramatic, as we do not see any peaks for the chosen values of γ/J as soon as $r > 1$.

IV.1 Propagation speed

In Fig. 2.3, we report the calculated entanglement propagation speeds versus J/U both for the case of photon injection (doublon excitation) and extraction (holon excitation). In the same plot, we have also reported the maximal propagation speed of doublons and that of holons for a closed system, namely $v^D_{max} = 4J\left[1 - \frac{4J^2}{U^2}\right] + \mathcal{O}\left(\frac{J^3}{U^4}\right)$, and $v^H_{max} = 2J\left[1 + \frac{17J^2}{2U^2}\right] + \mathcal{O}\left(\frac{J^3}{U^4}\right)$. The fact that the predicted entanglement speed for a closed system remains close to what we observed in the dissipative case is striking. This means that even in the presence of losses, the properties of the quasiparticle picture remain valid.

IV.2 Particle-hole asymmetry

The effect of dissipation for the two considered configurations is presented in Fig. 2.4 in which the peak value of $\mathcal{N}_{r=1}$ is plotted. To compare the genuine effect of dissipation and dephasing, we considered a holon propagation in a chain with hopping coupling $2J$ and a doublon propagation in a chain with hopping J in order to have the same speed (for $U/J \gg 1$ the speeds differ by a factor 2). From the peak value of the negativity (occurring at the same time), we see that pure dephasing acts on the two cases in the same identical way (dashed lines), with an exponential decay of the negativity. Indeed,

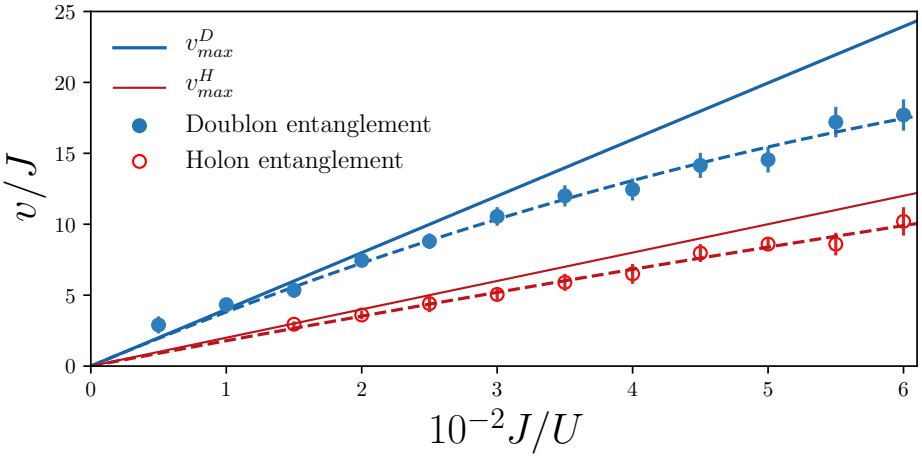

Figure 2.3: Entanglement propagation speed (units of J) versus the normalized hopping J/U. Filled (blue) circles: entanglement speed for the configuration corresponding to the injection of one additional photon (doublon injection). Empty (red) circles: entanglement speed for the holon injection. Solid (blue) thick line: maximal speed for doublon quasiparticles in a closed Hamiltonian system. Thin (red) solid line: maximal holon speed. Dashed lines are polynomial fits with linear and quadratic terms in J/U. Error bars were estimated taking into account uncertainty due to time discretization and to the finite bond link dimension in the MPO calculations. Parameters: $\gamma/J = 0.1$ and for the doublon (holon) configuration $U/\gamma = 100$ (1000).

pure dephasing conserves the total number of particles and does not break the particle-hole symmetry. On the other hand, in the presence of photon losses, our investigation reveals a striking asymmetry between the doublon (thick solid line) and holon (thin solid line) cases. Indeed, the negativity vanishes much faster for the holon configuration even when the speed is the same. The asymmetry can be qualitatively understood if we consider the quantum jump picture [53] for two sites that are entangled. In the quantum jump picture, the density matrix is seen as the statistical mixture of time-dependent quantum trajectories. These quantum trajectories represent the time-evolution of the wavefunction of the system along with sudden changes (quantum jumps) in the state of the wavefunction due to physical processes such as particle losses, that are inherently stochastic. In the holon case, by diagonalizing the reduced density matrix associated to two entangled sites, we have found that the entanglement is mostly due to the Bell state $|\psi_{H,+}\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_{i_{c-r}}|1\rangle_{i_{c+r}} + |1\rangle_{i_{c-r}}|0\rangle_{i_{c+r}}\right)$. In a single quantum trajectory picture, a single quantum jump due to photon decay in one of the two sites transforms such state into the factorized state $|0\rangle_{i_{c-r}} \otimes |0\rangle_{i_{c+r}}$. By contrast, in the doublon configuration, the entanglement is mostly due to the state $|\psi_{D,+}\rangle = \frac{1}{\sqrt{2}}\left(|2\rangle_{i_{c-r}}|1\rangle_{i_{c+r}} + |1\rangle_{i_{c-r}}|2\rangle_{i_{c+r}}\right)$. In this case, the quantum jump produced by a photon loss in site i_{c-r} produces the (normalized) state $\frac{1}{\sqrt{3}}\left(\sqrt{2}|1\rangle_{i_{c-r}}|1\rangle_{i_{c+r}} + |0\rangle_{i_{c-r}}|2\rangle_{i_{c+r}}\right)$, which is still entangled. A photon loss in site i_{c+r} produces an analogous state. The quantitative results in Fig. 2.4 show a remarkable non-exponential dependence on both doublon and holon negativity as a function of the

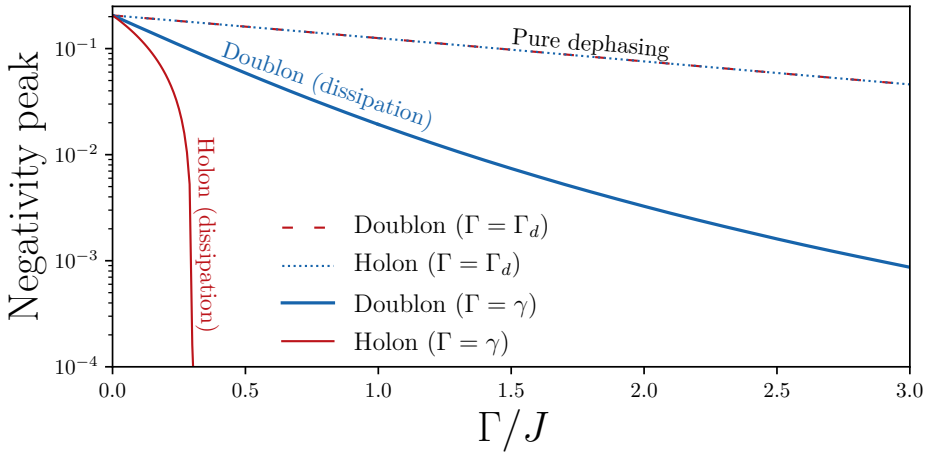

Figure 2.4: Peak value of $\mathcal{N}_{r=1}$ (log scale) as a function of the photon dissipation rate γ or dephasing Γ_d. Thick (thin) solid line: curve for the doublon (holon) configuration versus γ when only photon particle losses are present ($\Gamma_d = 0$). The holon injection case has been calculated for a chain with a hopping coupling $2J$ to have the same maximal speed of a doublon in a chain with hopping coupling J (see text). Dashed (doublon) and dotted (holon) lines: negativity peak versus Γ_d when there is only the pure dephasing channel ($\gamma = 0$). Parameters: $L = 5, U/J = 100$.

dissipation rate. The holon negativity becomes exactly zero for a finite value of the dissipation γ whereas the decay in the doublon case slows down for increasing values of γ. Indeed, the two-site reduced density matrix satisfies an effective master equation, which is in general non-markovian as obtained by tracing out the other degrees of freedom of the chain, whose dynamics is non-trivial.

V Conclusion

We have theoretically explored the physics of entanglement propagation in photonic Mott insulators in the presence of photon particle losses and dephasing. We have investigated a scheme where the entanglement is generated by injecting (or extracting) a photon from a site of a photonic Mott insulator. We have introduced a relatively simple quantum state tomography protocol, valid in the limit of strong photon-photon interactions, to study the bipartite entanglement properties. Our scheme is particularly suited for circuit QED platforms exhibiting strong photon-photon interactions and allowing the measurement of quantum optical correlation functions between distant sites. In spite of the losses, the propagation of the negativity peak exhibits a speed, which is close to the doublon (holon) quasiparticle propagation speed respectively in the case of the injection (extraction) of a photon. Remarkably, the impact of particle losses is highly asymmetric for these two configurations, while pure dephasing does not break the doublon-holon symmetry. Our work paves the way to new investigations on the entanglement propagation in open quantum systems, where multipartite entanglement spreading could also be investigated [119, 120].

A future interesting research direction is the characterization of the entanglement dynamics at long times (diffusive vs. ballistic) and the quest for universal features underlying the dynamics of correlations in this class of systems. Another challenging problem to be investigated in the future, given the recent experimental success in the dissipative stabilization of photonic Mott insulators [69], is the search for protocols to stabilize entanglement propagation in open quantum systems.

3 The dynamical corner-space method

In this chapter, we will introduce a numerical method to simulate the time-evolution of an open quantum system. This method is based on the same idea as the corner-space renormalization method, presented in Chapter 1, section III.6, which is that there exists a class of systems for which M pure states suffices to faithfully represent the density matrix, with $M \ll N$, N being the Hilbert space dimension. This new method, the *dynamical corner-space method*, efficiently computes the time evolution of open quantum systems that have a moderate entropy with controllable accuracy, depending on the effective rank M of the density matrix. It will here be applied to the continuous-time dynamics of noisy quantum circuits in the presence of dissipation and decoherence, beyond digital error models that are commonly used in the quantum information community [121].

I Introduction

As we have seen in the previous chapter, strongly-correlated open quantum systems have remained relatively unexplored due to the lack of existing numerical methods to tackle them. In particular, we are interested in methods that put no restriction on the amount of entanglement that quantum states exhibits, unlike tensor network and mean-field methods.

A standard method that does this job is the Monte-Carlo wave function method [53, 86]. As explained previously, this method reduces the problem to evolving N_{traj} wavefunctions, and in the $N_{\text{traj}} \to \infty$ limit one recovers the exact evolution. However, the number of required trajectories N_{traj} to obtain accurate results is not known in general [122] and, in the case of weak dissipation, the method can quickly become equivalent to a full integration of the master equation as a greater amount of trajectories are needed to reach convergence. In recent years, there has been a growing interest in the idea that for a certain class of low-entropy systems, a limited number of states, belonging to a so-called "corner" subspace, can efficiently and faithfully represent the density matrix [123–126]. As presented earlier, such a method, the corner-space renormalization method, exists and has proved to be extremely useful in the context of dissipative phase transitions. However, this method is restricted to finding steady-states. An alternative method is therefore needed to simulate the time-evolution of low-entropy strongly-correlated open quantum systems.

In the field of quantum information, the tremendous advances on the control of artificial quantum systems, such as superconducting Josephson qubits [67] and trapped ions [127], are allowing dramatic progress towards the realisation of devices for quantum computation [128, 129]. We are currently in the noisy intermediate-scale quantum (NISQ) era [30], where error correction is not yet possible due to daunting overheads [130], but where quantum advantage might be already exploited for applications in quantum chemistry [131], optimisation [132] and even finance [133]. Scaling up quantum circuits and designing practically efficient quantum correction protocols make it of crucial importance to precisely understand the effects of both incoherent and coherent sources of noise on quantum algorithms [134–137].

To meet these challenges, there is a strong need for accurate numerical simulations of quantum hardware on classical computers [138–142]. In this perspective, the application of tensor-network methods to quantum-circuit simulation has been shown to be effective to model circuits, although with restrictions on the amount of entanglement building up through the circuit [143–147]. In addition, most existing simulators of quantum hardware consider local and digital error models [121, 130, 148, 149], that consist in extending the quantum circuit model to noisy algorithms by including noise gates applied after each unitary gate. While in close proximity with classical error models, these two approximations do not necessarily hold, especially for highly-entangling circuits [150], and remain a challenge in quantum error correction [151]. To take into account realistic sources of noise for highly-entangling circuits one should revert to a continuous-time description, where noise is taken into account continuously during the dynamics associated to the quantum gates. If one neglects non-Markovian effects, this can be done in the framework of the Lindblad master equation [152]. However, such a description is numerically expensive; for a chain of L qubits with Hilbert space dimension $N = 2^L$, a density matrix of size N^2 must be evolved. Since quantum processors are conceived to be weakly dissipative and with low entropy, they belong to the class of systems that have a low-entropy. Stabilized arrays [69, 153], cat qubit systems [154] and quantum hardware with state-of-the-art dissipation rates [121, 155] also belong to this category.

In this chapter, we present the dynamical-corner space method that is particularly suited to simulate quantum processors in the NISQ era. We start by presenting the method and the steps that compose the corresponding algorithm in section II. In the next section, various computational details, such as the complexity of the algorithm, and the question of stiffness for time-integration are discussed in section III. Then, in section IV, we present an application of the method to the simulation of a noisy quantum Fourier transform processor. We discuss the applicability of our method to this quantum algorithm, and find scaling laws for the error as a function of the number of qubits, as well as the impact of dissipative channels on the performance of the algorithm as a function of the considered initial state. Finally, digital error models are discussed in more detail in section V, in which key differences between our approach to simulate quantum circuits and such models are discussed.

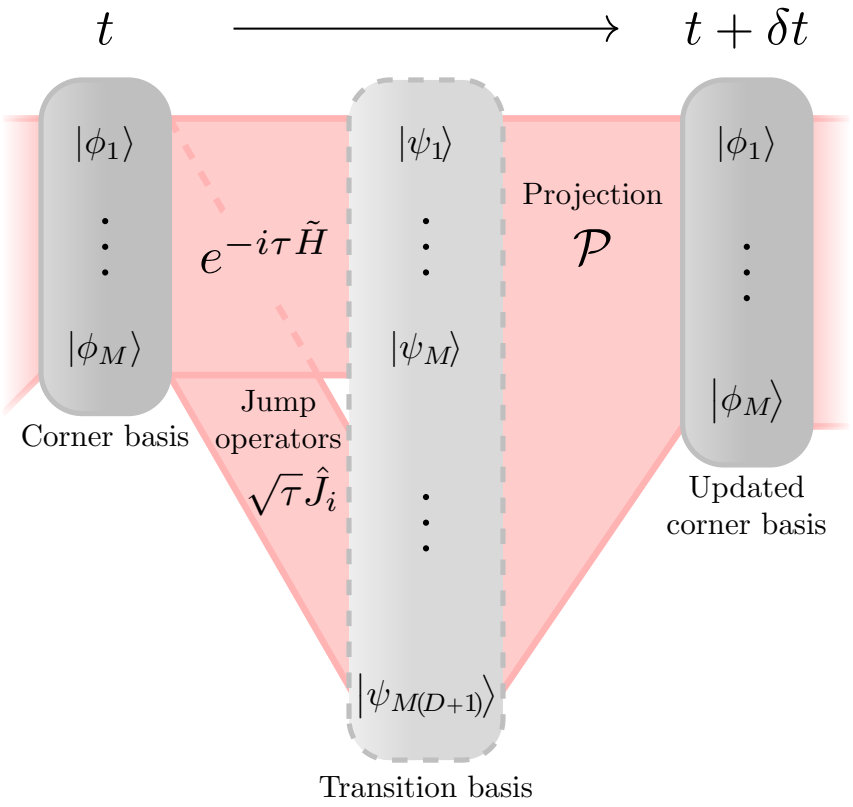

Figure 3.1: (a) Sketch depicting one iteration of the time-dependent corner-space method.

II The dynamical corner-space method

Let us consider an open quantum system whose dynamics is governed by the Lindblad master equation [152]:

$$\partial_t \hat{\rho} = \mathcal{L}\hat{\rho} = -\mathrm{i}[\hat{H}, \hat{\rho}] + \sum_{i=1}^{D} \Big(\hat{J}_i \hat{\rho} \hat{J}_i^\dagger - \frac{1}{2}\{\hat{J}_i^\dagger \hat{J}_i, \hat{\rho}\} \Big), \tag{3.1}$$

where $\hat{H}$ is the system Hamiltonian acting on a Hilbert space $\mathcal{H}$ of dimension N, and $\hat{J}_i$ is the ith jump operator. At any time t, the solution $\hat{\rho}$ may be approximated by

$$\hat{\rho}(t) \simeq \sum_{k=1}^{M(t)} p_k(t) \, |\phi_k(t)\rangle\langle\phi_k(t)| \,, \quad p_k(t) \geq p_{k+1}(t), \ \forall k, \tag{3.2}$$

where $p_k(t)$ are the $M(t)$ largest eigenvalues of $\hat{\rho}$ at the time t and $|\phi_k(t)\rangle$ their associated eigenvectors. By construction, the controlled truncation error introduced by such an approximation is monotonically decreasing with M and quantified by $\epsilon_M = 1 - \sum_{k=1}^{M} p_k$, so that the decomposition becomes exact for $M(t) = r(t)$, with $r(t)$ denoting the rank of $\hat{\rho}(t)$, equivalent to the $\alpha = 0$ Rényi entropy [156]. Therefore, in a wide class of low-entropy systems including most platforms relevant for quantum computing, $\hat{\rho}$ is very well approximated by $M \ll N$ basis vectors, and even by $M \gtrsim 1$ for close to pure states.

M will be referred to as the *corner dimension*. The accuracy of the calculations will be controlled by a fixed maximum error ϵ with $\epsilon_M \leq \epsilon$ enforced at any time.

It follows from the above that all the information of the density matrix is carried by a set of weighted corner base vectors of the form $\sqrt{p_k}\,|\phi_k\rangle$. In some arbitrary computational basis $\{|n\rangle\}_{n=1}^{N}$, these can be represented by a $N \times M$ matrix with elements $C_{nk}(t) = \sqrt{p_k(t)}\,\langle n|\phi_k(t)\rangle$. We have:

$$\hat{\rho}(t) = \sum_{k=1}^{M(t)} p_k(t)\,|\phi_k(t)\rangle\langle\phi_k(t)| = \hat{C}(t)\hat{C}^{\dagger}(t). \tag{3.3}$$

The idea behind the dynamical corner-space method is then to evolve $\hat{C}$ in time, without ever reconstructing $\hat{\rho}$. The evolution $\hat{C}(t) \mapsto \hat{C}(t+\delta t)$ over a small time step δt, schematically illustrated in Fig. 3.1, involves two computational operations: (i) the construction of the transition basis and (ii) the projection onto new principal components. Below we detail these two steps, respectively represented as the left and right parts of the figure.

Step (i): The weighted corner basis $\hat{C}(t)$ evolves into the transition basis $\hat{T}(t+\delta t)$ as

$$\hat{\rho}(t+\delta t) = \sum_{i=0}^{D} \hat{M}_i\hat{\rho}(t)\hat{M}_i^{\dagger} = \sum_{m=1}^{M(D+1)} |\psi_m(t+\delta t)\rangle\langle\psi_m(t+\delta t)| = \hat{T}(t+\delta t)\hat{T}^{\dagger}(t+\delta t). \tag{3.4}$$

Here, the following Kraus map [52], equivalent to Eq. (3.1), is used to time-evolve the density matrix:

$$e^{\delta t\mathcal{L}}\hat{\rho} = \sum_{i=0}^{D} \hat{M}_i\hat{\rho}\hat{M}_i^{\dagger}, \quad \hat{M}_0 = \exp(-i\delta t\tilde{H}), \quad \hat{M}_{i\geq 1} = \sqrt{\delta t}\hat{L}_i, \tag{3.5}$$

with $\tilde{H} = \hat{H} - \frac{i}{2}\sum_{i=1}^{D}\hat{L}_i^{\dagger}\hat{L}_i$ a non-Hermitian operator depending on the Hamiltonian and the quantum jump operators, which is reminiscent of the stochastic Schrödinger equation presented in section III.5. By construction, the transition basis $\hat{T}(t+\delta t)$ is a $N\times[M(t)(D+1)]$ rectangular matrix, where D is the number of dissipation channels. Its mth column is given by $|\psi_m(t+\delta t)\rangle = \sqrt{p_\mu}\hat{M}_\nu\,|\phi_\mu(t)\rangle$, with $\nu = (m-1) \div M(t)$ and $\mu = (m-1) \mod M(t)+1$. In the quantum trajectories picture, a Kraus operator is drawn at random at every time step according to the probability for an external observer to measure its outcome. This generated single trajectories conditioned by a specific record of the history of the outcomes, which by repeating the procedure could be averaged to recover a correct density matrix. We here employ a completely deterministic approach: at every time step, all the $(D+1)$ possible trajectories branching from each of the corner's M base state vectors are generated, naturally weighted by their likelihood. Although exact, no further processing would result in an exponential growth of the corner dimension with time, growing as $D^{n_{\delta t}}$, with $n_{\delta t}$ the number of timesteps. The second step of the algorithm solves this problem.

Step (ii): The transition matrix is now projected to a new weighted corner basis $\hat{C}(t+\delta t)$ of lower dimension $M(t+\delta t) \leq M(t)(D+1)$ via a new truncated eigendecomposition $\mathcal{P}$ of the form of Eq. (3.3). Crucially, this is possible without ever reconstructing the full

density matrix. Indeed, the $N \times N$ matrix $\hat{\rho}(t+\delta t) = \hat{T}(t+\delta t)\hat{T}^\dagger(t+\delta t)$ and the much smaller $[M(t)(D+1)] \times [M(t)(D+1)]$ matrix $\hat{\sigma}(t+\delta t) = \hat{T}^\dagger(t+\delta t)\hat{T}(t+\delta t)$ share the same non-zero eigenvalues p_k. The associated eigenvectors, $|\phi_{k,\rho}(t+\delta t)\rangle$ and $|\phi_{k,\sigma}(t+\delta t)\rangle$ are related by the identity [125, 157]:

$$\sqrt{p_k}\,|\phi_{k,\rho}(t+\delta t)\rangle = \hat{T}(t+\delta t)\,|\phi_{k,\sigma}(t+\delta t)\rangle\,. \tag{3.6}$$

The components of the new decomposition can then be judiciously truncated to only retain the leading $M(t+\delta t)$ eigenvalues p_k, yielding an updated weighted corner basis $\hat{C}(t+\delta t)$, with the same structure as the initial one $\hat{C}(t)$.

This procedure can be iterated for an arbitrary time with the possibility of tuning the length of the time step δt and the tolerance parameter ϵ to control the accuracy at the desired level. The time evolution of the corner basis involves no explicit reconstruction of the full density matrix. Indeed, the largest representation of the state of the system involved in the process, the transition basis $\hat{T}$, is comparable in size to that of $M(D+1)$ closed systems. As we will see, this method is comparable to MCWF in terms of complexity, where the the entropy is the essential quantity to check for: if the required corner dimension M is too large, the method fails terribly. For systems with moderate entropy instead, it can be very efficient.

III Computational details

III.1 Complexity

The method presented above amounts to evolving $M(t)$ closed systems, with $M(t)$ the corner dimension. For a quantum computation, the initial state is pure, $M(t=0) = 1$. As shown in Fig. 3.1(c), the dimension $M(t)$ grows moderately in time. At every time step, the most demanding operation is related to the construction of the matrix $\hat{\sigma} = \hat{T}^\dagger\hat{T}$. This involves a number of operations of order $O(M^2[D+1]^2 2^L)$. Indeed, the transition matrix $\hat{T}$ has a size $M(D+1) \times N$, where $N = 2^L$ denotes the dimension of the Hilbert space and L the size of the system under consideration. Note that the number of jump operators D typically scales as L for most relevant quantum computing platforms consisting of L coupled units. The complexity of the method is thus of order $O(M^2L^2 2^L)$. This represents an exponential reduction of the complexity with respect to that of a brute-force master equation integration, which is of order $O(4^L)$.

III.2 Memory use

The Kraus operators $\hat{K}_i$ are $N \times N$ matrices, the same size as the density matrix. However, their memory requirement is much smaller, thanks to their extreme sparsity for most. Indeed, jump operators corresponding to dissipative processes are typically single-body operators and thus as memory-consuming as state vectors[1]. The largest Kraus operator

[1]For fully-connected operators of the form $\hat{K} = \hat{a}_1 \otimes \hat{b}_2 \otimes \cdots \otimes \hat{z}_L$, this would be an issue, but such error models are quite unrealistic.

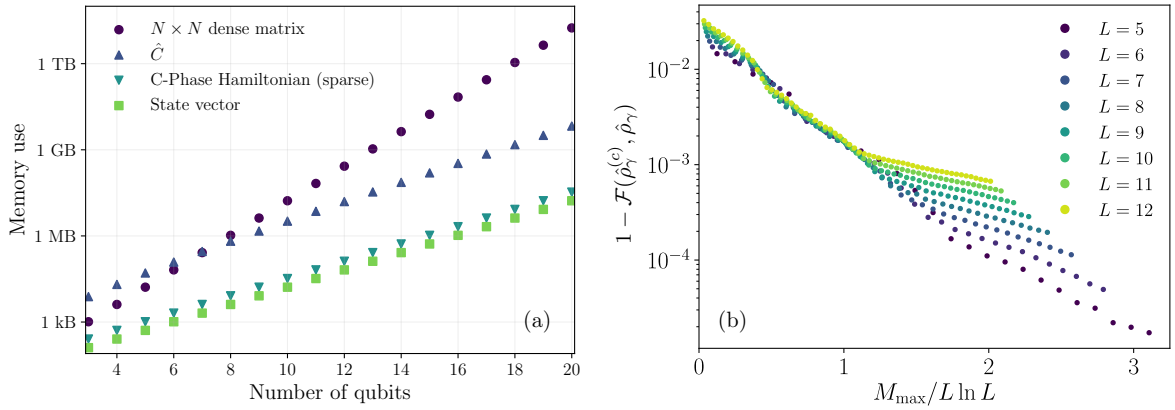

Figure 3.2: (a) Memory use of different quantities. One can see that the gain in memory requirement from our method scales exponentially. Note tha the Hamiltonian, being very sparse, is not the limitation. The largest stored object is the corner $\hat{C}$, that is composed of M states. Here, as an example, we considered $M = 100$. (b) Infidelity between the dissipative output density matrix $\hat{\rho}_\gamma^{(c)}$ (calculated via the corner-space method) and the exact integration density matrix $\hat{\rho}_\gamma$ as a function of the maximum corner dimension $M_{\max}$, for an L-qubit circuit and $\gamma T_{\mathrm{QFT}} = 1.42 \times 10^{-2}$. The initial state is $|\psi_0\rangle = \mathrm{QFT}^{-1}|\mathrm{GHZ}\rangle$.

is $\hat{K}_0$, which roughly corresponds to $\hat{H}$. The RAM needed to store it is shown in Fig. 3.2 together with that associated to other numerically relevant objects as a function of the number of qubits. Another feature of our implementation is that the dimension of the corner basis $M(t)$ is dynamically adapted to match the maximum allowed error that we impose, thus optimizing the computing and memory resources. This is done by means of memory-contiguous dynamically resizable arrays. This saves considerable computation time, as $M(t)$ grows in time when starting from a pure state ($M(0) = 1$).

III.3 Efficient evaluation of relevant observables

The evaluation of relevant metrics in quantum information is challenging with quantum-trajectory approaches, such as the Monte Carlo wave function method [51, 53, 158] (MCWF). Such an approach relies on the evolution of n stochastic trajectories $\{|\psi_i(t)\rangle\}_{i=1}^n$. The density matrix can then be reconstructed as $\hat{\rho}(t) = (1/n)\sum_{i=1}^n |\psi_i(t)\rangle\langle\psi_i(t)|$. To simulate a quantum computation, this has two downsides: firstly, for weakly dissipative systems ($\gamma T \ll 1$), most trajectories will experience no quantum jump and thus be identical. This means that most of the computing resources are wasted in performing a redundant task. Secondly, most relevant metrics involved in quantum-information processing, such as fidelity and entanglement measures, namely concurrence, negativity, or entanglement entropy [119], require constructing explicitly the (dense) density matrix of the system and diagonalizing it. In practice, the latter, of complexity $O(N^3)$, is not feasible for systems larger than $\sim$ 15 qubits. In contrast, the presented method yields explicitly both the eigenvalues $\{p_k(t)\}_{k=1}^M$ and the eigenvectors $\{|\phi_k(t)\rangle\}_{k=1}^M$ at every time step, with no need for additional calculations.

To give a concrete example, let us consider the evaluation of the fidelity between two arbitrary mixed states $\hat{\rho}$ and $\hat{\rho}'$ with rank M and M' respectively, as given by

$$\begin{aligned}\mathcal{F}(\hat{\rho},\hat{\rho}') &= \mathrm{Tr}[\sqrt{\sqrt{\hat{\rho}}\hat{\rho}'\sqrt{\hat{\rho}}}] \\ &= \sum_{m=1}^{M} \Big\langle \phi_m \Big| \Big\{ \textstyle\sum_{k,m=1}^{M} |\phi_k\rangle \, \mathcal{A}_{km} \, \langle\phi_m| \Big\}^{1/2} \Big| \phi_m \Big\rangle,\end{aligned}$$

with

$$\mathcal{A}_{km} = \sum_{k'=1}^{M'} p'_{k'} \sqrt{p_k p_m} \, \langle \phi_k | \phi'_{k'} \rangle \, \langle \phi'_{k'} | \phi_m \rangle \,, \tag{3.7}$$

where $p_k^{(\prime)}$ and $\phi_k^{(\prime)}$ correspond to the kth eigenvalue and eigenvector of $\hat{\rho}^{(\prime)}$. It appears from the above that the total complexity of this evaluation is of order

$$O(2N^3) + O(2M'N) + O(M^3) + O(MN) \,, \tag{3.8}$$

where the first term accounts for the eigendecomposition of the two density matrices, the second for the construction of the matrix $\mathcal{A}_{km}$, the third for its diagonalization and the last for the trace. An additional subleading complexity of order $O(n \times N)$ is to be added if the density matrices are obtained via Monte Carlo wave function calculations in order to account for their construction. In contrast, the order $O(2N^3)$ is to be discarded when using the dynamical corner-space method, as the eigendecompositions are known explicitly. Then, for each method, one finally has to leading order and for $M^{(\prime)} \ll N$ the following scaling figures of merit: In practice, the inconvenient scaling of the complexity for the

	MCWF	Dynamical corner-space
$\mathcal{F}(\hat{\rho},\hat{\rho}')$	$O(2N^3)$	$O(\max(M,2M')N)$
$\mathcal{F}(\hat{\rho},\lvert\phi\rangle)$	$O(nN)$	$O(MN)$
$S(\hat{\rho})$	$O(N^3)$	$O(M)$

Monte Carlo wave function approach, stemming from the two density-matrix diagonalizations, combined with the necessity of storing dense matrices well beyond the realistically available RAM makes it impossible to compute the fidelity between two mixed states from trajectories for systems larger than ~ 13 sites on a desktop computer. A similar discussion can be made for the entropy. We here see the superiority of the dynamical corner-space method, in particular in the context of quantum information where such metrics are often calculated.

III.4 Scaling of the corner dimension

In our method, the tolerance on the precision of the density matrix is tunable by design and the convergence versus M is ensured for moderately dissipative systems with low entropy. This is contrast to the case of quantum trajectories for which the number of needed trajectories n for a given problem is currently unknown *a priori* [122]. In Fig. 3.2(b), one can see that for $\gamma T_{\mathrm{QFT}} = 1.42 \times 10^{-2}$, an infidelity below 10^{-3} can be reached by choosing

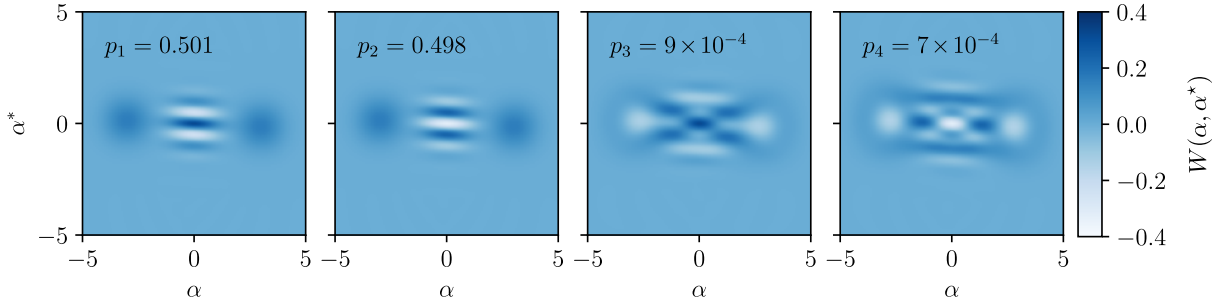

Figure 3.3: Wigner function of the 4 leading components of the corner-space describing a cavity subject to a Lindblad evolution as described in section III.5, after time $\gamma t = 10$. The first two correspond to states $|C_\alpha^\pm\rangle$ with opposite parities. Parameters here are $\kappa/\gamma = 2$, $\omega_c/\gamma = 1$, $K/\gamma = 10$, $G/\gamma = 50$.

$M_{\max} \sim L \ln L$. Hence, the corner dimension grows only polynomially with the size L, in particular sub-quadratically. Note that this value of γT_{QFT} is slightly higher than the state-of-the-art, hence the method is well suited to simulate experimental platforms in the coming years.

III.5 Integration method and stiffness

An issue with the map $\hat{C}(t) \mapsto \hat{C}(t + \delta t)$ that updates the corner basis at each time step is that it involves terms proportional to $\sqrt{\delta t}$ in the Kraus operators. This seems to restrict the method to a first order explicit integration scheme, which could result in a poor stability of the method when dealing with stiff dynamics ensued from nonlinear Hamiltonians. However, we circumvent this limitation by separating the pseudo-unitary evolution generated by $\hat{K}_0$ from that of the rest of the Kraus operators. Instead of computing $T_{im} = \sum_k K_{0,ik} C_{km} = \sum_k (\delta_{ik} - i\delta t \tilde{H}_{ik}) C_{km}$, we perform an exact numerical integration of the differential equation $\partial_t \hat{C} = -i\tilde{H}\hat{C}$ over the time interval $[t, t + \delta t]$ via an ordinary differential equation (ODE) solver well-adapted to the level of stiffness induced by the Hamiltonian. This allows us to treat stiff problems, via adapted implicit methods, and to use adaptive time-stepping.

To illustrate this point, let us consider a Kerr-cat qubit [159] as described by the following Hamiltonian:

$$\hat{H} = K\hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a} + \omega_c\hat{a}^\dagger\hat{a} + G(\hat{a}^2 + \hat{a}^{\dagger 2})\,, \tag{3.9}$$

where $\hat{a}$ ($\hat{a}^\dagger$) is a cavity mode annihilation (creation) operator, ω_c is the frequency of the cavity, K its Kerr nonlinearity, and G the two-photon driving frequency. If one considers jump operators $\hat{J}_1 = \sqrt{\gamma}\hat{a}$, $\hat{J}_2 = \sqrt{\kappa}\hat{a}^2$ that respectively describe single- and two-photon losses, a logical qubit can be conceived by considering logical states $|+\rangle = |C_\alpha^+\rangle = (|\alpha\rangle + \mathrm{i}\,|-\alpha\rangle)/\sqrt{2}$ and $|-\rangle = |C_\alpha^-\rangle = (|\alpha\rangle - \mathrm{i}\,|-\alpha\rangle)/\sqrt{2}$ as these are steady states of the system in the $\gamma \to 0$ limit. Numerically, the simulation of such systems cannot be efficiently treated with explicit first-order methods as the differential equation corresponding to their time evolution is stiff because of the Kerr nonlinearity. Thanks to

the numerical integrator used for the coherent part of the evolution of the corner $\hat{C}$, our method is capable of describing such systems, whose entropy is limited when $\kappa \gg \gamma$. In Fig. 3.3, the Wigner functions of each of the 4 most populated states of the corner are shown with their associated probabilities p_i, after having evolved the system for a time $t = 10/\gamma$ via the corner-space method, setting a photon cutoff $n_{\mathrm{ph}} = 20$. One sees that the low-dimensional basis found by the corner indeed closely matches that of a qubit with Schrödinger-cat logical states while keeping track of the effects of the single-body loss on the lowly probable states. By tuning the tolerance parameter of the method ϵ, such dissipative effects can be captured to any desired order. Our method therefore enables one to investigate the dynamics of such systems in an efficient way, and could be used to understand how single-photon dissipation impacts quantum operations in multi-Kerr-cat-qubit systems, among other applications involving bosonic qubits.

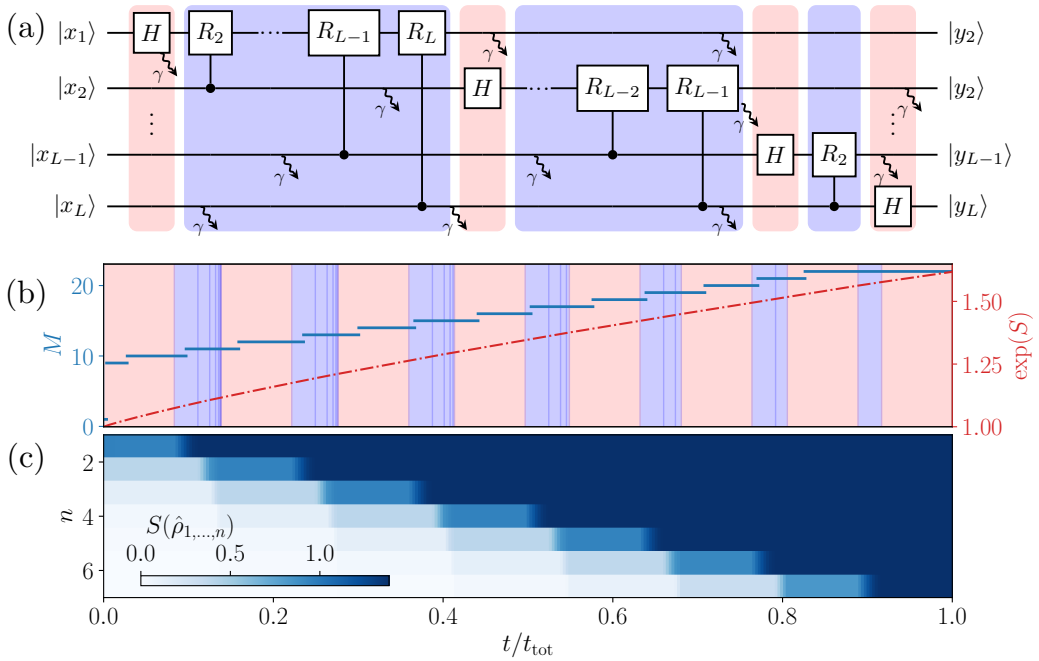

Figure 3.4: (a) Quantum circuit representing the quantum Fourier transform (QFT) in the presence of dissipation. (b) Continuous-time evolution of the exponential of the von Neumann entropy S (dash-dotted) with the input state $|\psi_0\rangle$ defined in Eq. (3.17), for $L = 8$ qubits and $\gamma/\delta = 1 \times 10^{-3}$. The corner-space dimension $M(t)$ is also plotted (plain line). Temporal intervals corresponding to Hadamard and controlled-phase gates are indicated by lighter (red) and darker (blue) background colors, respectively. (c) Temporal build-up of the entanglement entropy $\mathcal{S}_{\mathrm{ent}}(n)$, as defined in Eq. (3.19), as the different gates of the QFT are performed.

IV Application to the noisy QFT

IV.1 The quantum Fourier transform

We will now apply the method to the numerical simulation of noisy quantum circuits. A quantum circuit consists in a set of gates $\mathcal{G}$, that correspond to unitary operators $\hat{U}_g$, with $g \in \mathcal{G}$. These operators are applied to a register of L qubits that have as input state $|\boldsymbol{x}\rangle$, that are equivalent to the random-access memory of a classical computer. The output of a quantum circuit is given by state $|\boldsymbol{y}\rangle = f(|\boldsymbol{x}\rangle)$, where f denotes any recursive function [160], and $|\boldsymbol{y}\rangle$ corresponds to the answer we want from the algorithm.
We consider the noisy quantum Fourier transform (QFT) circuit, that is the cornerstone of the Shor algorithm and many algorithms related to the hidden subgroup problem [28]. In doing so, we will consider the architecture depicted in Fig. 3.4 (a). The input state to be Fourier transformed through the circuit is encoded in the state of L qubits. Given an orthonormal computing basis $\{|n\rangle\}_{n=0}^{N}$ spanning the Hilbert space of the system, of dimension $N = 2^L$, this circuit linearly transforms any input of the form

$$|\boldsymbol{x}\rangle = \sum_{j=0}^{N-1} x_j \, |j\rangle \,, \tag{3.10}$$

where $\boldsymbol{x}$ is a vector of complex amplitudes, into

$$|\boldsymbol{y}\rangle = \mathrm{QFT}(|\boldsymbol{x}\rangle) = \sum_{k=0}^{N-1} y_k \, |k\rangle \,, \quad \text{with} \quad y_k = \frac{1}{\sqrt{N}} \sum_j e^{2i\pi jk/N} x_j. \tag{3.11}$$

This therefore corresponds to the discrete Fourier transform of the N wave-function amplitudes. This operation can be reduced to subsequent applications of only two different types of quantum gates [28], the Hadamard and the controlled-phase gates. In the local basis:

$$\hat{U}_H|_i = \frac{1}{\sqrt{2}} \begin{pmatrix} \langle 0_i| & \langle 1_i| \\ 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{matrix} \\ |0_i\rangle \\ |1_i\rangle \end{matrix}, \qquad \hat{U}_R|_{jk} = \begin{pmatrix} \langle 0_j 0_k| & \langle 0_j 1_k| & \langle 1_j 0_k| & \langle 1_j 1_k| \\ 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & e^{2\pi i/2^k} \end{pmatrix} \begin{matrix} |0_j 0_k\rangle \\ |0_j 1_k\rangle \\ |1_j 0_k\rangle \\ |1_j 1_k\rangle \end{matrix}. \tag{3.12}$$

The number of involved such elementary gates, $L(L+1)/2$, is only polynomial in $L = \log_2(N)$; the complexity of the QFT algorithm is hence of order $O(L^2)$. In contrast, that of the usual fast Fourier transform (FFT) is exponential: $O(N \log N) \sim O(L2^L)$. This quantum advantage makes the QFT a central building block of many other algorithms.

In the following, the quantum gates introduced above will be executed via a continuous-time evolution defined by an appropriate master equation taking the form of Eq. (3.1). The effect of two types of noise will be addressed. We will first consider dissipative processes induced by a weak coupling to a zero-temperature environment, as described by jump operators of the form:

$$\hat{J}_i = \sqrt{\gamma}\hat{\sigma}_i^- , \tag{3.13}$$

ubiquitous in superconducting platforms. This corresponds to local decay processes from the excited qubit state $|\uparrow\rangle_j \equiv |0\rangle$ to the lower energy qubit state $|\downarrow\rangle_j \equiv |1\rangle$. We will treat the effect of decoherence as well. This will be described by local pure-dephasing jump operators:

$$\hat{J}_i = \sqrt{\gamma}\hat{\sigma}_i^z, \tag{3.14}$$

whose microscopic origin can be related to different sources such as disorder and phase fluctuations in the drive [161]. We consider a drive rotating at the qubit Rabi frequency. The Hadamard gate acting on the ith qubit can be realized via the subsequent application of the two following Hamiltonians:

$$\hat{H}_i^1 = \frac{\delta}{2}\hat{\sigma}_i^y, \qquad \hat{H}_i^2 = \frac{\delta}{2}\hat{\sigma}_i^z, \tag{3.15}$$

for a time $\Delta t_{H^1} = \pi/2\delta$ and $\Delta t_{H^2} = \pi/\delta$, respectively. The controlled-phase gates with control qubit j and target qubit k can be performed through the Hamiltonian:

$$\hat{H}_{j,k} = \frac{\delta}{2}\hat{\sigma}_j^z + \frac{\delta}{2}\hat{\sigma}_k^z - \frac{\delta}{2}\left(\hat{\sigma}_j^z\hat{\sigma}_k^z + \hat{\mathbb{1}}\right), \tag{3.16}$$

applied for a time $\Delta t_{R,k} = \pi/\delta 2^k$. For simplicity, we will assume sudden switching between gate Hamiltonians and will not include coherent errors (related to some switching time,or to some unwanted detuning between the drive and the qubit frequency), although both effects could be accurately described by the dynamical corner-space method.

A first example of the dynamics of such a circuit is presented in Fig. 3.4 (b) and Fig. 3.4 (c). There, the initial state is chosen as

$$|\psi_0\rangle = \mathrm{QFT}^{-1}(|\mathrm{GHZ}\rangle) = \frac{1}{\sqrt{2N}}\sum_{n=0}^{N-1}(1 + e^{2i\pi n/N})\,|n\rangle\,, \tag{3.17}$$

so as to be the inverse of the Greenberger–Horne–Zeilinger (GHZ) state [162–164]:

$$|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}\left(|00...0\rangle + |11...1\rangle\right). \tag{3.18}$$

The latter is known to be a maximally entangled state. Through this choice, the output state is ensured to be highly entangled, therefore demonstrating that entanglement is not a limiting factor for the corner-space method, as the intermediate states build up long-range entanglement. In panel (b) of Fig. 3.4, the time evolution of the corner-space dimension M as well as the exponential $\exp(S)$ of the von Neumann entropy $S(\hat{\rho}) = \mathrm{Tr}[\hat{\rho}\ln\hat{\rho}]$ are shown, exhibiting similar trends. Panel (c) depicts the spatial entanglement propagation as the circuit's gates are progressively applied from the first to the last qubit. This is quantified by the entanglement entropy

$$\mathcal{S}_{\mathrm{ent}}(n) \equiv S(\mathrm{Tr}_{1,\ldots,n}[\hat{\rho}]), \tag{3.19}$$

that evaluates entanglement between bipartitions of the form $\{\{1,\ldots,n\},\{n+1,\ldots,L\}\}$. While this is a rigorous measure of entanglement only for pure states [165], it still gives a valid qualitative description of the entanglement temporal build-up for states that are

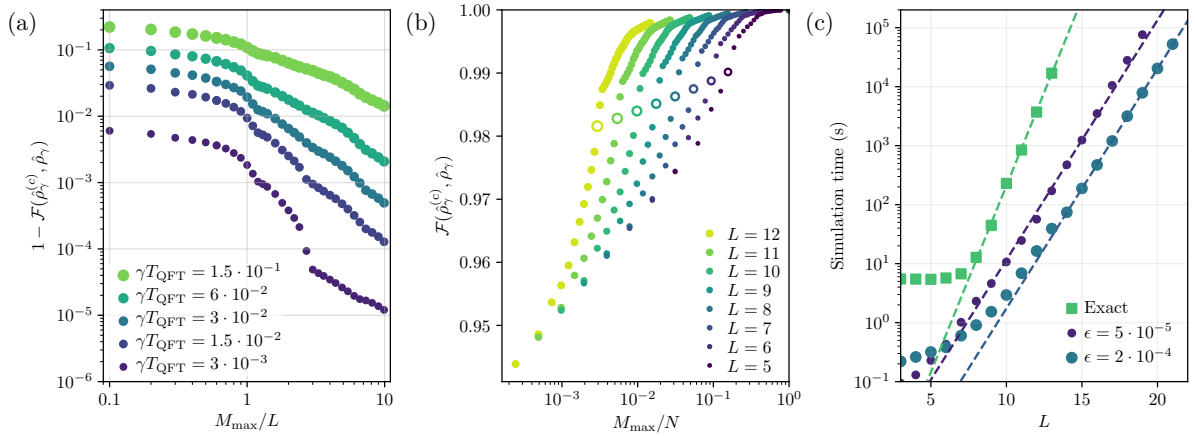

Figure 3.5: (a) Infidelity between the output density matrix $\hat{\rho}_\gamma^{(c)}$ calculated with the corner-space method and the output $\hat{\rho}_\gamma$ obtained via an exact integration as a function of the maximum corner dimension $M_{\max}$ for different values of γT_{QFT} and a fixed number of qubits $L = 10$. (b) Fidelity as a function of $M_{\max}$ and L for $\gamma T_{\mathrm{QFT}} = 2.5 \times 10^{-2}$. The parameter N denotes the dimension of the total Hilbert space. Values corresponding to $M_{\max} = L$ are highlighted by hollow markers (we find a fidelity $\mathcal{F} \gtrsim 0.997$ for $M_{\max} \sim L \ln L$). The initial state is $|\psi_0\rangle$ as in Fig. 3.1 for all three panels. (c) Simulation time of the QFT algorithm in the presence of dissipation versus the number L of qubits for the exact solution of the master equation (squared markers) and for the time-dependent corner-space method (circles) for two different values of the control parameter ϵ. The dissipation rate is set by $\gamma T_{\mathrm{QFT}} = 2.5 \times 10^{-2}$.

close to pure. Note that $\mathcal{S}_{\mathrm{ent}}(n) > S$ is a sufficient condition[2] for showing quantum entanglement between the two subsystems $1 \oplus \ldots \oplus n$ and $n+1 \oplus \ldots \oplus L$. One sees that the initial localized entanglement spreads through the system in close relation to the architecture of the circuit. As clearly follows from the shading in panel (b), this takes place when the (entangling) controlled-phase gates are applied.

IV.2 Benchmarking the method

Let us benchmark the accuracy of our calculations against the results of an exact integration of the master equation for small values of L, the numbers of qubits. In what follows, $\hat{\rho}_\gamma$ and $\hat{\rho}_\gamma^{(c)}$ denote the output density matrices of the noisy QFT obtained via the exact integration and via the corner method, respectively. Instead, $\hat{\rho}_0$ denotes the ideal outcome of the noiseless QFT, which is a pure state. The results of this benchmarking process are presented in Fig. 3.5 for fixed values of γT_{QFT}, where T_{QFT} denotes the physical duration of the QFT operation. This ensures that the output infidelity with respect to $\hat{\rho}_0$ remains constant as the circuit size is increased. In particular, Fig. 3.5 (a) shows the infidelity of the method $1 - \mathcal{F}(\hat{\rho}_\gamma^{(c)}, \hat{\rho}_\gamma)$ as a function of the rescaled maximum corner dimension $M_{\max}/L$ for $L = 10$. One sees that for $M_{\max} \sim 10L$ the exact results are excellently approximated by the time-dependent corner-space method for all the considered

[2]This corresponds to the violation of the classical identity $S_A \leq S_{A\oplus B}$.

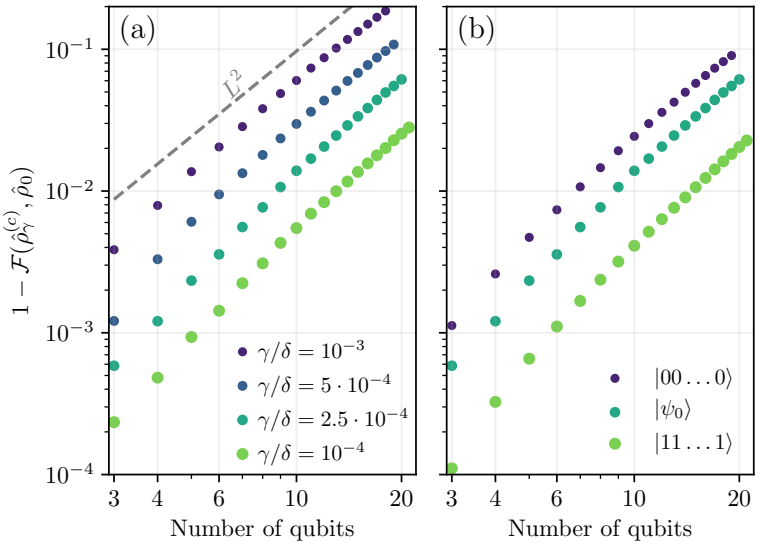

Figure 3.6: (a) Infidelity between the noisy output density matrix $\rho_\gamma^{(c)}$ and the noiseless output $\hat{\rho}_0 = |\mathrm{GHZ}\rangle\langle\mathrm{GHZ}|$ obtained from the initial state $|\psi_0\rangle$ for different values of γ/δ. The dashed line is a guide to the eye showing a growth $\propto L^2$. (b) Infidelity for three different initial states, described in the legend, for the same dissipation rate $\gamma/\delta = 2.5 \times 10^{-4}$.

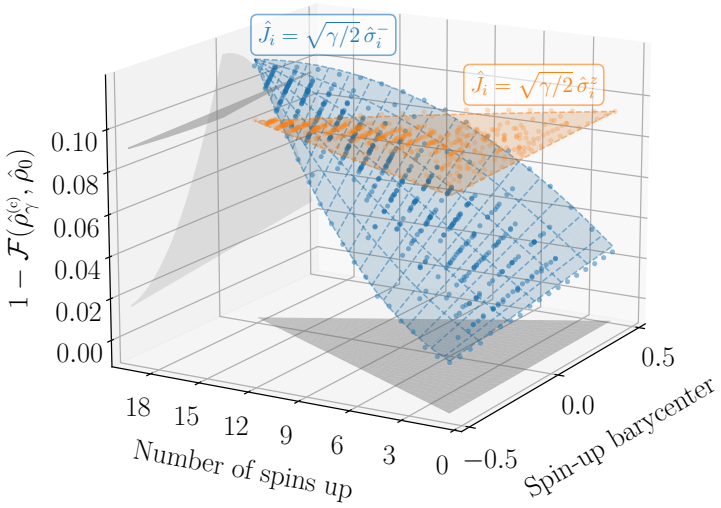

Figure 3.7: Infidelity for $L = 18$ qubits with dissipation (dark blue) or pure dephasing (orange) for 517 initial states per incoherent process, randomly sampled from the canonical basis. The infidelity is plotted versus the total number n_S of spins up in a given state and the spin-up barycenter defined in Eq. 3.20. The infidelity in both cases can be fitted by $\mathcal{I}(n_S, B) = a(n_S - n_{S_0})(B - B_0) + \mathcal{I}_0$. The surface for pure dephasing shows a negligible dependence on the initial state. Parameters: $\gamma/\delta = 3.7 \times 10^{-4}$ ($\gamma T_{\mathrm{QFT}} = 2.5 \times 10^{-2}$) and $\epsilon = 1 \times 10^{-4}$.

values of γT_{QFT}. The method still performs reasonably well for noise rates as high as $\gamma T_{\mathrm{QFT}} = 1.5 \times 10^{-1}$, where the fidelity to the output of the noiseless circuit is as low as $\mathcal{F}(\hat{\rho}_0, \hat{\rho}_\gamma) = 0.758$. Fig. 3.5 (b) shows the fidelity of the method $\mathcal{F}(\hat{\rho}_\gamma^{(c)}, \hat{\rho}_\gamma)$ for different numbers of qubits L as a function of the fraction $M_{\max}/N$ between the dimension of the basis spanning the corner and that of the basis of the entire Hilbert space. These results show that the advantage of the dynamical corner-space method over exact integration of the master equation increases with L. In particular, to match a desired fidelity, the required corner dimension M is found to grow as $L \ln L$ with the system size. For $L = 12$, an excellent agreement of the corner-space method with the exact integration is already obtained for $M_{\max}/N = 10^{-2}$. Finally, in Fig. 3.5 (c), we compare the computation time of the corner-space method to the exact integration, for two different values of ϵ[3]. The corner-space method achieves an exponential speed-up with respect to the master equation integration. This leads to simulations faster by more than three orders of magnitude for $L \sim 15$. Moreover, tuning the tolerance ϵ from 1×10^{-4} down to 5×10^{-5} preserves the scaling of the simulation time with L. Simulations of up to 21 qubits are presented. This represents a Hilbert-space dimension ($N = 2^L$) above 2 million states, corresponding to (dense) density matrices weighting 64 TiB that could never have been handled with a brute-force integration of the master equation (3.1). The method presented above is thus capable of efficiently evolving such high-dimensional objects by means of parsimonious corner representations of the density matrix. The calculations were performed on a single six-core Intel Xeon processor at a clock speed of 1.9GHz.

IV.3 Scaling laws

One can now evaluate the impact of incoherent processes on intermediate-scale devices via a continuous-time description and determine the scaling of errors. In Fig. 3.6, the fidelity $\mathcal{F}(\hat{\rho}_\gamma^{(c)}, \hat{\rho}_0)$ is shown for up to $L = 21$ qubits, for different values of γ/δ. Here, we consider dissipation channels described by the jump operators $\hat{J}_i = \sqrt{\gamma}\hat{\sigma}_i^-$. Remarkably, the infidelity scales only quadratically as a function of the number of qubits L. This scaling dependence allows one to precisely estimate the impact of $\gamma = 1/T_1$ on the QFT algorithm, T_1 being the energy relaxation time of the considered system.

IV.4 Impact of initial states

Another key property is the dependence of the fidelity on the initial state, crucial to redesign algorithms that rely preferentially on a certain class of states. In Fig. 3.7, we address this question for the QFT by sampling initial states. Either energy relaxation produced by the jump operators $\hat{J}_i = \sqrt{\gamma}\hat{\sigma}_i^-$ or pure dephasing described by $\hat{J}_i = \sqrt{\gamma}\hat{\sigma}_i^z$ are considered. Only two simple parameters that characterise the initial state are found to be crucial for the considered architecture: the total number n_S of spins up and the spin-up "barycenter"

$$B(\hat{\rho}) = \frac{1}{n_S} \sum_\ell \ell \times \mathrm{Tr}\Big[|{\uparrow}\rangle\langle{\uparrow}|_\ell \, \hat{\rho}\Big]. \tag{3.20}$$

[3]In evaluating these execution times, both the exact integration and corner-space calculations were carried out on a single six-core Intel Xeon E5-2609 v3 processor at 1.9 GHz.

Our findings show that, in presence of energy relaxation, the fidelity of the noisy QFT decreases linearly with the number of spins up in the initial state. This is in stark contrast to the case of pure dephasing, which shows no significant dependence on n_S. The fidelity also exhibits a strong dependence on the spin barycenter. Indeed, energy relaxation only affects excited states and the circuit's Hadamard gates are applied one qubit at a time starting from the beginning of the chain. As a result, excited qubit states (spin up) close to the end of the chain are rotated down to the Bloch-sphere equator by the Hadamard gates later than those on the opposite end. Thus, they are globally more affected by dissipation.

V Quantum errors

A fundamental question that quickly appeared in the debates on quantum computing in the early 90s is the question of how quantum errors occur. In classical computing, this was a problem for a long time and it took some time for classical error correction to be developed, leading to fault-tolerant classical computers. Quantum error correction is trickier, as for instance quantum information cannot simply be copied because of the no-cloning theorem, so all repetition codes developed for classical computing cannot be reused. Quantum error correction codes are based on digital error models, that are a simplification of how quantum errors microscopically occur. Here we will examine the differences between such a model and a full master equation approach.
As presented in section IV, a convenient and widely used model for quantum computation is the gate-based model [28]. For closed systems, this model is strictly equivalent to the successive application of unitary time-evolution operators of the form $\hat{U}_G = e^{-i\hat{H}_G\tau}$, with $\hat{H}_G$ the gate Hamiltonian and τ the gate time. In most current classical simulations of noisy quantum processors, errors are accounted for by extending the gate-based model to what has been coined digital error models [121]. These consist in applying noise gates after each unitary gate, expressed as Kraus operators, in analogy to error models for classical processors. However, in general, a quantum system is subject to a completely-positive, trace-preserving (CPTP) map acting on the system density matrix $\hat{\rho}$ in continuous time. The generator $\mathcal{L}$ of such a map can always be expressed in the Lindblad form as a Liouvillian whose action on the density matrix takes the form $\mathcal{L}[\hat{\rho}] = \mathcal{U}[\hat{\rho}] + \mathcal{D}[\hat{\rho}]$, with $\mathcal{U}$ and $\mathcal{D}$ the unitary and dissipative contributions to the time evolution of $\hat{\rho}$. Explicitly, for a quantum gate G, one has:

$$\mathcal{U}[\hat{\rho}] = -\mathrm{i}[\hat{H}_G, \hat{\rho}], \tag{3.21}$$

$$\mathcal{D}[\hat{\rho}] = \sum_{i=1}^{D} \Big(\hat{J}_i \hat{\rho} \hat{J}_i^\dagger - \frac{1}{2}\{\hat{J}_i^\dagger \hat{J}_i, \hat{\rho}\} \Big), \tag{3.22}$$

with $\hat{J}_i$ the jump operators that describe the dissipative channels, which take a simple form when the environment can be treated within the Born-Markov approximation. Given an initial density matrix $\hat{\rho}$, after a time interval τ, the density matrix at the output of the gate $\hat{\rho}(\tau)$ is given by

$$\hat{\rho}(\tau) = e^{\tau\mathcal{L}}\hat{\rho}. \tag{3.23}$$

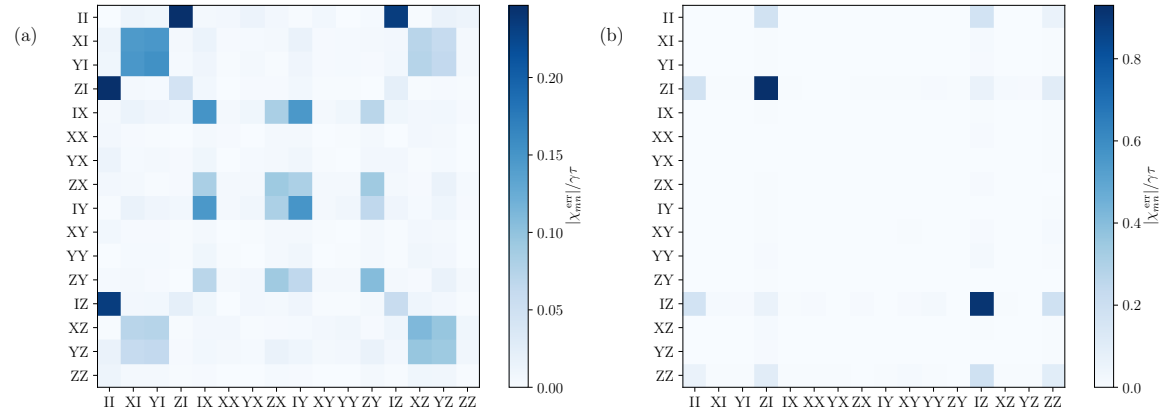

Figure 3.8: Absolute value of the elements of the error matrices χ^{err} (normalized by the characteristic error magnitude $\gamma\tau$) after the continuous-time evolution described in section V. In panel (a), the system is subject to dissipation ($\hat{J}_i = \sqrt{\gamma}\sigma_i^-$) and in panel (b) to decoherence channels ($\hat{J}_i = \sqrt{\gamma}\sigma_i^z$). Here $\gamma/\delta = 10^{-3}$. For each error matrix the element χ_{11} has been set to zero for readability (for such weak values of γ, $\chi_{11} \sim 1$).

One can always separate the unitary and dissipative parts of $\mathcal{L}$ to obtain

$$\hat{\rho}(\tau) = \mathcal{E}_G\left[\hat{U}_G(\tau)\hat{\rho}\hat{U}_G^\dagger(\tau)\right] \equiv \mathcal{E}_G(\hat{\rho}_{\hat{U}_G}), \tag{3.24}$$

where we have defined the density matrix after the ideal unitary process $\hat{\rho}_{\hat{U}_G} = \hat{U}_G(\tau)\hat{\rho}\hat{U}_G^\dagger(\tau)$, with $\hat{U}_G(\tau)$ the time evolution operator that corresponds to the application of Hamiltonian $\hat{H}_G$ for time τ, and $\mathcal{E}_G$ a non-unitary map. In cases where $\mathcal{D}$ and $\mathcal{U}$ commute, one explicitly has $\hat{U}_G(\tau) = e^{-i\hat{H}_G\tau}$ and $\mathcal{E}_G = e^{\tau\mathcal{D}}$, for the unitary and the error processes, respectively. However, in general, obtaining $\mathcal{E}_G$ is non-trivial, and experimentally one would need to perform a tomography on each gate to determine its exact error process: this is known as quantum process tomography [166].

Here we numerically simulate the quantum process tomography of a controlled phase gate of duration $\tau = \pi/2\delta$, δ being the Rabi frequency of the qubits. To do so, we decompose the error process $\mathcal{E}$ in the Pauli basis as:

$$\hat{\rho}(\tau) = \mathcal{E}(\hat{\rho}_{CR_{\pi/2}}) = \sum_{m,n} \chi^{\text{err}}_{mn}\hat{P}_m\hat{\rho}_{CR_{\pi/2}}\hat{P}_n^\dagger, \tag{3.25}$$

with $\hat{\rho}_{CR_{\pi/2}}$ the output density matrix corresponding to an ideal controlled phase gate, $\hat{P}_n \in \{I, X, Y, Z\}^{\otimes 2}$ the generators of the Pauli group on 2 qubits and χ^{err} the error matrix that completely characterizes the error process and that we aim at numerically determining.

In Fig. 3.8, the error matrix corresponding to a noisy $CR_{\pi/2}$ gate is shown for two different cases. Panel (b) corresponds to a Lindblad evolution with jump operators $\hat{J}_i = \sqrt{\gamma}\hat{\sigma}_i^-$, and panel (a) corresponds to a Lindblad evolution with jump operators $\hat{J}_i = \sqrt{\gamma}\hat{\sigma}_i^z$. By using a Choi decomposition [167], one can numerically estimate the error matrix that recovers the density matrix $\hat{\rho}(\tau)$ of the realistic Lindblad evolution. In the first case, $\hat{J}_i = \sqrt{\gamma}\hat{\sigma}_i^-$ was considered, which does not commute with the Hamiltonian. This leads

to a complex error process accounting for the spatial propagation of local errors upon (non-local) Hamiltonian time evolution. In this case, this manifests through the presence of single- and two-qubit error events with different magnitudes in the components of the error matrix. Similar error matrices have also been found experimentally [150]. Each non-negligible element corresponds to a Kraus operator to be applied after the ideal gate in the digital-error approach. While two local jump operators suffice to describe the qubit-environment interaction, up to 16 Kraus operators could be necessary in such an approach. Note that the noise process acting on two qubits at the same time generally has a smaller magnitude, which can justify that one can neglect them for low-depth circuits (the quantum Fourier transform has $L(L+1)/2$ gates, meaning these errors can accumulate in the long run). By contrast, in panel (b) pure-dephasing jump operators were considered, $\hat{J}_i = \sqrt{\gamma}\hat{\sigma}_i^z$, which commute with the controlled phase Hamiltonian. Hence, the error process only involves Z errors in this case, although not strictly local (notice the ZZ component).

As appears from the above discussion, modeling errors in continuous time is of fundamental and applied importance for the following reasons:

(1) The errors induced by the presence of local dissipative events, as captured by local jump operators, cannot be accounted for via local noise gates, and are affected by the applied Hamiltonian, an effect that is generally not described by digital error models. For d-qubit gates, the number of noise gates to be applied goes up to 4^d in a digital-error approach.

(2) For circuits such as the quantum Fourier transform, there are L controlled rotation gates that are applied for different times, hence one would need to perform tomography on each of them to recover their error matrices (as well as for the Hadamard gate). For multi-qubit gates such as the Toffoli gate, tomography becomes even more expensive.

(3) Characterizing the jump operators of an experimental platform is much easier than performing tomography to obtain the error process for each gate. For example, in the case of superconducting circuits, measuring the T_1 and T_2 relaxation times is enough.

(4) Our method is also able to treat collective dissipative processes described by jump operators such as $\hat{J}_z = (1/\sqrt{L})\sum_i \hat{\sigma}_i^z$ [168, 169]. This corresponds to a single Kraus operator in our method, and is hence inexpensive. With a digital error model, describing collective effects would require to perform tomography on a L-qubit system, which is intractable since it requires $O(4^L)$ measurements.

VI Conclusion

We investigated the role of dissipation and decoherence in noisy intermediate-scale quantum circuits. Focusing on the key algorithm of the QFT, we revealed the scaling behavior for the fidelity and explored its dependence on the initial state. To achieve this goal, we have

introduced and demonstrated a numerical time-dependent corner-space method that performs a judicious compression of the Hilbert space to faithfully represent the system density matrix. The method is not limited by entanglement and is suitable for systems with moderate entropy. Furthermore, our approach could be combined with efficient representations of the corner-space wavefunctions, such as neural-network ansätze [170–173]. These qualities make our approach ideally tailored for the NISQ era, where the noise processes will become weaker as progress is made. This provides a new tool to improve our understanding of quantum hardware. The presented method can indeed be applied in many contexts related to quantum information: algorithm design for quantum feedback [174], machine learning for quantum control [175] and quantum error mitigation [176].

4 Neural-network quantum states

In section 1, we have seen that there exists a plethora of methods to simulate quantum systems, both in closed and open settings. Each methods has its strengths and weaknesses; for instance mean-field methods are computationally extremely efficient while they poorly represent entanglement. Tensor networks slightly improve on this, being more resource-intensive but accurately representing area-law entangled states. Finally, corner methods have less limitations on entanglement, but deal poorly with high-entropy quantum systems. In this chapter, we present neural quantum states (NQS), a new class of methods that enable to treat systems in the strongly correlated regime, with a higher representative power than tensor networks. These methods are based on machine learning algorithms, more precisely on *generative models*.

I Introduction

Since LeCun's revolutionary paper on image classification in the 1990s [177], the broad field of machine learning has exploded. It can perhaps be considered as the latest technological revolution, nowadays permeating every field of science and having immense commercial impact. Its impact on physics has already been considerable [21]. In the braodest sense, machine learning consists in building algorithms that are not symbolically written: for instance, in the AlphaGo program [178], the instructions of how to play Go are not explicitly programmed, only a general purpose architecture is built-in, and it is only through some training procedure that the performance of the algorithm can become satisfying.

Generative models have recently been at the forefront of machine learning research. These models enable to reconsitute objects of a certain class from having observed a large number of instances of objects of this class. For instance, after observing a large number of human faces, new realistic faces can be constructed [179]. As such, one can claim the *function* corresponding to generating a realistic face will have been learned. The same idea has been applied to language processing, where models are now able to complete sentences in a high-fidelity and realistic fashion, thus being able to write, e.g., newspaper columns [180].

It is therefore a somewhat natural question to ask if such powerful models could help

for our understanding of complex many-body systems. Wavefunctions are indeed high-dimensional objects that must carry some structure, much like images and language. In this context, neural quantum states (NQS), a variant of generative models, that consist in finding faithful low-dimensional representations of many-body wavefunctions by a neural network were introduced [181]. Many important applications of NQS have been explored, including finding the ground state [182, 183], investigating the dynamics of strongly correlated systems [184, 185], quantum tomography [186], open quantum systems [187–190] and classical simulation of quantum circuits [191, 192]. In fact, NQS have proved to be the most accurate variational method in approximating the J_1–J_2 model's ground state in the frustrated regime [183]. In addition, several works have recently demonstrated the superior capacity of some neural-network architectures over tensor-network states in representing volume-law entangled states [193, 194] or area-law entangled states in 2D [195]. All these results indicate that NQS will become a standard tool to numerically study many-body systems in the coming years.
In this chapter, we introduce neural quantum states as well as applications to simulating dynamics of closed and open quantum systems. In section II, we introduce the task of generative modeling, and how its building blocks can be applied to learning quantum states. In section III, we present present Monte-Carlo Markov chains and their application to NQS. Then, we show how many-body wavefunctions are represented in practice in section IV, and present a few of the main architectures for NQS that were introduced in recent years. The question of their expressive power, or capacity, with respect to other approximation methods will be examined in section V.
We then turn to results obtained for the dynamics of many-body systems using NQS in section VI. We present some of the issues that can occur when trying to propagate an NQS in time, and provide an alternative solution for these problems. As such, we are able to obtain results for the Kibble-Zurek mechanism in large two-dimensional lattices, and can compare these results to state-of-the-art numerical methods.
Finally, in section VII, we present the application of NQS to open quantum systems. We explain why it is general trickier to construct an NQS for a density matrix, and present some of the solutions to these issues that have been developed recently. One of these solutions, that will work well for low-entropy systems, is to use a low-rank representation of the density matrix and to represent each state of its corner basis by an NQS.

II Generative models

Let us start by introducing the basic ingredients of generative modelling. A generative model is a statistical model $f_{\boldsymbol{\theta}}$, that depends on a set of variational parameters $\boldsymbol{\theta}$, of a function f that operates on elements $\boldsymbol{x}$ of a configuration space $\mathcal{X}$. For simplicity, we will consider a discrete configuration space in what follows, with N the total number of configurations. As such, the goal of generative modelling is to find the statistical model $f_{\boldsymbol{\theta}}$ that approximates the function f as best as possible. This means that at the end of the execution of some training algorithm

$$D\left(f_{\boldsymbol{\theta}}, f\right) = \epsilon, \tag{4.1}$$

with $\epsilon \to 0$ if the number of training steps goes to infinity with a carefully selected distance measure D, provided the generative model $f_{\boldsymbol{\theta}}$ has enough representative power[1]. For instance, a commonly used distance measure in machine learning is the mean square error [196]:

$$D\left(f_{\boldsymbol{\theta}}, f\right) = \sum_{\boldsymbol{x}_i}^{\boldsymbol{x}_N} \left(f_{\boldsymbol{\theta}}(\boldsymbol{x}_i) - f(\boldsymbol{x}_i)\right)^2 \tag{4.2}$$

which we see approaches 0 as $f_{\boldsymbol{\theta}} \to f$. To minimize this, one can use many different algorithms, that are all variants of gradient descent. Gradient descent consists in calculating the gradient of the distance, and updating the variational parameters so as to follow the steepest descent of the distance in the variational space. The update rule for the parameters $\boldsymbol{\theta}$ is given by:

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \eta \nabla_{\boldsymbol{\theta}} D(f, f_{\boldsymbol{\theta}}), \tag{4.3}$$

where η, known as the learning rate, is a hyperparameter that is adjusted for better convergence, and n denotes the nth iteration of the training procedure.
In practice, the number N of possible configurations may be exponentially large, hence the sum in Eq. (4.2) can become intractable. One must therefore approximate the distance D as well as its gradient. We can therefore identify two major requirements to perform generative modelling:

1. The evaluation of $f_{\boldsymbol{\theta}}(\boldsymbol{x})$ must be tractable.

2. The sums appearing in $D(f, f_{\boldsymbol{\theta}})$ and $\nabla_{\boldsymbol{\theta}} D(f, f_{\boldsymbol{\theta}})$ must be tractable or made tractable by approximation.

Before presenting how these requirements are fulfilled in practice, let us specialize the discussion to quantum states.
In the general case, one wishes to approximate a certain quantum state $|\psi\rangle$. The generative model, also known as *neural quantum state* (NQS) will therefore be denoted by $|\psi_{\boldsymbol{\theta}}\rangle$ in the rest of the chapter. Throughout the chapter we will restrict the discussion to spin systems, where spin configurations $\boldsymbol{s} = (\sigma_1, \sigma_2, \ldots, \sigma_L)$ are associated to states $|s\rangle$ that belong to a Hilbert space $\mathcal{H}$. $|\psi_{\boldsymbol{\theta}}\rangle$ and $|\psi\rangle$ can be expanded in the basis spanned by these states as:

$$|\psi\rangle = \sum_{\boldsymbol{s}} \psi(\boldsymbol{s}) \, |s\rangle \tag{4.4}$$

$$|\psi_{\boldsymbol{\theta}}\rangle = \sum_{\boldsymbol{s}} \psi_{\boldsymbol{\theta}}(\boldsymbol{s}) \, |s\rangle \, . \tag{4.5}$$

The functions ψ and $\psi_{\boldsymbol{\theta}}$ operate on vectors $\boldsymbol{s}$ and the resulting amplitudes $\psi(\boldsymbol{s})$ and $\psi_{\boldsymbol{\theta}}(\boldsymbol{s})$ are complex numbers. Thus the generative model is fully specified by its amplitudes $\psi_{\boldsymbol{\theta}}(\boldsymbol{s})$, that form a 2^L vector of complex numbers. Note that the requirement that $\psi_{\boldsymbol{\theta}}(\boldsymbol{s})$ be tractable means that the number of variational parameters $\boldsymbol{\theta}$ must be polynomial in the system size, or else computing amplitudes will include a super-polynomial number

[1] We will come back to question in section V.

of operations.

Originally, these ideas were applied to finding the ground state of a many-body Hamiltonian $\hat{H}$. In this case, the cost function to minimize is the energy, given by

$$E_{\boldsymbol{\theta}} = \frac{\langle \psi_{\boldsymbol{\theta}} | \hat{H} | \psi_{\boldsymbol{\theta}} \rangle}{\langle \psi_{\boldsymbol{\theta}} | \psi_{\boldsymbol{\theta}} \rangle}. \tag{4.6}$$

As the energy converges to its minimum, $|\psi_{\boldsymbol{\theta}}\rangle \to |\psi_0\rangle$, with the ground state defined by $\hat{H}\,|\psi_0\rangle = E_0\,|\psi_0\rangle$. By doing this one effectively *learns* the ground state. The expression for the expectation value of any operator $\hat{A}$ can be written in the following form, for a given NQS $|\psi_{\boldsymbol{\theta}}\rangle$:

$$\langle \hat{A} \rangle = \frac{\langle \psi_{\boldsymbol{\theta}} | \hat{A} | \psi_{\boldsymbol{\theta}} \rangle}{\langle \psi_{\boldsymbol{\theta}} | \psi_{\boldsymbol{\theta}} \rangle} \tag{4.7}$$

$$= \frac{\sum_{s,s'} \langle \psi_{\boldsymbol{\theta}} | s \rangle \langle s | \hat{A} | s' \rangle \langle s' | \psi_{\boldsymbol{\theta}} \rangle}{\sum_s |\langle \psi_{\boldsymbol{\theta}} | s \rangle|^2} \tag{4.8}$$

$$= \frac{\sum_s \langle \psi_{\boldsymbol{\theta}} | s \rangle \frac{\langle s | \psi_{\boldsymbol{\theta}} \rangle}{\langle s | \psi_{\boldsymbol{\theta}} \rangle} \sum_{s'} \langle s | \hat{A} | s' \rangle \langle s' | \psi_{\boldsymbol{\theta}} \rangle}{\sum_s |\langle \psi_{\boldsymbol{\theta}} | s \rangle|^2} \tag{4.9}$$

$$= \frac{\sum_s |\langle \psi_{\boldsymbol{\theta}} | s \rangle|^2 \sum_{s'} \langle s | \hat{A} | s' \rangle \frac{\langle s' | \psi_{\boldsymbol{\theta}} \rangle}{\langle s | \psi_{\boldsymbol{\theta}} \rangle}}{\sum_s |\langle \psi_{\boldsymbol{\theta}} | s \rangle|^2}. \tag{4.10}$$

Two main terms can be identified:

$$P(\boldsymbol{s}) = \frac{|\langle \psi_{\boldsymbol{\theta}} | s \rangle|^2}{\sum_s |\langle \psi_{\boldsymbol{\theta}} | s \rangle|^2} \tag{4.11}$$

$$A_{\mathrm{loc}}(\boldsymbol{s}) = \sum_{s'} \langle s | \hat{A} | s' \rangle \frac{\langle s' | \psi_{\boldsymbol{\theta}} \rangle}{\langle s | \psi_{\boldsymbol{\theta}} \rangle}. \tag{4.12}$$

where $A_{\mathrm{loc}}(\boldsymbol{s})$ is the so-called *local estimator* of $\hat{A}$. Therefore, the expectation value of an observable $\hat{A}$ can be written as the expectation value of its local estimator A_{loc} over the probability distribution $P(\boldsymbol{s})$:

$$\langle \hat{A} \rangle = \sum_s P(\boldsymbol{s}) A_{\mathrm{loc}}(\boldsymbol{s}) = \langle A_{\mathrm{loc}} \rangle_P \ . \tag{4.13}$$

Let us stress that these calculations are feasible for operators with the property that the number of states $\boldsymbol{s}'$ such that the elements $|\langle s | \hat{A} | s' \rangle| \neq 0$, for arbitrary $\boldsymbol{s}$ is at most polynomial in the number of spins. This is known as k-locality, with k the maximum number of sites a single term in the Hamiltonian acts on (for the transverse-field Ising model, $k = 2$). Conversely, evaluating $A_{\mathrm{loc}}(\boldsymbol{s})$ would not be tractable, given that the sum over s' in 4.12 would be over an exponential number of elements. By defining the logarithmic derivatives, $O_k(\boldsymbol{s}) = \partial_{\theta_k} \psi_{\boldsymbol{\theta}}(\boldsymbol{s}) / \psi_{\boldsymbol{\theta}}(\boldsymbol{s})$, the components of the gradient can be

calculated straightforwardly:

$$\begin{aligned}\partial_{\theta_k}\langle\hat{A}\rangle &= \partial_{\theta_k}\frac{\sum_{\boldsymbol{s},\boldsymbol{s'}}\psi^*_{\boldsymbol{\theta}}(\boldsymbol{s})\,\langle s|\hat{A}|s'\rangle\,\psi_{\boldsymbol{\theta}}(\boldsymbol{s'})}{\sum_{\boldsymbol{s}}|\psi_{\boldsymbol{\theta}}(\boldsymbol{s})|^2}\\ &= \frac{\sum_{\boldsymbol{s},\boldsymbol{s'}}\psi^*_{\boldsymbol{\theta}}(\boldsymbol{s})\,\langle s|\hat{A}|s'\rangle\,O_k(\boldsymbol{s'})\psi_{\boldsymbol{\theta}}(\boldsymbol{s'})}{\sum_{\boldsymbol{s}}|\psi_{\boldsymbol{\theta}}(\boldsymbol{s})|^2} + \frac{\sum_{\boldsymbol{s},\boldsymbol{s'}}\psi^*_{\boldsymbol{\theta}}(\boldsymbol{s})O^*_k(\boldsymbol{s})\,\langle s|\hat{A}|s'\rangle\,\psi_{\boldsymbol{\theta}}(\boldsymbol{s'})}{\sum_{\boldsymbol{s}}|\psi_{\boldsymbol{\theta}}(\boldsymbol{s})|^2}\\ &\qquad - \langle\hat{A}\rangle\frac{\sum_{\boldsymbol{s}}|\psi_{\boldsymbol{\theta}}(\boldsymbol{s})|^2(O_k(\boldsymbol{s})+O^*_k(\boldsymbol{s}))}{\sum_{\boldsymbol{s}}|\psi_{\boldsymbol{\theta}}(\boldsymbol{s})|^2},\end{aligned}$$

which implies $\partial_{\theta_k}\langle\hat{A}\rangle = \langle A_{\mathrm{loc}}O^*_k\rangle - \langle A_{\mathrm{loc}}\rangle\langle O_k\rangle + \text{c.c.}$ (4.14)

III Monte-Carlo Markov Chains

To calculate these quantities in order to perform gradient descent, one needs to truncate the sum over $\boldsymbol{s}$ in Eq. (4.13) for it to become tractable. A way to do this is to obtain samples according to P, so as to have a correct estimation of the averages in the expression of the gradient. This will lead to averages given by

$$\langle\hat{A}\rangle \approx \frac{1}{M}\sum_{i=1}^{M}A_{\mathrm{loc}}(\boldsymbol{s}^{(i)}), \tag{4.15}$$

where $\boldsymbol{s}^{(i)}$ are samples drawn according to P. The statistical error associated with such an estimate scales as $\varepsilon \propto \sqrt{\sigma^2/M}$, and it is bounded as long as the variance σ^2 of A_{loc} is finite[2]. For instance, when $\hat{A}$ is a k-local spin operator with bounded coefficients, its variance is strictly finite.[3] Therefore, the error in the estimate of expectation values decreases as $\varepsilon \sim 1/\sqrt{M}$, which allows us to reach arbitrary accuracy in the estimation by increasing the number of samples M, given that $\lim_{M\to\infty}\varepsilon = 0$. However, generating a set of samples according to the Born distribution, $\{\boldsymbol{s}^{(i)}\} \sim P(\boldsymbol{s})$, is in general a non-trivial computational task in the case where the variational ansatz, $\Psi_{\boldsymbol{\theta}}(\boldsymbol{s})$, is parameterized by an efficiently computable, yet arbitrary function. One of the most commonly adopted strategies to sample from $P(\boldsymbol{s})$ is through *Markov chain Monte Carlo* [197] methods, including the Metropolis-Hastings method [198], which generate a sequence of correctly distributed samples $\boldsymbol{s}^{(i)}$.

Metropolis-Hastings methods construct a markovian stochastic process which satisfies the *detailed balance* condition for the target probability distribution

$$P(\boldsymbol{s})\mathcal{T}(\boldsymbol{s}\to\boldsymbol{s'}) = P(\boldsymbol{s'})\mathcal{T}(\boldsymbol{s'}\to\boldsymbol{s})\,, \tag{4.16}$$

where $\mathcal{T}(\boldsymbol{s}^{(i)}\to\boldsymbol{s}^{(i+1)})$ is the probability that the state $\boldsymbol{s}^{(i)}$ at step i transitions to the state $\boldsymbol{s}^{(i+1)}$ at the following step. As the process is Markovian, the transition probability at

[2] If A_{loc} is Gaussian, then the error is equal to $\sqrt{\sigma^2/M}$.

[3] When $|\psi_\theta\rangle$ approaches an eigenstate of $\hat{A}$, it can be shown that the variance vanishes. Consequently, considering $\hat{A} = \hat{H}$, the statistical error vanishes as we approach the ground (or any excited) state. This is known as the zero-variance property, and is convenient for ground state search as it means the variance of the error will vanish as the solution is approached, thus ensuring a well-behaved training procedure.

every step depends exclusively on the current configuration. Assuming that the process is ergodic, the detailed balance condition ensures that, regardless of the initial configuration $\boldsymbol{s}^{(0)}$, the sequence eventually converges to the correct distribution $P(\boldsymbol{s})$ in the long time limit.

One possible choice of the transition probability $\mathcal{T}$ is given by the *Metropolis-Hastings algorithm* [198]. The main idea is to express $\mathcal{T}$ in terms of a local transition kernel T and an acceptance probability A such that

$$\mathcal{T}(\boldsymbol{s} \to \boldsymbol{s}') = T(\boldsymbol{s} \to \boldsymbol{s}')A(\boldsymbol{s} \to \boldsymbol{s}')\,. \tag{4.17}$$

This way, we split the global stochastic process into the product of two local subprocesses that we can compute efficiently. For instance, it is very easy to find a normalized local transition kernel that allows us to modify only a few degrees of freedom, like flipping a single spin in a given configuration. Conversely, it is hard to find a normalized global kernel that would act on all spins.

The acceptance probability to go from a configuration $\boldsymbol{s}$ to $\boldsymbol{s}'$ through a local transition is defined as

$$A(\boldsymbol{s} \to \boldsymbol{s}') = \min\left(1, \frac{P(\boldsymbol{s}')T(\boldsymbol{s}' \to \boldsymbol{s})}{P(\boldsymbol{s})T(\boldsymbol{s} \to \boldsymbol{s}')}\right)\,. \tag{4.18}$$

Notice that the normalization of the Born probabilities cancel out, giving the tractable expression

$$\frac{P(\boldsymbol{s}')}{P(\boldsymbol{s})} = \left|\frac{\langle s'|\psi_{\boldsymbol{\theta}}\rangle}{\langle s|\psi_{\boldsymbol{\theta}}\rangle}\right|^2\,, \tag{4.19}$$

which allows us to consider unnormalized variational ansätze. Therefore if the variational state is computationally tractable, the transition probability is also tractable.

Choosing a valid transition rule $T(\boldsymbol{s} \to \boldsymbol{s}')$ is important, particularly in the case of systems whose states obey certain symmetries. For example, if the total magnetization along the direction of the computational basis is known, it may be fixed during the sampling procedure, and we can use a transition rule that does not project the Markov chain outside of a certain region in Hilbert space. For ground state search, a somewhat expensive yet effective choice for the transition kernel is to use the Hamiltonian itself:

$$T(\boldsymbol{s} \to \boldsymbol{s}') = \frac{\left|\langle s|\hat{H}|s'\rangle\right|(1-\delta_{\boldsymbol{s},\boldsymbol{s}'})}{\sum_{s'\neq s}\left|\langle s|\hat{H}|s'\rangle\right|}\,, \tag{4.20}$$

which is known as the Hamiltonian transition rule [181].

Finally, note that obtaining samples through Markov chains yields issues, as the samples are generally correlated. In addition, since the proposed moves are generally local, this can fail terribly for states that have multimodal distributions in Hilbert space. We will come back to these issues while discussing direct sampling in the next section.

IV Representing the wavefunction

Now that we have seen how to compute the quantities of interest using parametrized quantum states, let us dive into how to devise expressive variational states in practice. The

main idea is that we need to represent high-dimensional functions with a parametrization that is flexible and general enough to describe physical systems, while involving only a polynomial amount of parameters.

Traditionally, researchers have relied on physically-inspired variational ansätze. An extension of the mean-field ansätze presented in the introduction is the Jastrow wave function [199, 200], which stands out as one of the most successful and widely used ones for quantum chemistry. It is based on the assumption that two-body interactions are the most physically relevant, and it assigns a trainable potential to every interacting pair. Formally, the Jastrow wave function is given by:

$$\psi^J_{\boldsymbol{\theta}}(\boldsymbol{s}) = e^{-\frac{1}{2}\sum_{i\neq j}\theta_{ij}\sigma_i\sigma_j}\,, \tag{4.21}$$

where the sum runs over all possible spin pairs, and θ_{ij} are the parameters encoding pairwise spin correlations. Therefore, for a system of L spins, the resulting wave function has L^2 parameters. Moreover, in translationally invariant systems, the parameters θ_{ij} can be made depend exclusively on the distance between i and j, resulting in a reduced number of parameters L.

To design more general variational ansätze, one can use neural networks. In this case, $\psi_{\boldsymbol{\theta}}(\boldsymbol{s})$ corresponds to the output of a neural network, that takes the configuration $\boldsymbol{s}$ as input in the form of a vector. In a feed-forward neural network of depth D, every layer l consists of a nonlinear activation function $g^{(l)}$ that acts, component-wise, on a vector resulting from applying the weight matrix $\boldsymbol{W}^{(l)}$ to the output of the previous layer. This way, it is possible to write the variational state as the composition of operations $g^{(l)}\odot\boldsymbol{W}^{(l)}$, where $\odot$ indicates point-wise operation, such that

$$\psi_{\boldsymbol{\theta}}(\boldsymbol{s}) = g^{(D)}\odot\boldsymbol{W}^{(D)}\ldots g^{(2)}\odot\boldsymbol{W}^{(2)}g^{(1)}\odot\boldsymbol{W}^{(1)}\boldsymbol{s}\,. \tag{4.22}$$

Hence, the output is a complex scalar representing the probability amplitude of configuration $\boldsymbol{s}$.

From a mathematical perspective, these ansätze are of great interest given that neural networks are subject to universal representation theorems [201]. Consider a neural network without nonlinear activation functions. The function realized by such a network is a simple affine map that consists of multiplying the input by a weight matrix and adding to it an additional bias vector. This is evidently not enough to represent nonlinear functions. Thus, the addition of nonlinear activation functions is crucial for neural networks to be able to represent a larger class of functions. Kolmogorov and Arnold [201] have shown that any arbitrary continuous high-dimensional function can be expressed as a linear combination of the composition of a set of nonlinear functions

$$f(\boldsymbol{x}) = \sum_{i=0}^{2m}\zeta_i\left(\sum_{j=1}^{m}\varsigma_{i,j}(x_j)\right), \tag{4.23}$$

where $\zeta_i, \varsigma_{i,j}$ are nonlinear functions that act on the individual components of the input $\boldsymbol{x}\in\mathbb{R}^m$. This means that we could represent any function $f(\boldsymbol{x})$ with a polynomial number $O(m^2)$ of one-dimensional nonlinear functions. This strongly resembles the structure of a neural network with two hidden layers. Note, however, that the nonlinear functions

must be carefully chosen depending on the target function. In most machine learning architectures, the nonlinearities are typically fixed, i.e. $\zeta_i = \varsigma_{i,j}\ \forall i, j$. It turns out that fully-connected neural network composed of a single hidden layer and nonlinear activation functions are also *universal function approximators*. That is, given that the target function is reasonably well-behaved it can be approximated to any desired accuracy given that its hidden layer contains enough nodes [196, 202, 203]. Note that this may still require a hidden layer that is exponentially large in the number of nodes. This raises the question what one can achieve with neural networks that have multiple hidden layers.

The universal approximation theorem guarantees that there exists a neural network, i.e., choice of the architecture, which approximates the given target function arbitrarily well. However, it does not guarantee that we are able to find this choice. It turns out that, in practice, deep neural networks are capable of solving many problems with much less nodes, i.e., trainable parameters, compared to shallow neural networks. However, these results hold for arbitrary nonlinear functions, $\zeta_i, \varsigma_{i,j}$ in 4.23, that must be appropriately found in order to represent the target function. In practice, neural networks use a fixed nonlinear activation, and we can only adjust the number of operations. In these cases, the number of neurons does not have a strict polynomial scaling and it can be, in the worst case, exponential in L [202]. Nevertheless, the state-of-the-art results in computer vision and natural language processing [180, 204, 205] should be sufficient motivation to employ similar techniques to represent quantum states. In recent years, a large number of variational ansätze based on neural networks have also been developed for fermionic [206, 207] and bosonic systems [208, 209].

IV.1 Restricted Boltzmann machines

NQS were first introduced using restricted Boltzmann machines (RBMs) [181]. RBMs are shallow models featuring two fully-connected layers: a *visible* layer, consisting of L units, and a *hidden* layer, consisting of M units. A scheme of an RBM architecture is presented in 4.1. The wave function amplitudes of an RBM ansatz are given by:

$$\Psi_{\boldsymbol{\theta}}(\boldsymbol{s}) = \sum_{\boldsymbol{h}} e^{\boldsymbol{b}_v^\dagger \boldsymbol{s} + \boldsymbol{b}_h^\dagger \boldsymbol{h} + \boldsymbol{h}^\dagger \boldsymbol{W} \boldsymbol{s}}. \tag{4.24}$$

where $\boldsymbol{s}, \boldsymbol{h}$ represent the visible and hidden units, respectively, and the parameters $\boldsymbol{\theta} = \{\boldsymbol{b}_v, \boldsymbol{b}_h, \boldsymbol{W}\}$ represent the visible and hidden biases and the weight matrix, respectively. In the NN picture, the RBM is a single-layer nonlinear feed-forward NN, with the visible units serving as inputs and the exponential serving as the activation function.

By construction, RBMs are designed in such a way that computing the summation over hidden units, as in Eq. (4.24), can be done analytically. To see this, we can rewrite Eq. (4.24) in a tractable form considering binary hidden units $h_i \in \{-1, 1\}$, leading to

$$\psi_{\boldsymbol{\theta}}(\boldsymbol{s}) = e^{\boldsymbol{b}_v^\dagger \boldsymbol{s}} \prod_{i=1}^{M} 2\cosh\left(\boldsymbol{b}_{h,i} + \boldsymbol{W}_{i\cdot}\boldsymbol{s}\right), \tag{4.25}$$

where $\boldsymbol{b}_{h,i}$ and $\boldsymbol{W}_{i\cdot}$ denote the i-th hidden bias and weight matrix row, respectively. To treat spin systems, the visible units will represent the N physical spins, thus the input of

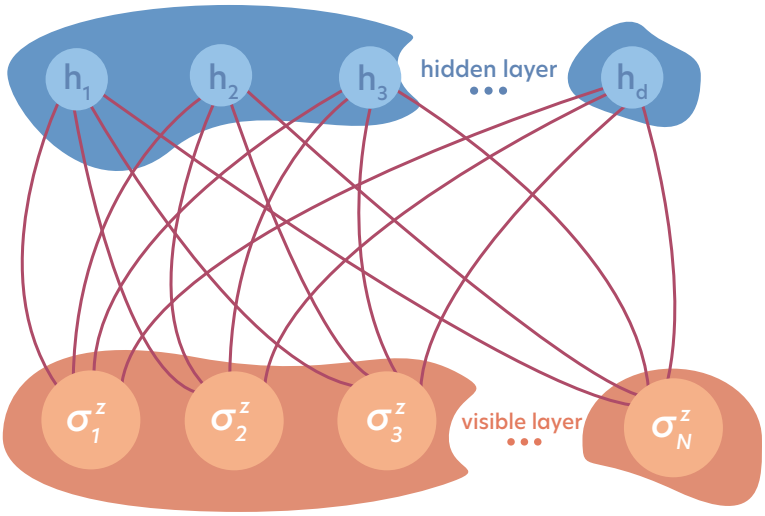

Figure 4.1: Pictorial representation of a RBM that represents the wave function of an L-spin system, with $\boldsymbol{s} = (\sigma_1, \sigma_2, \ldots, \sigma_L)$ and $\boldsymbol{h} = (h_1, h_2, \ldots, h_d)$ the hidden units. Adapted from [22].

the RBM is simply the spin configuration $\boldsymbol{s}$. In this way we obtain an analytical expression to evaluate the amplitude for a given spin configuration, and thus represent the full wave function with this ansatz. One can also interpret the hidden units as M hidden spins, and in this picture the RBM can be thought of as an interacting spin model with interaction strengths W_{ij}. Moreover, we can treat an RBM as a model with an associated energy depending on its parameters, input, and hidden spin values. This is known as an energy-based model, and explains why input biases are present in Eq. (4.24). In fact, the RBM is equivalent to a Hopfield network, a type of spin glass [210].

Being the first to be introduced in this context, most of the early works about NQS employ RBMs, but other architectures have been explored in more recent years, that we present in the next section. The capacity of RBMs and its relationship to quantum entanglement has been examined in various works [211, 212]. An extension of this architecture, the deep RBM, has also been introduced to solve more complex problems [213], which consists in stacking more than two fully connected layers.

IV.2 Autoregressive neural networks

Autoregressive neural networks can also be used for constructing NQS, as introduced in [182] and applied to both quantum [214] and classical problems [215]. Their main advantage is that the wave function outputs of such models are normalized, allowing for direct sampling, which is easier to parallelize than Markov chain Monte Carlo and leads to independent samples. A pictorial representation of both the network and the sampling algorithm is presented in 4.2. Here, the many-body wave function is expressed in terms

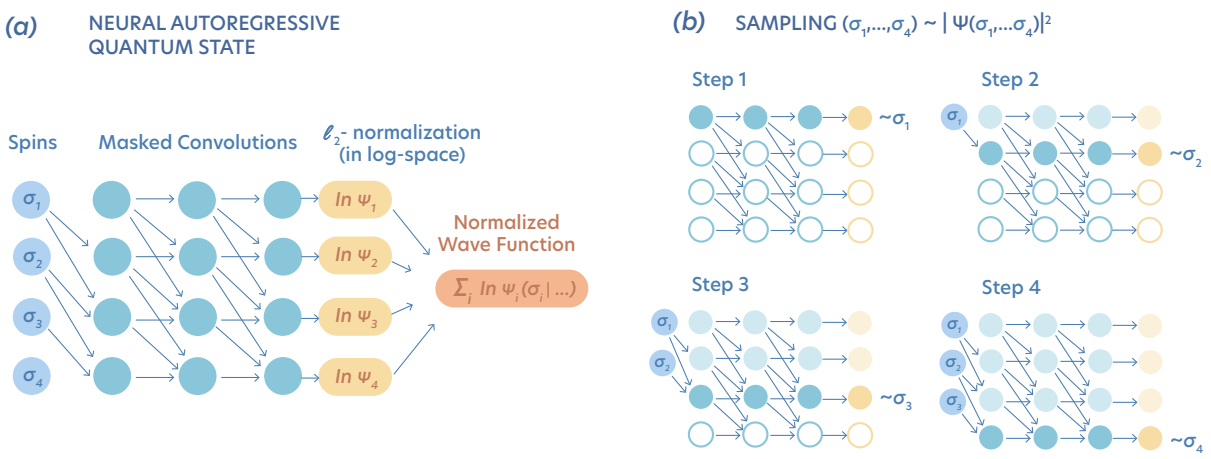

Figure 4.2: Example of an ARNN quantum state for four spins. (a) Pictorial representation of the network. The arrows representing the weights of the model are skewed in order not to break the conditional structure of the output probability distribution. These layers are "masked", due to some connections being deleted. (b) Sampling algorithm. One samples consecutive spins using direct sampling on the conditional probabilities at each step. From [22].

of a product of conditional complex amplitudes:

$$\psi_{\boldsymbol{\theta}}(\boldsymbol{s}) = \prod_{i=1}^{L} \phi_i(\sigma_i|\sigma_{i-1},\ldots,\sigma_1)\,, \tag{4.26}$$

which is subject to the normalization condition $\sum_\sigma \left|\phi_i(\sigma|\sigma_{i-1},\ldots,\sigma_1)\right|^2 = 1$. With this kind of architecture, we can compute expectation values by directly sampling state configurations instead of building a Markov Chain through the Metropolis-Hastings algorithm. This is done by iteratively sampling one spin after the other: we start sampling the first spin σ_1 from the reduced probability distribution $|\phi(\sigma_1)|^2$. Then, we sample the second one σ_2 according to the conditional probability distribution $|\phi_2(\sigma_2|\sigma_1)|^2$, then the next one $|\phi(\sigma_3|\sigma_2,\sigma_1)|^2$, and so on until σ_N. This sampling procedure is embarrassingly parallel.[4]

This sampling procedure yields independent identically distributed samples. Conversely, Markov chain Monte Carlo methods may suffer from highly correlated consecutive samples, which is problematic for complex probability distributions, e.g. that are far from Gaussian. Markov chain Monte Carlo methods such as the Metropolis-Hastings algorithm, generally rely on performing modifications to the spin configurations to sample subsequent states. Therefore, this process could yield highly correlated consecutive samples that may have a negative impact on the results. In order to compute expectation values, we need to estimate the autocorrelation time to draw uncorrelated samples from the resulting chain. Moreover, when approaching a phase transition, such methods suffer from critical slowing down, making the sampling of uncorrelated configurations unfeasible in many situations.

[4]We can use the intermediate conditional probabilities to draw samples for a low computational cost, e.g., use the probabilities for $N-1$ spins and sample from the last one, to obtain new samples; with Markov chain Monte Carlo we cannot do this.

Consider a quantum state that spans several separated regions in the Hilbert space, where the probability is concentrated. In this case, Markov chains generally remain stuck in one of the regions, given that it must take several penalizing steps to travel from one to another, resulting into a highly inaccurate sampling. In contrast, the direct sampling procedure can seamlessly draw spin configurations belonging to all the regions according to the probability distribution, yielding much better samples.

Let us finish by mentioning that a plethora of different architectures have been implemented as NQS in recent years, such as convolutional neural network wave functions [184], and, more recently, group convolutional neural networks [216], which are especially convenient to implement certain symmetries directly built-in to the wave function.

IV.3 The gated recurrent unit ansatz

An ansatz that we will employ in section VI is the gated recurrent unit (GRU) ansatz, a type of recurrent neural network (RNN). This ansatz is autoregressive, and its conditional amplitudes are obtained through the following transformation, which depends on s_i, the local spin, and $\boldsymbol{h}_{i-1}$, the hidden unit coming out of the previous recurrent cell:

$$\phi_i(s_i, \boldsymbol{h}_{i-1}) := \mathrm{GRU}(s_i, \boldsymbol{h}_{i-1}) = A(s_i) \exp\{i\varphi(s_i)\}, \tag{4.27}$$

$$A(s_i) = \varsigma(\boldsymbol{U}_A \boldsymbol{h}_i + \boldsymbol{b}_A), \tag{4.28}$$

$$\varphi(s_i) = \varsigma\varsigma(\boldsymbol{U}_\varphi \boldsymbol{h}_i + \boldsymbol{b}_\varphi), \tag{4.29}$$

where ς and $\varsigma\varsigma$ denote a softmax and softsign transformation respectively, and where we have defined the conditional modulus $A(s_i)$ and phase $\varphi(s_i)$ that depend on matrices $\boldsymbol{U}_A$ and $\boldsymbol{U}_\varphi$ respectively, which each contain variational parameters. The hidden vector $\boldsymbol{h}_i$ is given by

$$\boldsymbol{h}_i = (1 - \boldsymbol{z}_j) \odot \boldsymbol{h}_{i-1} + \boldsymbol{z}_j \odot \tilde{\boldsymbol{h}}_j, \tag{4.30}$$

with the *latent hidden vector* $\tilde{\boldsymbol{h}}_j$ given by

$$\tilde{\boldsymbol{h}}_j = \tanh\left(\tilde{\mathbf{W}}[\boldsymbol{r}_j \odot \boldsymbol{h}_{j-1}, \boldsymbol{s}_j] + \tilde{\boldsymbol{b}}\right) \tag{4.31}$$

and the *update gate* $\boldsymbol{z}_j$ and *reset gate* $\boldsymbol{r}_j$:

$$\boldsymbol{z}_j = \mathrm{sig}\left(\boldsymbol{W}_z[\boldsymbol{h}_{j-1}; \boldsymbol{s}_{j-1}] - \boldsymbol{b}_z\right), \tag{4.32}$$

$$\boldsymbol{r}_j = \mathrm{sig}\left(\boldsymbol{W}_r[\boldsymbol{h}_{j-1}; \boldsymbol{s}_{j-1}] - \boldsymbol{b}_r\right), \tag{4.33}$$

where sig denotes the sigmoid funtion and where we have defined the matrices $\tilde{\boldsymbol{W}}, \boldsymbol{W}_r, \boldsymbol{W}_z$ and bias vectors $\tilde{\boldsymbol{b}}, \boldsymbol{b}_r, \boldsymbol{b}_z$ which are all variational parameters. $[\boldsymbol{h}_{j-1}; \boldsymbol{s}_{j-1}]$ denotes a concatenation of vectors $\boldsymbol{h}_{j-1}$ and $\boldsymbol{s}_{j-1}$, the latter corresponding to the one-hot encoding of the local spin configuration s_{j-1}. The important point here is that the total number of variational parameters scales quadratically with the dimension of the hidden vector d_h, since the $\boldsymbol{W}$ matrices each contain $d_h \times (d_h + 2)$ variational parameters for a one-hot encoding of the local spins s_i. One can see from this transformation that the GRU ansatz naturally respects the autoregressive property.

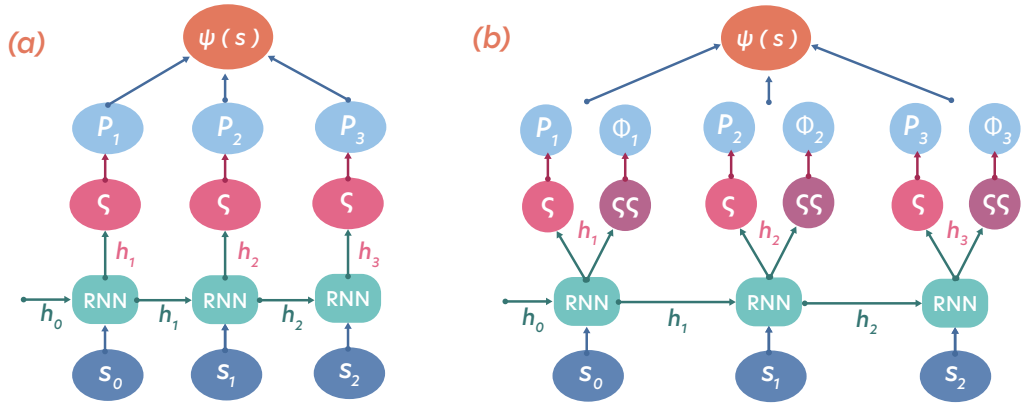

Figure 4.3: Pictorial representation of a recurrent neural network architecture for NQS. Panel (a) is for real-valued wave functions, which can be relevant for a certain class of problems, and panel (b) is for complex-valued wave functions. In both schemes, a local spin configuration s_i and a hidden vector h_i are fed into an RNN cell, which performs a nonlinear transformation. Then an activation function (ς, for softmax and/or $\varsigma\varsigma$, for softsign) is applied to obtain the final probability and/or phase corresponding to the configuration. At the end, the probabilities (and phases) are combined to obtain the final wave function amplitudes $\psi(\boldsymbol{s})$.

V Capacity and entanglement

A natural question arises regarding the expressive capacity of NQS and how different model architectures compare to each other in this respect. In particular, understanding how they compare to tensor network methods is crucial, as it would be natural that there is some limit, perhaps stemming from entanglement structure as for tensor networks. However, as is common in machine learning, this question is tricky to answer. A point we do not cover here but that is also important is that trainability is also a common issue for machine learning methods: an ansatz may be highly expressive but not so trainable, hence in practice it will not be useful.

Tensor networks have been a recurrent tool to study many-body systems. They are now well established and characterized, and their limits are well understood, hence they constitute a conceptual framework to study the complexity of quantum states. For this reason, there has been a significant community effort to study the relationship between tensor network states and NQS [194, 212, 217] , which provides insight about the expressive capacity of NQS [193]. Following the first introduction of NQS implementing RBM [181], early works focused on finding direct relationships between various kinds of RBM-based states and tensor network states [212, 217]. Recently, it has been proven that NQS can efficiently approximate, in logarithmic space, all efficiently contractible tensor network states with arbitrary precision [194]. Therefore, for every tensor network state there exists an equivalent NQS of polynomial size. Conversely, there are quantum states that can be efficiently described by NQS, whose representation in terms of tensor networks requires an exponential amount of parameters. Hence, tensor network states are a subset

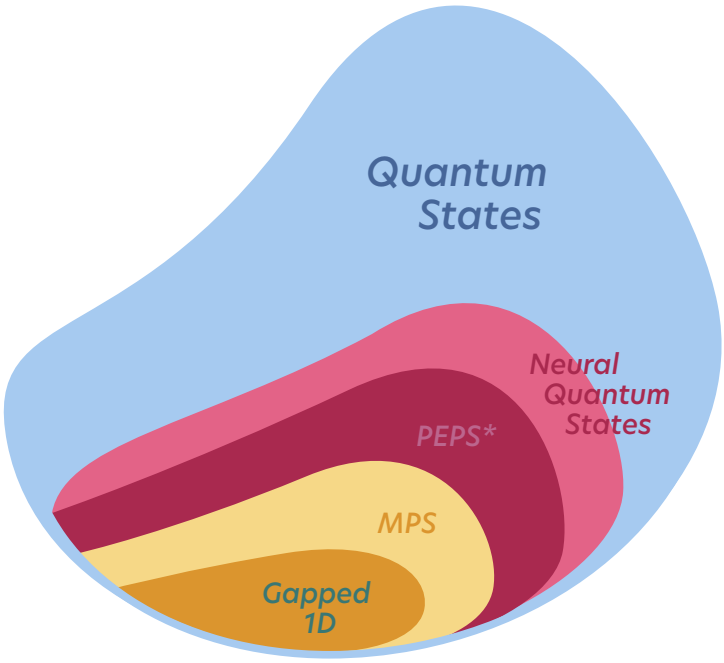

Figure 4.4: Expressive capacity of different classes of variational states, as explicitly proven in Ref. [194] by mapping tensor networks to NQS. PEPS* refers to a sub-class of projected entangled pair states, a generalization of MPS for two-dimensional systems. Adapted from [194].

of NQS [194], as depicted in 4.4.

More precisely, one study the entanglement scaling captured by the different ansätze. In a generic quantum many-body system with density matrix ρ, the entanglement entropy is defined as

$$S(\hat{\rho}) = -\mathrm{Tr}\left[\hat{\rho}\log_2\hat{\rho}\right] , \tag{4.34}$$

which is zero for any pure state. Let us consider a partition of the system in two subsets: I and its complementary O, as well as the reduced density matrix $\rho_I = \mathrm{Tr}_O[\rho]$. In general, ρ_I represents a mixed state, which can have nonzero von Neumann entanglement entropy. For a generic quantum state, the entanglement entropy of $\hat{\rho}_I$ grows with the volume of the cut. Thus, it corresponds to a *volume-law* scaling. NQS can efficiently capture such scaling with architectures ranging from very basic shallow ones, such as RBM [211], to more modern and deeper approaches, such as convolutional neural networks or RNNs [193].

We can understand most differences between the ansätze at an intuitive level by, simply, looking at how they are built. In 4.5, we provide a pictorial representation of the different connections that some ansätze can draw in a bi-dimensional system. Clearly, the MPS ansatz, depicted in 4.5(b), is the most locally restricted one, as it can only account for nearest neighbour connections in a snake-like pattern. This effectively limits

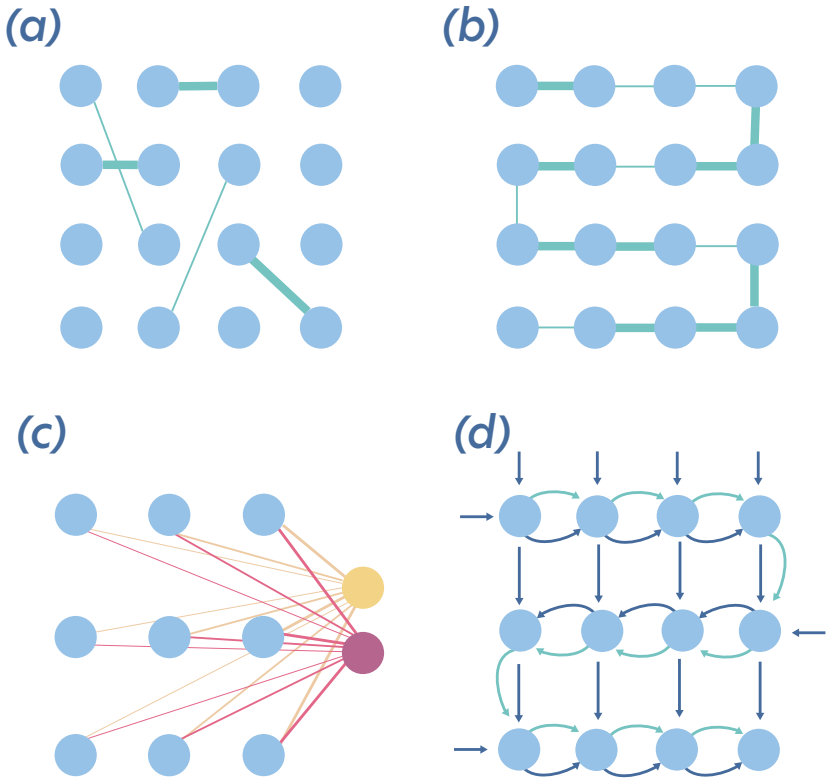

Figure 4.5: Schematic representation of various ansätze inspired by [217]. (a) The Jastrow ansatz draws connections between all possible pairs of sites. (b) The MPS ansatz draws connections between nearest-neighbour sites along a line. (c) The RBM ansatz connects all the sites to every hidden neuron, illustrated in different colours. (d) The RNN ansatz processes the state sequentially, following the green arrows. The dark blue arrows indicate the flow of information within the model. Arrows without a starting site correspond to free parameters. Sketch inspired by [218].

the entanglement that MPS can capture. The RNN ansatz, illustrated in 4.5(d), while it is limited to parse the state in the same pattern as the MPS, it has the freedom to account for additional information, allowing it to capture richer correlations.

In contrast, other ansätze such as the Jastrow or RBM wave functions, respectively illustrated in 4.5(a) and (c), can draw connections between arbitrary sites. The Jastrow ansatz can account for all possible pairs in the system, regardless of the distance. Then, the RBM anstaz is a generalization of the Jastrow by means of an auxiliary hidden layer of variable size. Through the hidden neurons, the ansatz is no longer limited to pairs, and it can actually consider up to all-to-all connections. This non-local character allows them to capture volume-law entanglement.

VI Dynamics of closed systems with NQS

We now turn to the application of NQS to dynamics of nonequilibrium quantum systems. Impressive results for the transverse-field Ising model on large lattices have been obtained using convolutional neural networks [184]. It is therefore of particular interest to continue investigating such problems that could yield important insights into nonequilibrium phenomena. In particular, studying the quantum Kibble-Zurek mechanism in higher dimensions has remained relatively untouched [219], and NQS is deemed to be employed for the dynamics of time-dependent finite-size systems.
While most previous works focus on the restricted Boltzmann machine (RBM) ansatz, more recent works on ground state search have employed networks that are closer to the state-of-the-art in machine learning such as autoregressive convolutional models [182] or recurrent neural networks [220], for which the accuracy of the variational ground state energy was significantly improved. These networks have a so-called autoregressive structure, which means that one can perform direct sampling of uncorrelated configurations for arbitrarily large system sizes, thereby reducing the number of required samples. It is therefore of crucial importance to investigate the use of similar networks for quantum dynamics.

However, the widespread application of more complex and autoregressive neural networks for quantum dynamics has been held back by issues arising with time-dependent variational Monte-Carlo (t-VMC) [221]. This procedure, used to propagate an NQS in time, involves the inversion of a stochastically constructed singular matrix, which makes it particularly prone to noise. While some regularization techniques have helped improving the accuracy of the method [184, 222], accessing all regimes at long times via t-VMC remains a challenge [22, 222–224].

In this section, we show that the stability of t-VMC strongly depends on the chosen ansatz, and that, in particular, it fails when applied to those based upon recurrent neural networks (RNN). To circumvent this issue, we propose an alternative scheme to numerically solve the dynamics of quantum systems. The scheme consists in casting an arbitrary Runge-Kutta integration scheme of any order into minimizing a variational distance at each time step, while only involving a polynomial overhead. This enables one to employ recurrent neural networks for quantum dynamics, which leads to a drastic reduction the the number of required samples. Our scheme can be implemented for any order of a chosen integration method, for a polynomial overhead in memory. We then apply our scheme to both time-dependent and sudden quenches, enabling one to recover Kibble-Zurek scaling laws for large system sizes and high precision on the dynamics during the full quench.

VI.1 Issues with t-VMC

In order to propagate a NQS in time according to some Hamiltonian $\hat{H}$, one should minimize the following variational distance:

$$D(\tilde{\boldsymbol{\theta}}) = \mathrm{dist}\left(|\psi_{\tilde{\boldsymbol{\theta}}}\rangle\,, \exp\Big\{-i\hat{H}\tau\Big\}\,|\psi_{\boldsymbol{\theta}}\rangle\right) \tag{4.35}$$

over the variational parameters $\tilde{\boldsymbol{\theta}}$ after each time step τ. By expanding $|\psi_{\tilde{\boldsymbol{\theta}}}\rangle$ as a Taylor expansion in the variational parameters to obtain the following equation to be solved at

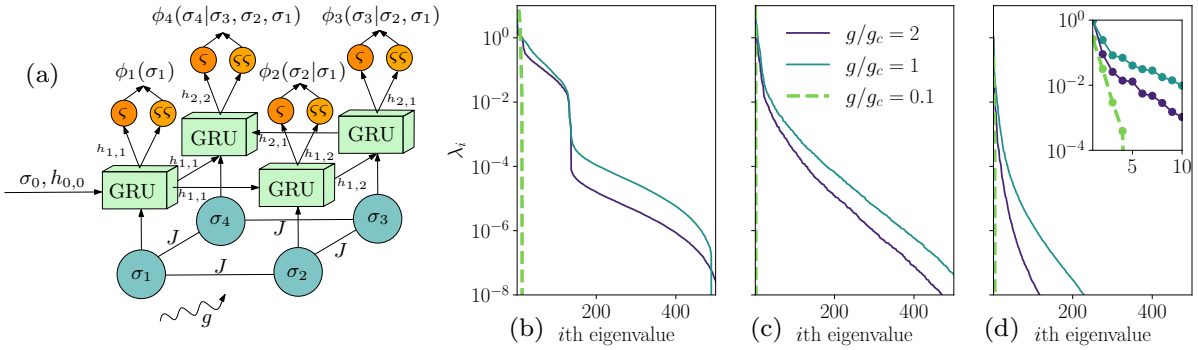

Figure 4.6: (a): Scheme of the GRU ansatz for a two-dimensional system of $L = 4$ spins with open boundary conditions. The system is subject to the TFI Hamiltonian, and configurations are fed into GRU in an autoregressive fashion, each cell being fed with a hidden vector $\boldsymbol{h}_{i,j}$ and a spin σ_i. After the GRU cell, a softmax (ς) and a softsign ($\varsigma\varsigma$) transformation are applied to obtain the amplitudes and the phases of the conditional amplitudes $\phi_i(\sigma_i|\sigma_{i-1},..,\sigma_1)$. (b) Spectra of the quantum geometric tensors $\mathbf{S}$ for different ansätze optimized to the ground state of the TFI model, for various values of g/J. The left panel is for a complex RBM ansatz with $\alpha = 2$, the central panel for a two three layer convolutional autoregressive network with complex parameters (ARNN), and the right panel for a GRU ansatz with real parameters with $d_h = 10$, for which is plotted the spectrum of the imaginary part of the $\mathbf{S}$ matrix, as this is the quantity to invert for ansätze with real parameters.

each time step, so as to obtain $\dot{\theta}_k := \partial_t \theta_k$ [181]:

$$S_{k,k'}\dot{\theta}_{k'} = -iF_k, \tag{4.36}$$

with $F_k = \langle O_k^* E_{\text{loc}}\rangle - \langle O_k^*\rangle\langle E_{\text{loc}}\rangle$, and $S_{k,k'} = \langle O_k^* O_{k'}\rangle - \langle O_k^*\rangle\langle O_{k'}\rangle$ when keeping only second-order terms. The O_k terms are the log-derivatives of the NQS ansatz, as given by

$$O_k(\boldsymbol{\sigma}) = \partial_{\theta_k} \ln \psi_{\boldsymbol{\theta}}(\boldsymbol{\sigma}). \tag{4.37}$$

With most machine learning algorithms, these are obtained through automatic differentiation. The local energy E_{loc} is local estimator which evaluates as

$$E_{\text{loc}}(\boldsymbol{\sigma}) = \sum_{\boldsymbol{\sigma}'} \frac{\psi_\theta(\boldsymbol{\sigma}')}{\psi_\theta(\boldsymbol{\sigma})} \langle \boldsymbol{\sigma}|\hat{H}|\boldsymbol{\sigma}'\rangle\,. \tag{4.38}$$

The parameters $\boldsymbol{\theta}$ are then updated at each time step according to their derivatives $\dot{\theta}_k$ using a numerical solver. This procedure is known as time-dependent variational Monte Carlo (t-VMC), and is related to natural gradient methods in machine learning [181]. Several issues can arise: (i) The $\mathbf{S}$ matrix to be inverted is in general singular, which makes the process extremely sensitive to stochastic variations coming from sampling. Hence, regularization must be used to obtain a non-diverging parameters derivative $\dot{\theta}_k$ [222]. This increases the stability of the method, although impacting its accuracy, making it challenging to obtain accurate long-time dynamics. (ii) State-of-the-art regularization

techniques involve the inversion of the $\mathbf{S}$ matrix via its singular-value decomposition, which makes the complexity of the method of order $O(N_{\text{par}}^3)$, with N_{par} the number of parameters for the NQS.[5] This is one of the reasons why natural gradient methods is rarely used for other machine-learning applications, which can involve models with up to billions of parameters. (iii) For NQS with real parameters, one may split the real and imaginary parts of Eq. (4.36) and solve either of the two resulting equations [225]:

$$\begin{aligned}\text{Re}(S_{kk'})\dot{\theta}_{k'} &= 0\\ \text{Im}(S_{kk'})\dot{\theta}_{k'} &= -F_k.\end{aligned}$$

Solving the first forces one to impose a condition on $\dot{\theta}_k$ to obtain a nontrivial solution ($\dot{\boldsymbol{\theta}} \neq \mathbf{0}$), and solving the second one is difficult in general since the diagonal elements of $\text{Im}(\mathbf{S})$ are zero ($\mathbf{S}$ is Hermitian) and remaining off-diagonal elements are close to 0 for various ansätze with real parameters we have tested. When the minimum and maximum eigenvalues of a matrix are highly separated, $\lambda_{\text{min}}/\lambda_{\text{max}} \sim \epsilon$, $\epsilon \to 0$, this is known as poor conditioning [196]. It makes regularization schemes even harder to implement and yields large parameter time derivatives, thereby requiring impractically small time steps for t-VMC to work.

To further investigate such issues, let us consider the transverse-field Ising (TFI) Hamiltonian:

$$\hat{H}_{\text{TFI}} = -J \sum_{\langle m,n \rangle} \hat{\sigma}_m^z \hat{\sigma}_n^z + g \sum_m \hat{\sigma}_m^x, \tag{4.39}$$

with J the nearest-neighbor coupling strength, and g the transverse field strength. As mentioned in section II.1, at zero temperature this model exhibits a quantum phase transition for $g_c = J$ in one dimension and for $g_c \approx 3.044J$ in two dimensions [35]. For $J > 0$, the transition separates a ferromagnetic [6] phase from a paramagnetic phase, where the spins tend to align with the transverse-field. In the former, when $g \ll J$, the ground state is degenerate and spanned by $|\uparrow, \uparrow, \ldots, \uparrow\rangle$ and $|\downarrow, \downarrow, \ldots \downarrow\rangle$, while in the latter, when $g \gg J$ the ground state is $|\rightarrow, \rightarrow, \ldots, \rightarrow\rangle$, with $|\rightarrow\rangle$ the eigenstate of $\hat{\sigma}^x$. This model serves as a convenient benchmark for the dynamics of NQS [181, 184, 185].

In Fig. 4.6(b), the spectra of $\mathbf{S}$, computed at the ground state of the TFI Hamiltonian is shown for three different ansätze: a restricted Boltzmann machine (RBM), an autoregressive convolutional network with complex parameters (ARNN), and the GRU ansatz presented in the previous section, for different values of g/g_c (we have fixed $J = 1$) for a two-dimensional array of $N = 16$ spins. Two key observations can be made from this figure: first, all spectra become more singular as g/J decreases, something that makes t-VMC less applicable to ferromagnetic-like states, whether it is for ground state search or for dynamics, which involves the full spectrum of the Hamiltonian. One can also see that the RBM spectrum is the less singular, and presents a shell-like structure, making it simpler to regularize. The ARNN ansatz is similar to this, and presents many eigenvalues of order $O(1)$. In stark contrast, the GRU spectrum indicates that the variational

[5] Alternatively, one can also perform the inversion with iterative methods such as conjugate gradients, thereby reducing the complexity. This, however, does not enable all forms of regularization.

[6] or anti-ferromagnetic if $J < 0$. For simplicity we will consider $J > 0$ throughout the paper.

space that is extremely flat for all values of g/J, as very few eigenvalues are of a high enough value, and the magnitude of the eigenvalues decrease extremely rapidly. This inherent poor conditioning implies that if such states are the initial states of a dynamical evolution, the inversion of $\mathbf{S}$ will yield a parameter update $\dot{\boldsymbol{\theta}}$ whose norm is large with respect to $\|\boldsymbol{\theta}\|$, forcing one to dramatically reduce the time step which makes calculations impractical. This last spectrum is tricky to regularize, as there is no clear separation of the eigenvalues.

VI.2 Variational Runge-Kutta algorithms

As identified in the previous section, t-VMC suffers from a number of limitations that make it unsuitable for use with certain ansätze, in particular for the GRU ansatz. We propose to circumvent these issues by resorting back to the original Dirac-Frenkel variational principle (Eq. (4.35)) and solving an optimization problem at each time step. The general form of this problem is the minimization of the distance

$$D(\tilde{\boldsymbol{\theta}}) = \mathrm{dist}\Big(|\psi_{\tilde{\boldsymbol{\theta}}}\rangle\,, \hat{T}\,|\psi_{\boldsymbol{\theta}}\rangle\Big), \tag{4.40}$$

where $\tilde{\boldsymbol{\theta}}$ denotes the set of variational parameters to optimize, $\boldsymbol{\theta}$ those at the previous time step t, and $\hat{T}$ a propagator evolving the state of the system between times t and $t+\tau$ under the action of the Hamiltonian of interest. While expanding the propagator to first order in τ yields a valid first-order update of the variational state, we instead propose to build a variational principle upon a s-order Runge-Kutta approximant, as generated by a propagator $\hat{T}_s$ such that:

$$|\psi(t+\tau)\rangle = \hat{T}_s(t)\,|\psi(t)\rangle + O(\tau^{s+1}). \tag{4.41}$$

General explicit expressions for this operator are derived in Appendix A for arbitrary diagonally implicit Runge-Kutta schemes. In particular, for the standard (second-order) Heun method used throughout the paper, one has:

$$\hat{T}_{\mathrm{Heun}}(t) = \hat{\mathbb{1}} - i\tau\hat{H}(t) - \frac{\tau^2}{2}\hat{H}(t+\tau)\hat{H}(t). \tag{4.42}$$

The variational method induced by this scheme involves an integration error per time step of third order in τ. In practice, rather than the Fubini-Study distance, we use the following numerically well-behaved metric based upon the quantum fidelity:

$$\mathrm{dist}\Big(|\psi\rangle\,, |\phi\rangle\Big) = 1 - \frac{|\langle\psi|\phi\rangle|^2}{\langle\psi|\psi\rangle\,\langle\phi|\phi\rangle}, \tag{4.43}$$

which for normalized quantum states reduces to

$$D(\tilde{\boldsymbol{\theta}}) = 1 - \frac{|\,\langle\psi_{\tilde{\boldsymbol{\theta}}}|\hat{T}_s|\psi_{\boldsymbol{\theta}}\rangle|^2}{\langle\psi_{\tilde{\boldsymbol{\theta}}}|\hat{T}_s^\dagger\hat{T}_s|\psi_{\boldsymbol{\theta}}\rangle}. \tag{4.44}$$

For the Heun method, one has

$$= 1 - |\,\langle\psi_{\tilde{\boldsymbol{\theta}}}|\hat{T}_{\mathrm{Heun}}|\psi_{\boldsymbol{\theta}}\rangle|^2 + O(\tau^4) \tag{4.45}$$

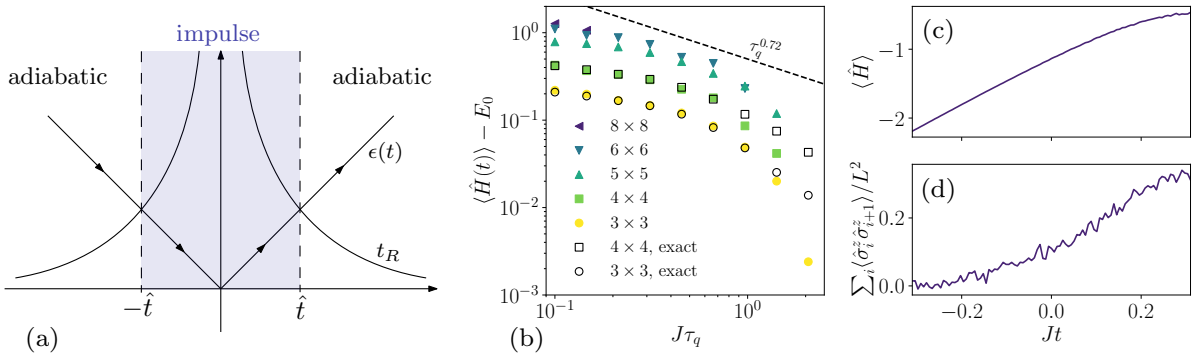

Figure 4.7: (a) Scheme of the Kibble-Zurek mechanism according to the adiabatic-impulse approximation. At times $\pm\hat{t}$, the relaxation time of the system t_R crosses the time scale of the quench, which defines the limit between the adiabatic region, where the system essentially stays in the ground state of the system at that time, and the impulse regime, where nonequilibrium dynamics take place. (b) Scaling of the injected energy $\hat{H}(\tau_q) - E_0$ as a function of the quench time τ_q for different system sizes $N = L \times L$. One can see that for fast quenches, the NQS dynamics predict very well the injected energy. The data for system sizes with more than $L = 5$ cannot be simulated via exact diagionalization, hence they are compared to the $\tau_q^{0.72}$ scaling obtained in ref. [219]. (c) Average energy for a quench of total time $J\tau_q = 0.31$, for a 6×6 system and a GRU ansatz with $d_h = 8$, with $N_s = 200$ samples. (d) Average correlation function along the $z-$axis for the same quench as in (c).

with

$$\langle\psi_{\tilde{\boldsymbol{\theta}}}|\hat{T}_s|\psi_{\boldsymbol{\theta}}\rangle \equiv \langle\psi_{\tilde{\boldsymbol{\theta}}}|\hat{T}_{\text{loc}}|\psi_{\tilde{\boldsymbol{\theta}}}\rangle\,, \tag{4.46}$$

where $\hat{T}_{\text{loc}}$ is an operator acting as a local estimator and whose non-zero entries are given by

$$\begin{aligned} T_{\text{loc}}(\boldsymbol{\sigma}) &:= \langle\boldsymbol{\sigma}|\hat{T}_{\text{loc}}|\boldsymbol{\sigma}\rangle \\ &= \sum_{\boldsymbol{\sigma}'} \frac{\psi_{\boldsymbol{\theta}}(\boldsymbol{\sigma}')}{\psi_{\tilde{\boldsymbol{\theta}}}(\boldsymbol{\sigma})} \langle\boldsymbol{\sigma}|\hat{T}_s|\boldsymbol{\sigma}'\rangle\,. \end{aligned} \tag{4.47}$$

It thus appears from the above that any s-order Runge-Kutta update can be cast into a variational problem consisting in maximizing the squared expectation value of an observable. Furthermore, this expectation value can be efficiently sampled at every step of the optimization process as the average

$$\langle\psi_{\tilde{\boldsymbol{\theta}}}|\hat{T}_{\text{loc}}|\psi_{\tilde{\boldsymbol{\theta}}}\rangle = \mathbb{E}_{\boldsymbol{\sigma}\sim|\psi_{\tilde{\boldsymbol{\theta}}}|^2}[T_{\text{loc}}(\boldsymbol{\sigma})]. \tag{4.48}$$

The error of an integration using the Heun propagator is of the order of τ^3, which will be the lower bound for the error at each time step for the variational method we present. Note that this involves summing over the connected elements of $\hat{T}$ as can be seen in Eq. (4.47). This will in general become more expensive as the order s increases, as powers of the Hamiltonian will be contained in $\hat{T}$. The number of connected elements will depend

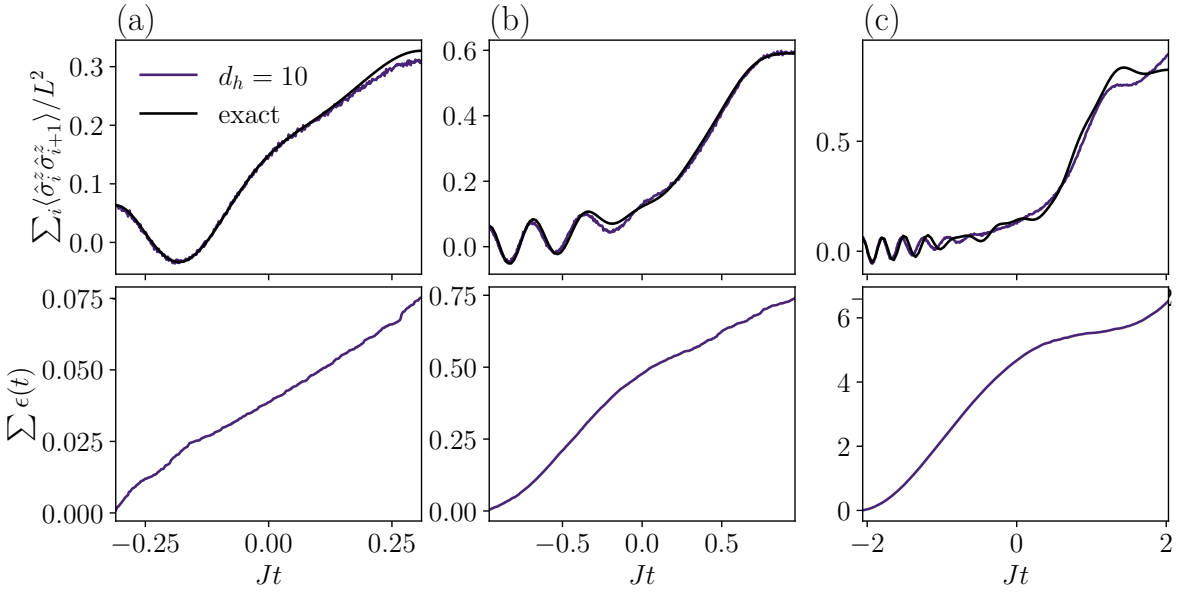

Figure 4.8: Numerical results obtained for a linear quench for various values of the total quench time τ_q. On the upper panels, the average correlation function is shown as a function of time for both the exact (dashed line) and the GRU simulations (solid lines). The error as a function of time $\epsilon(t)$ is shown in the lower panels. For panels (b) and (c), the time step was chosen to be linear and non-constant, as the characteristic time scale changes as a function of time. Parameters for each quench are are given in the appendix.

on the sparsity of the Hamiltonian; for the transverse-field Ising model for instance, the number of connected elements in $\hat{H}^s$ scales as N^s (which means N^2 for a second-order explicit integration scheme).

VI.3 Application: critical quench dynamics

The Kibble-Zurek mechanism

The Kibble-Zurek mechanism [25, 26] predicts the formation of topological defects in quenched systems undergoing a second-order phase transition as the parameters are linearly swept across a critical point. This universal behavior stems from the fact that there exists a time, denoted $\hat{t}$, at which the characteristic evolution time of the system (the relaxation time t_R, related to the inverse of the energy gap) becomes larger than the characteristic quench time τ_q, as schematically shown in Fig. 4.7(a). Before $\hat{t}$, the dynamics remains quasi-adiabatic, as the Hamiltonian parameters are tuned slowly with respect to the time scale of the system. After this time, non-equilibrium dynamics takes place as critical slowing down sets in and the system dynamics becomes slower than the Hamiltonian parameters sweep. Ignoring specific details of this dynamics, one can derive a scaling law for the density of created defects at the end of the linear quench [26], which reads

$$\langle \hat{n}_d(\tau_q) \rangle \sim \tau_q^{-d\nu/(z\nu+1)} \tag{4.49}$$

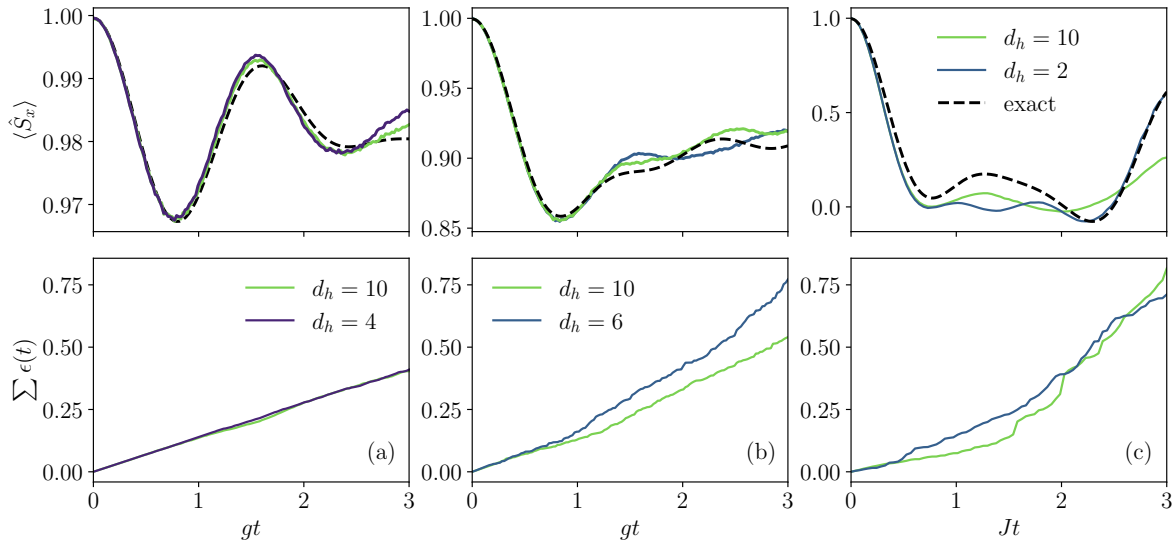

Figure 4.9: Sudden quench dynamics induced by quenching the Hamiltonian parameters to $g/g_c = 2$, $g/g_c = 1$ and $g/g_c = 1/10$ for various hidden vector dimensions d_h, starting from an initial paramagnetic state. Upper panels show the average magnetization along the x axis $\langle \hat{S}_x \rangle = \sum_i \hat{\sigma}_i^x / L^2$ and lower panels show the cumulative error as a function of time.

with d the dimensionality of the system and z, ν universal critical exponents. Note that $\hat{n}_d$ will depend on the geometry of the system, and is not always easily accessible experimentally. For this reason some works consider other observables, such as the number of domain walls [226], or the injected energy density [219], defined as

$$\mathcal{E} = \frac{1}{L^2} \left[\langle \hat{H}(t) \rangle - E_0(t) \right]. \tag{4.50}$$

This quantity is a witness of the injected defects, which are excitations of the system with respect to its ground state at time t, hence $\mathcal{E}$ must scale as the density of defects. This quantity is convenient as it does not depend on the geometry of the physical system under consideration. In Ref. [219], authors extract a scaling law in $\tau_q^{0.72}$ for the 2D transverse-field Ising model based upon various numerical methods that they used to estimate the gap of the system at different times to extract a correlation length. We consider a similar protocol in this paper, in which both J and g are varied in time. The Hamiltonian reads

$$\hat{H}_{\mathrm{TFI}} = -J(t) \sum_{\langle m,n \rangle} \hat{\sigma}_m^z \hat{\sigma}_n^z + g(t) \sum_m \hat{\sigma}_m^x, \tag{4.51}$$

with $J(t) = J(1 - t/\tau_q)$, $g(t) = g_0(1 + t/\tau_q)$, with $g_0 = g_c$ and $J = 1$, which will be used as a reference unit in the rest of the paper. Note that this type of quench involves all regimes ($0 \leq g/J < +\infty$), hence stochastic reconfiguration does not enable one to access the full dynamics [219]. Our approach does not suffer from such issues, which is one of its major strengths. In Fig. 4.7(b), we show the injected energy (not rescaled for clarity) $\langle \hat{H}(t) \rangle - E_0(t)$ for various system sizes, for different values of the quench time. One can

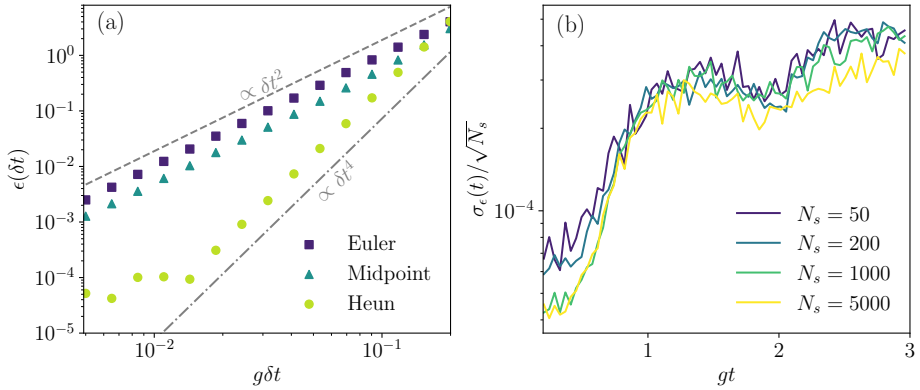

Figure 4.10: (a): residual error ϵ as a function of the time step δt for the Euler method, the midpoint method, and the Heun integration methods averaged over the 10 first time steps of quench dynamics of the TFI model to $g/g_c = 2$. (b) Standard deviation of the error as a function of time rescaled by $\sqrt{N_s}$. Curves collapse, which is what is predicted by a direct sampling scheme. This enables us to estimate the required number of samples to obtain a given error. Note that as the dynamics progress, the variance of the residual error rises, indicating a rising difficulty to find the best minima in time.

see that the exact results are closely matched for fast quenches, and precision drops for slower quenches ($J\tau_q > 1$). This is related to the fact that dynamics become almost fully adiabatic, hence the NQS dynamics is not precise enough to capture the small amount of injected energy. The presented method most accurately reproduces the dynamics of the system in the non-adiabatic region, where the dynamics presents universal features; this is the regime of relevance when studying the Kibble-Zurek mechanism. For lattices above 5×5, one cannot compare our results with exact results, but the scaling predicted in [219] is recovered, indicating that the results are reliable. We also show correlation functions for various quench times, as well as the residual error corresponding to the full dynamics, as given by

$$\epsilon(t) = \mathrm{dist}\left(|\psi_{\tilde{\boldsymbol{\theta}}}\rangle, \hat{T}\,|\psi_{\boldsymbol{\theta}}\rangle\right) \tag{4.52}$$

$$= 1 - |\langle\psi_{\tilde{\boldsymbol{\theta}}}|\hat{T}\,|\psi_{\boldsymbol{\theta}}\rangle\,|^2. \tag{4.53}$$

VI.4 Sudden quenches

One can also investigate the nonequilibrium dynamics of many-body systems without relying on regularization hyperparameters and instabilities stemming from stochastic reconfiguration. As a benchmark of the presented method, we consider the two-dimensional time-independent transverse-field Ising model, defined in Eq. (4.39). We prepare the system in the ground state of the TFI Hamiltonian for $g \gg J$, which is $|\psi_0\rangle = |\rightarrow, \rightarrow, \ldots, \rightarrow\rangle$, and quench the magnetic field to the values $g/g_c = 2$, 1, and 1/10; results are displayed in panels (a), (b), and (c) of Fig. 4.9 respectively. By doing this, one spontaneously creates excitations of all the eigenstates of the Hamiltonian corresponding to the final value of the parameters. This is of high interest experimentally, as it can be used to probe

the properties of non-integrable systems close to the critical point. It is expected that a quench to the critical point is the most difficult to simulate, since it involves states that are correlated at all scales in the thermodynamical limit. In contrast to other approaches, the use of direct sampling here ensures that the dynamics will be accurate for a lower number of samples. Here, we have imposed the reflection and $\mathbb{Z}_2$ symmetries for panels (a) and (b), and no symmetry for panel (c), where, surprisingly, imposing symmetries was found to degrade the accuracy.

VI.5 Error analysis

The residual error ϵ depends on the chosen time step δt. In Fig. 4.10, this is shown as a function of the chosen time step δt for different integration schemes, namely Euler, (implicit) midpoint, and Heun methods. The propagators $\hat{T}$ for the Euler and the implicit midpoint method are respectively given by

$$\hat{T}_{\text{Eul}} = \hat{\mathbb{1}} - i\tau\hat{H} \tag{4.54}$$

$$\hat{T}_{\text{mid}} = \hat{\mathbb{1}} - i\tau\hat{H} + \frac{\tau^2}{4}\hat{H}^2, \tag{4.55}$$

and the Heun propagator was given by Eq. (4.42). Euler is a first-order method, while the midpoint and Heun methods are both second-order methods. The midpoint method conserves a symplectic symmetry in the exact case, which for the Schrödinger equation corresponds to the energy. Note, however, that propagating an NQS in time, this is no longer guaranteed, as the time evolution is approximated stochastically. Using $\hat{T}_{\text{mid}}$ as a propagator is equivalent to minimizing the distance proposed in Ref. [185]. Interestingly though, when using the midpoint method, the norm of the denominator of Eq. (4.44) contains terms of order 2, whereas they vanish for the Heun method. Indeed, the implicit midpoint method yields an update of the form

$$|\psi(t+\delta t)\rangle = \hat{T}_{\text{mid}}(t)\,|\psi(t)\rangle + O(\delta t^3), \tag{4.56}$$

and as such may in principle be regarded as being of second order, provided one is able to approximate $|\psi(t+\delta t)\rangle$ with an NQS up to an error of order $O(\delta t^3)$. However, this condition cannot be satisfied with a normalized ansatz such as those used in this work. Indeed, the norm of the ideal updated state to be matched is given by:

$$\langle\psi(t)|\hat{T}_{\text{mid}}^\dagger\hat{T}_{\text{mid}}|\psi(t)\rangle = \langle\psi(t)|\left(\hat{\mathbb{1}} + \frac{3\delta t^2}{2}\hat{H}^2\right)|\psi(t)\rangle + O(\delta t^3) = 1 + O(\delta t^2)\,. \tag{4.57}$$

This must depart from 1 to second order in δt, implying that the implicit midpoint method effectively yields a first-order update when using ansätze normalized by design.

In contrast to the midpoint method, the Heun method does not suffer from this issue, indeed:

$$\hat{T}_{\text{Heun}}^\dagger\hat{T}_{\text{Heun}} = \left(\hat{\mathbb{1}} + i\delta t\hat{H} - \frac{\delta t^2}{2}\hat{H}^2\right)\left(\hat{\mathbb{1}} - i\delta t\hat{H} - \frac{\delta t^2}{2}\hat{H}^2\right) = \hat{\mathbb{1}} + O(\delta t^4)\,. \tag{4.58}$$

This property allowed us to simplify the expression of the distance in Eq. (4.44), considerably reducing the complexity of evaluating the gradients of the loss function.

These considerations generalize to the time-dependent scenario, where for the Heun method we have

$$\hat{T}^{\dagger}_{\text{Heun}}(t)\hat{T}_{\text{Heun}}(t) = \left(\hat{\mathbb{1}} + i\delta t\hat{H}(t) - \frac{\delta t^2}{2}\hat{H}(t)\hat{H}(t+\delta t)\right)\left(\hat{\mathbb{1}} - i\delta t\hat{H}(t) - \frac{\delta t^2}{2}\hat{H}(t+\delta t)\hat{H}(t)\right)$$
$$= \hat{\mathbb{1}} + \delta t^2\hat{H}(t)^2 - \frac{\delta t^2}{2}\{\hat{H}(t), \hat{H}(t+\delta t)\} + O(\delta t^3). \tag{4.59}$$

Provided one can expand $\hat{H}(t+\delta t) = \hat{H}(t) + \delta t\partial_t\hat{H}(t) + O(\delta t^2)$, as is the case for any analytic quench, the second-order terms in Eq. (4.59) cancel out exactly, implying that the Heun method remains of second order when using normalized ansätze. Therefore, in this context, this method will effectively be of first-order.

One can clearly observe in Fig. 4.10(a) a δt^4 scaling for the Heun method as expected from the form of the distance we minimize, and a δt^2 error for the two other methods. Hence the midpoint method carries the disadvantage of having an error comparable to an Euler integration while having the computational complexity of the Heun method. One can also see that for the Heun method there is a cutoff around $g\delta t = 0.01$, for which the error cannot be reduced by decreasing the time step. This is due to the fact that the method displays a systematic error coming from the optimization for a given set of hyperparameters (number of samples, learning rate, number of training step). Nevertheless, the error can always be decreased by increasing the number of samples. Note that in the constant error regime, reducing the time step will be penalizing as the total number of time steps at fixed total simulation time. Hence there is an optimal time step to be chosen, which we find to be $g\delta t \approx 0.01$ for various hyperparameter choices. In Fig. 4.10(b), the standard deviation of the error is shown for each time step of a sudden quench from a paramagnetic phase to $g/g_c = 2$. One can distinguish two regimes, separated by $gt \approx 1$, where before this time the standard deviation of the residual error grows, and then saturates after this time. This can be attributed to a harder optimization error after a given time, where the minimum of the distance is found but with a greater variance. These results are reminiscent of those presented in a recent work where authors demonstrate that quantum states become harder to fit with an NQS at lated times after a sudden quench, which is not always related to the growth of entanglement [227]. Standard deviations are shown for various values of N_s, the number of samples considered for each optimisation. One can see that $\sigma_\epsilon(t)$ decreases as $\sqrt{N_s}$, as predicted by a direct sampling scheme.

VII Dynamics of open quantum systems with NQS

VII.1 Representing density matrices

Extensions of neural network techniques to the time evolution of open quantum systems have been proposed. A first approach relies on purification [186], which obliges one to use Restricted Boltzmann Machine (RBM) as the ansatz [187–189, 228] for the density matrix, a requirement that bounds the representative power of such methods. More recently, new approaches based on autoregressive networks were proposed for the time

evolution of open quantum systems. These are based on the Positive-Operator-Value-Measurement (POVM) formalism, in which the density matrix is represented as an average over outcomes of measurements [229, 230]. By recasting the Lindblad master equation into a differential equation on a probability distribution, a time-dependent variational principle can be obtained. An issue with this approach is that the space of density matrices (positive semi-definite) is more restricted than that of arbitrary probability distributions, therefore the time evolving map may project distributions out of the space of density matrices, for example by violating positivity. Finally, a more recent approach consists in building an autoregressive and positive density matrix that relies on a construction based on Gram matrices [231].
We also present the extension of the dynamical-corner space method, which consists in representing each state of the corner basis by an NQS. This restricts the method to low-entropy quantum systems, therefore will not be practical to study, for instance, steady-states of critical systems (where the entropy can be high), but ensures positivity and avoids enforcing a certain structure for the ansatz. In principle, one could even combine different neural network architectures with such an approach, but this is generally computationally impractical.

The RBM approach

One wants to obtain a generative model for the coefficients of the density matrix $\rho_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{s'})$, defined as:

$$\hat{\rho}_{\boldsymbol{\theta}} = \sum_{s,s'} \rho_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{s'}) \, |s\rangle\langle s'| \, . \tag{4.60}$$

The first approach consists in using an RBM to represent a pure state $|\psi_{\boldsymbol{\theta}}\rangle$ that belongs to an enlarged Hilbert space $\mathcal{H} = \mathcal{H}_s \otimes \mathcal{H}_a$, with $\mathcal{H}_s$ describing the physical spins and $\mathcal{H}_a$ describing the ancilla spins. Provided that there are at least as many ancilla spins as physical spins, one can always write:

$$\rho_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{s'}) = \sum_{\boldsymbol{a}} \psi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a}) \psi^*_{\boldsymbol{\theta}}(\boldsymbol{s'}, \boldsymbol{a}). \tag{4.61}$$

However, in the general case, the sum over $\boldsymbol{a}$ is intractable, as there are 2^L elements to sum over. The key point is that since the RBM ansatz has an analytical form, this sum can be performed analytically and a closed expression for $\rho_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{s'})$ can be obtained [187]. This is a calculation much like that which led to Eq. (4.25), where a trace over hidden spins was performed analytically. This leads to a positive and Hermitian ansatz for the density matrix, with derivatives of the density matrix coefficients also having an analytical form. However, as we have seen, RBMs are restricted in terms of capacity, hence other approaches were quickly proposed to obtain more freedom on the choice of the neural network.

The Gram-Hadamard matrix approach

The Gram-Hadamard approach proposes to re-express the coefficients of a density matrix by using a Gram matrix form and useful properties of the Hadamard product [231]. The

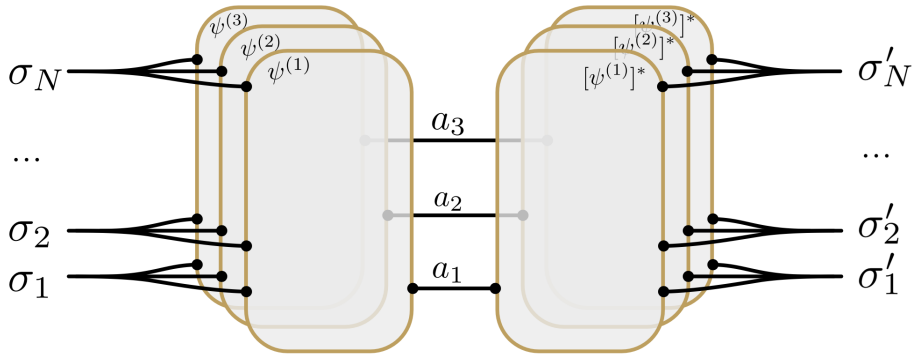

Figure 4.11: Tensor network diagram of the Gram-Hadamard ansatz for the density matrix. The plaquette structure of the ansatz is shown: with K plaquettes of rank R, the obtained density matrix has rank R^K, which enables for states that have a high entropy to be efficiently represented. Adapted from [231].

Gram matrix of a set of R vectors or quantum states $\left|\psi^{(a)}\right\rangle, a \in 1, \ldots, R$ is given by

$$G = \sum_{a=1} \left|\psi_{\boldsymbol{\theta}}^{(a)}\right\rangle\!\left\langle\psi_{\boldsymbol{\theta}}^{(a)}\right| . \tag{4.62}$$

Note the similarity of this form to that of the low-rank density matrix presented in chapter 3, $\hat{\rho} = \hat{C}\hat{C}^\dagger$. Now, by noting that the element-wise or Hadamard product of two positive semi-definite matrices remains positive semi-definite [231], we can define K Gram matrices, and the final Gram-Hadamard ansatz is built as:

$$\hat{\rho}_{\boldsymbol{\theta}} = \bigodot_{h=1}^{K} G^{(h)} \tag{4.63}$$

$$= \bigodot_{h=1}^{K} \left(\sum_{a=1}^{R} \left|\psi_{\boldsymbol{\theta}}^{(a)}\right\rangle\!\left\langle\psi_{\boldsymbol{\theta}}^{(a)}\right| \right)^{(h)} . \tag{4.64}$$

where the element-wise product is taken in the same basis for each Gram matrix. The elements of the density matrix are therefore given by

$$\hat{\rho}_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{s'}) = \prod_{h=1}^{K} \sum_{a=1}^{R} \psi_{\boldsymbol{\theta}}^{(a,h)}(\boldsymbol{s}) \psi_{\boldsymbol{\theta}}^{*(a,h)}(\boldsymbol{s'}). \tag{4.65}$$

where $\psi_{\boldsymbol{\theta}}^{(a,h)}(\boldsymbol{s})$ are complex amplitudes. Note that this ansatz closely resembles the purification of Eq. 4.61. A diagram of the ansatz is shown in Fig. 4.11. The power of such a construction is that by restricting the rank R of the Gram matrices to a small number, the rank of the final density matrix is obtained as R^K. This form is reminiscent of the locally-purified density operator presented in section III.6 of Chapter 1, where a partial trace is taken on finite-rank states, therefore is tractable. However, there is no local purification here, as $\psi^{(h)}$ are global states. A natural question that arises is whether this construction could also be used with tensor networks architectures as an alternative to LPDOs, hence creating an alternative positive tensor network ansatz. Note that Gram-Hadamard construction may constrain the entanglement structure of $\hat{\rho}_{\boldsymbol{\theta}}$, and further studies are needed

to verify this, in which case the full power of machine learning methods could be deployed to study open quantum systems. Finally, in the same work [231], a similar construction has been proposed that explictily enforces the autoregressive property, to enable for direct sampling and normalized wavefunctions. However, again this construction may constrain the entanglement structure of the ansatz.

The POVM approach

Another approach that does not enforce positivity but allows absolute freedom on the choice of the ansatz will be referred to as the POVM approach. Here, we decompose the density matrix as

$$\hat{\rho}_{\boldsymbol{\theta}} = \sum_b \Pi^b_{\boldsymbol{\theta}} \hat{M}^b, \tag{4.66}$$

where $\{\hat{M}^b\}$ is a set of tomographically complete measurement operators, with 4^L elements. A simple choice for this is the Pauli set, composed of strings of Pauli operators, i.e. $\{\hat{M}^b\} = \{\mathbb{1}, \hat{\sigma}^x, \hat{\sigma}^y, \hat{\sigma}^z\}^{\otimes L}$. $\Pi^b_{\boldsymbol{\theta}} = \mathrm{Tr}\{\hat{\rho}_{\boldsymbol{\theta}} \hat{M}^b\}$ is the probability for the operator string to be measured. As such, $\Pi_{\boldsymbol{\theta}}$ is a parametrized probability distribution over 4^L outcomes, and can be parametrized by a neural network, with a freedom on the choice of architecture. In [214] and [230], the authors use a Transformer architecture and a recurrent neural network, respectively, to represent $\Pi_{\boldsymbol{\theta}}$.
Violating the positivity of the ansatz, although unphysical, has not proved to be such an issue in practical applications, much like for MPOs. However, we expect such approaches to work best for systems with strong dissipation. Consider a singe qubit system, whose state is described by a density matrix $\hat{\rho}$ which lives inside a Bloch sphere. If dissipation is strong, the state will tend to occupy central regions of the Bloch sphere, thus remaining in the space of valid quantum states. For states close to pure, thus living on the edge of the Bloch sphere, if positivity is not enforced the reconstructed state from $\Pi_{\boldsymbol{\theta}}$ may reside outside the Bloch sphere, i.e. outside the space of physical quantum states.

VII.2 Combining the dynamical corner-space method with NQS

As a perspective to be further investigated in the future, an alternative to the POVM and Gram-Hadamard matrix approaches is to employ the low-rank representation of the density matrix detailed in chapter 3. By parametrizing each state in the corner-space by a NQS and propagating the corner in time, that is now parametrized by a set of NQS, the time-evolution of the open quantum system is recovered. This approach is particularly suited to weakly dissipative systems, thus being complementary to POVM approaches. Its advantage with respect to Gram-Hadamard approaches is that it bears no constraint on the form of the ansatz, hence the entanglement structure of the final density matrix is the same as that of the states composing it. In this representation, the corner $C(t)$ at each time t fully specifies the density matrix, and is defined by its elements as $C_{sk} = \sqrt{p_k}\psi_{\boldsymbol{\theta}^k}(\boldsymbol{s})$, with $\boldsymbol{\theta}^k$ the weights vector of the network that parametrises the k-th corner state. In the NQS language, the algorithm consists in four important steps:

(i) Evaluate the $M(D+1)$ states generated by the Kraus operators $\hat{K}_i$ acting on the M states of $\hat{C}$. The state generated by Kraus operator $\hat{K}_0$ can be obtained as

for closed systems, i.e. using t-VMC with a non-Hermitian Hamiltonian or an optimisation scheme as presented in the previous section. This step consists of $O(MD)$ optimizations. This leads to $M(D+1)$ states $\left\{\left|\phi_{\boldsymbol{\eta}^l}\right\rangle\right\}_{l=1}^{M(D+1)}$ with updated weights $\boldsymbol{\eta}^l$ (that form the transition basis).

(ii) Probabilities of each new state are obtained by using probability conservation; the probability π_l of each state $\left|\phi_{\boldsymbol{\eta}^l}\right\rangle$ of the initial corner basis has to lead to $D+1$ states whose probabilities sum to p_l.

(iii) Now that the parameters $\boldsymbol{\eta}^l$ represents the $M(D+1)$ states of the transition basis and π_l their respective probabilities, one can construct $\hat{T}^\dagger\hat{T}$ through a sampling procedure, diagonalize, and truncate the eigenvalues below ϵ to obtain the new probabilities $\tilde{p}_m$ that will correspond to $\tilde{M}$ new states.

(iv) Obtain the new corner basis $\tilde{C}$ from $\hat{T}$. In the neural network picture, obtain the parameters $\tilde{\boldsymbol{\theta}}^m$ that correctly represent the states of $\tilde{C}$ from the eigenvalues and eigenvectors of $\hat{T}^\dagger\hat{T}$.

Below each step is detailed. Note that a quantum trajectory approach is also possible, where a single trajectory is propagated and subject to random jumps, with the same optimization procedure as the one described below.

Evaluation of corner states generated by jump operators

This parts amounts to minimizing, at each time step t, for every jump operator $\hat{J}_i$, the distance

$$C\left(\boldsymbol{\eta}^l\right) := \mathrm{dist}\left(\left|\phi_{\boldsymbol{\eta}^l}\right\rangle, \hat{J}_i\left|\psi^k_{\boldsymbol{\theta}}\right\rangle\right), \tag{4.67}$$

with $\boldsymbol{\eta}^l$ the weights of the ansätz to optimize (once optimized, these will be the parameters of the lth state of the transition basis), and $\boldsymbol{\theta}^k$ the weights of the states of the kth state of the corner basis at time step t. For ease of notation, we drop the superscripts l and k in the calculations that follow. For normalized ansätze, the cost function is simply the fidelity. We denote a configuration. For local jump operators $\hat{J}_i = \hat{\mathbb{1}}_1 \otimes \hat{\mathbb{1}}_2 \otimes \ldots \otimes \hat{\jmath}_i \otimes \ldots \otimes \hat{\mathbb{1}}_N$ (with $\hat{\jmath}_i$ a local operator acting on a single spin) we have:

$$\begin{aligned}
C\left(\boldsymbol{\eta}\right) &= \frac{|\langle\phi_{\boldsymbol{\eta}}|\hat{J}_i|\psi_{\boldsymbol{\theta}}\rangle|^2}{\langle\phi_{\boldsymbol{\eta}}|\hat{J}_i^\dagger\hat{J}_i|\phi_{\boldsymbol{\eta}}\rangle}, \\
\langle\phi_{\boldsymbol{\eta}}|\hat{J}_i|\psi_{\boldsymbol{\theta}}\rangle &= \sum_{s,s'}\phi^*_{\boldsymbol{\eta}}(\boldsymbol{s})\psi_{\boldsymbol{\theta}^k}(s')\langle s|\hat{J}_i|s'\rangle \\
&= \sum_s |\phi_{\boldsymbol{\eta}}(\boldsymbol{s})|^2 \sum_{s'}\frac{\psi_{\boldsymbol{\theta}^k}(s')}{\phi_{\boldsymbol{\eta}}(\boldsymbol{s})}\langle s|\hat{J}_i|s'\rangle \\
&= \sum_s |\phi_{\boldsymbol{\eta}}(\boldsymbol{s})|^2 \sum_{s'_i}\frac{\psi_{\boldsymbol{\theta}}(s'_i|s\backslash s_i)}{\phi_{\boldsymbol{\eta}}(\boldsymbol{s})}\langle s_i|\hat{\jmath}_i|s'_i\rangle, \\
\langle\phi_{\boldsymbol{\eta}}|\hat{J}_i|\psi_{\boldsymbol{\theta}}\rangle &= \sum_s |\phi_{\boldsymbol{\eta}}(\boldsymbol{s})|^2 \hat{J}\psi_\phi(\boldsymbol{s}) \\
&= \langle\hat{J}\psi_\phi\rangle_{|\phi_{\boldsymbol{\eta}}|^2}
\end{aligned}$$

where we have defined

$$\hat{J}\psi_\phi(\boldsymbol{s}) = \sum_{\sigma_i'} \frac{\psi_{\boldsymbol{\theta}}(\sigma_i'|\boldsymbol{s}\backslash\sigma_i)}{\phi_{\boldsymbol{\eta}}(\boldsymbol{s})}\langle\sigma_i|\hat{j}_i|\sigma_i'\rangle,$$

and $\langle ...\rangle_{|\phi|^2}$ denotes an average over the probability distribution given by $|\phi(\boldsymbol{s})|^2$. The point here is that $\hat{J}\psi_\phi(\boldsymbol{s})$ can be evaluated exactly (in fact, for a k-local operator, it will be a sum of 2^k terms) and its average can therefore be approximated by sampling the probability distribution over which it is averaged. Therefore, one has:

$$C\left(\boldsymbol{\eta}\right) = \frac{|\langle\phi_{\boldsymbol{\eta}}|\hat{J}_i|\psi_{\boldsymbol{\theta}}\rangle|^2}{\langle\phi_{\boldsymbol{\eta}}|\hat{J}_i^\dagger\hat{J}_i|\phi_{\boldsymbol{\eta}}\rangle} = \frac{\langle\hat{J}\psi_\phi\rangle_{|\phi_{\boldsymbol{\eta}}|^2}\langle\hat{J}\phi_\psi\rangle_{|\psi_{\boldsymbol{\theta}}|^2}}{\langle\hat{J}_i^\dagger\hat{J}_i\rangle_{|\phi_{\boldsymbol{\eta}}|^2}}. \tag{4.68}$$

$C(\theta)$ can be minimized via a stochastic gradient descent, exactly like the cost functions considered in the chapter. The components of the gradient $\nabla C(\boldsymbol{\theta})$ are again given by the log-derivative trick, with $O_k = \partial_{\boldsymbol{\theta}_k}\ln\psi_{\boldsymbol{\theta}}(\boldsymbol{s})$:

$$\partial_k C\left(\boldsymbol{\eta}\right) = \left\langle\left(O_k - \langle O_k\rangle_{|\psi_{\boldsymbol{\theta}}|^2}\right)\hat{J}\phi_\psi\right\rangle_{|\psi_{\boldsymbol{\theta}}|^2}\langle\hat{J}\psi_\phi\rangle_{|\phi_{\boldsymbol{\eta}}|^2}/\langle\hat{J}_i^\dagger\hat{J}_i\rangle_{|\phi_{\boldsymbol{\eta}}|^2} \tag{4.69}$$

which has the same computational complexity as that of minimizing the energy as presented at the beginning of the chapter.

Evaluate the probabilities

After step (i), one obtains the correct parameters that represent the states $\left|\phi_{\boldsymbol{\eta}^l}\right\rangle$ of the transition basis. We can almost build $T^\dagger T$ to diagonalize it, but we still need the associated probabilities π_l. In a regular corner approach, these would naturally arise from step (i). This is not the case here, so we have to obtain them in another way from the corner basis C. Using the facts that each state in $\hat{C}$ generates $D+1$ states and that probability is conserved, we have:

$$\sum_{i=1}^{D+1}\pi_i^{(k)} = p_k,$$

and we know that the states generated by the jumps ($i \neq 1$) will have a probability $\pi_i^{(k)} = \delta t\langle\psi_{\boldsymbol{\theta}^k}|\hat{J}_i^\dagger\hat{J}_i|\psi_{\boldsymbol{\theta}^k}\rangle$. It naturally follows that $\pi_1^{(k)} = p_k - \sum_{i=2}^{D+1}\pi_i^{(k)}$, and we can therefore obtain the complete set of probabilities $\{\pi_l\}_{l=1}^{M(D+1)}$. This step is computationally for free since the averages $\langle\hat{J}_k^\dagger\hat{J}_k\rangle$ were computed at the previous step. At this step, one can discard all the probabilities that are below some truncation error ϵ, and renormalize the probability distribution given by π_l.

Evaluation of $T^\dagger T$

From the previous steps one obtains probabilities π_l and the parameters $\boldsymbol{\eta}^l$ corresponding to each state of the transition basis, $\left|\phi_{\boldsymbol{\eta}^l}\right\rangle$. Since we cannot explicitly reconstruct $T^\dagger T$,

we can sample it. Explicitly, the elements of $T^\dagger T$ are

$$[T^\dagger T]_{ll'} = \sum_s \phi^*_{\boldsymbol{\eta}^l}(\boldsymbol{s})\phi_{\boldsymbol{\eta}^{l'}}(\boldsymbol{s})\sqrt{\pi_l \pi_{l'}} \tag{4.70}$$

$$= \sum_s |\phi_{\boldsymbol{\eta}^l}(\boldsymbol{s})|^2 \frac{\phi_{\boldsymbol{\eta}^{l'}}(\boldsymbol{s})}{\phi_{\boldsymbol{\eta}^l}(\boldsymbol{s})}\sqrt{\pi_l \pi_{l'}} \tag{4.71}$$

$$= \left\langle \frac{\phi_{\boldsymbol{\eta}^{l'}}}{\phi_{\boldsymbol{\eta}^l}}\sqrt{\pi_l \pi'_l} \right\rangle_{|\phi_{\boldsymbol{\eta}^l}|^2}. \tag{4.72}$$

We once again find the form of an expectation value. A difference here is that these expectation values are complex, hence may be more difficult to estimate due to a sign problem that comes from adding phases together. However there is no optimization here so it is possible to use a very large number of samples to reconstruct $\hat{T}^\dagger\hat{T}$ with a low computational cost. After reconstructing $\hat{T}^\dagger\hat{T}$, one can diagonalize it and obtain $\tilde{M}$, the updated corner dimension, as well as $\tilde{p}_m$, the corresponding probabilities. However, we do not obtain the set of parameters $\tilde{\boldsymbol{\theta}}_m$ corresponding to these states, thus another optimization is needed. This step implies the diagonalization of an $M(D+1)$ square matrix, $\hat{T}^\dagger\hat{T}$, with no dependence on N.

Obtention of the updated corner basis $\tilde{C}$

In the original dynamical corner-space algorithm, one easily obtains $\tilde{C}$ from $\hat{T}$ by multiplying the eigenstates of $\hat{T}$ by $\hat{T}$ itself. By denoting $|t^m\rangle$ the m-th eigenvector of $\hat{T}^\dagger\hat{T}$ (which was explicitly diagonalized in step 3) and t^m_q its q-th component, the elements of $\tilde{C}$ are obtained as:

$$\tilde{C}_m = \hat{T}\,|t^m\rangle\,,$$
$$\tilde{C}_{sm} = \sum_q T_{sq} t^m_q,$$
$$|\psi_{\boldsymbol{\theta}^m}\rangle = \sum_q \frac{\sqrt{\pi_q}}{\sqrt{\tilde{p}_m}} t^m_q\,|\phi_{\boldsymbol{\eta}^q}\rangle\,.$$

By noting $|\Psi_{\text{tar}}\rangle$ the target state, one can therefore minimize

$$\text{dist}\left(|\psi_{\boldsymbol{\theta}^m}\rangle\,, |\Psi_{\text{tar}}\rangle\right) = |\langle\psi_{\boldsymbol{\theta}^m}|\Psi_{\text{tar}}\rangle|^2.$$

In analogy with the previous steps, one has:

$$\langle\psi_{\boldsymbol{\theta}^m}|\Psi_{\text{tar}}\rangle = \sum_s |\psi_{\boldsymbol{\theta}^m}(\boldsymbol{s})|^2 \frac{\Psi_{\text{tar}}(\boldsymbol{s})}{\psi_{\boldsymbol{\theta}^m}(\boldsymbol{s})} = \left\langle \frac{\Psi_{\text{tar}}}{\psi_{\boldsymbol{\theta}^m}} \right\rangle_{|\psi_{\boldsymbol{\theta}^m}|^2},$$

hence

$$\text{dist}\left(|\psi_{\boldsymbol{\theta}^m}\rangle\,, |\Psi_{\text{tar}}\rangle\right) = \left\langle \frac{\Psi_{\text{tar}}}{\psi_{\boldsymbol{\theta}^m}} \right\rangle_{|\psi_{\boldsymbol{\theta}^m}|^2} \left\langle \frac{\psi_{\boldsymbol{\theta}^m}}{\Psi_{\text{tar}}} \right\rangle_{|\Psi_{\text{tar}}|^2}. \tag{4.73}$$

Each state of the updated corner is therefore a particular superposition of the states of the transition basis for which we know the corresponding amplitudes $\frac{\sqrt{\pi_q}}{\sqrt{\tilde{p}_k}} t^m_q$. By considering normalized ansätze, we are able to evaluate the correct amplitudes $\Psi_{\text{tar}}(\boldsymbol{s})$ and evaluate this distance and its gradient.

Computational complexity

The computational complexity of such a method bears no dependence on N, the Hilbert space dimension. Combining the four steps of the algorithm, the complexity of a single time step is dominated by the first and third steps. The first step consists in MD optimizations that have a number of optimization steps n_{steps}, where a gradient of size N_{par} (the number of variaitonal parameters) is calculated by estimating averages over N_{samples} samples. The third step consists in diagonalizing a $M(D+1) \times M(D+1)$ matrix. The complexity of a single time step is therefore given by

$$O\left(\max\left\{[M(D+1)]^3, MDn_{\text{steps}}N_{\text{par}}N_{\text{samples}}\right\}\right). \tag{4.74}$$

VIII Conclusion

We have seen how neural-network quantum states presents itself as a powerful numerical toolbox to study the dynamics of complex quantum systems.
We have presented an alternative scheme for the real time evolution of quantum many-body systems with an NQS ansatz. This scheme does not rely on hyperparameter tuning for regularization and is found to be stable, and is not limited by the regime of the physical system. We therefore expect it to be useful when stochastic reconfiguration fails, when one uses autoregressive models or for when the network contains a large number of variational parameters. This scheme could be used to further investigate the limits of the Kibble-Zurek mechanism in nontrivial quantum systems, such as frustrated systems in triangular lattices or in systems with multiple critical points. As other applications, quantum control and pulse optimization problems could also be considered, where accurate methods to simulate time-dependent dynamics are required, as well as the quench dynamics of bosonic systems [232]. As an outlook, the scheme could further be improved by considering different cost functions that lead to weaker optimization errors and variance. We also expect the scheme to be helpful for the dynamics of open quantum systems [188, 230, 231], where t-VMC can present similar difficulties to those exposed. Finally, we expect the various approaches for the simulation of open quantum systems to be applied to noisy quantum circuits. Indeed, ideal quantum circuits have already been shown to be simulatable by NQS more efficiently than with tensor network methods [192]. In the regime of low noise, therefore weak entropy, a hybrid dynamical corner-NQS approach would be particularly suited. We also expect the Gram-Hadamard approach to be effective for the study of dissipative phase transitions, as well as the study of an extended form of the Kibble-Zurek mechanism to dissipative critical systems.

General conclusion and perspectives

In this manuscript, we explored the nonequilibrium dynamics of closed and open quantum systems and presented some novel methods to deal with the difficult problem of simulating such dynamics, along with new physical results for various interesting problems entering this class of systems.

After explaining the key questions and the main numerical methods surrounding the topic, we investigated the dynamics of entanglement in a strongly-interacting dissipative quantum system was presented. This chapter was motivated by recent experimental progress towards large chains of dissipatively stabilized Mott insulator of photons, which open up a new playground for many-body physics, where solid-state phases can be achieved in photonic systems. We introduced a new simple protocol to monitor entanglement dynamics, where bipartite entanglement can be investigated by looking at the negativity of reduced density matrices describing pairs of distant sites. We observe a ballistic propagation of entanglement, in agreement with the quasiparticle picture provided by a closed system analysis. However, a strikingly asymmetric behaviour is observed as a function of the chosen protocol, which depends on whether we initially inject or extract a particle.

Then, we presented the *dynamical corner-space method*. This new method to simulate the dynamics of open quantum systems fills up a gap in the literature that existed for low-entropy, highly-entangled states. If the rank of the density matrix describing the system of interest is small with respect to the Hilbert space dimension, the method can outperform other methods that do not restrict entanglement, such as Monte-Carlo wave function methods. We applied this new method to the simulation of a noisy quantum Fourier transform processor by considering a piecewise continuous-time implementation of the algorithm, so that microscopic dissipative channels were taken to account as opposed to digital error models. By doing so, we were able to derive scaling laws for the error as a function of the dissipation rate. In addition, the impact of the initial state on the performance of the algorithm was unveiled, revealing a highly biased algorithm. We believe this method will enable the study of dissipatively stabilized quantum systems such as cat qubits, and error mitigation protocols that are strongly needed in the quantum information community until the era of fault-tolerant quantum computation is reached.

Finally, the general class of ansätze of neural-network quantum states (NQS) was presented. These were introduced in the context of the general task of generative modeling in standard machine learning. After presenting Monte-Carlo sampling algorithms, we introduced the various architectures used to represent quantum states. In particular,

autoregressive models, which enable for normalized quantum states and efficient sampling algorithms were exposed. We believe such architectures, that are becoming standard in state-of-the-art machine learning routines, will become a key element of the numerical physics toolbox to study complex quantum systems. We presented various issues that arise when trying to simulate quantum dynamics with NQS, as well as an alternative scheme to perform this task that we developed. This scheme allows one to study time-dependent quenches, and the transition from adiabatic to genuine nonequilibrium dynamics that occur during such quenches, as explained by the Kibble-Zurek mechanism. We also presented the issue of representing the state of an open quantum system with NQS, whose difficulty resides in preserving the positivity of the density matrix. We displayed key ingredients of a new hybrid method, which uses ideas from the dynamical corner-space method combined with NQS. We believe such a method will be particularly useful for the simulation of noisy quantum circuits, to reach sizes well beyond what was achieved using the original dynamical corner-space method and examine, for example, the entanglement dynamics and spreading that occur in noisy quantum circuits.

Very recently, a new efficient algorithm for matrix multiplications was discovered by a machine learning model, coined AlphaTensor [233]. Prior to this discovery, the last algorithm for such a task was about fifty years old. We believe such phenomena will only amplify with the increase of computing power along with the improvement of algorithms. The work presented in this thesis is part of the efforts to design and improve effective models, to ultimately achieve an era of computationally-assisted scientific discovery [234], where the task of the scientist becomes that of formulating a problem as precisely as possible, and algorithms provide solutions to these problems, as complex as they may be.

A High-order integration methods for NQS

I Generalization to higher-order methods

Let us consider some linear ordinary differential equation of the form

$$\partial_t \left|\psi(t)\right\rangle = \hat{\Phi}_t \left|\psi(t)\right\rangle, \tag{A.1}$$

whose solution $|\psi\rangle$ is discretized over a set of well-defined times $\{t_n\}_{n=0}^{N}$ such that $|\psi^{(n)}\rangle \equiv |\psi(t_n)\rangle$. Linear Runge-Kutta (RK) methods provide us with the following prescription for approximately updating $|\psi^{(n)}\rangle$:

$$|\psi^{(n+1)}\rangle = |\psi^{(n)}\rangle + h_n \sum_{i=1}^{s} b_i |\kappa_i^{(n)}\rangle, \tag{A.2}$$

$$|\kappa_i^{(n)}\rangle = \hat{\Phi}_{t_{n,i}}\Big(|\psi^{(n)}\rangle + h_n \textstyle\sum_{j=1}^{s} a_{ij} |\kappa_j^{(n)}\rangle\Big), \tag{A.3}$$

where s denotes the number of stages of the method, $h_n = t_{n+1} - t_n$, $t_{n,i} = t_n - c_i h_n$ and the coefficients $\{a_{ij}\}_{i,j}$, $\{b_i\}_i$ and $\{c_i\}_i$ are completely determined by the Butcher tableau of the considered method:

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \,. \end{array} \tag{A.4}$$

Note that the entries are nonzero only on the lowest triangular matrix for explicit methods. We will here consider methods whose Butcher coefficients satisfy either $a_{i,j>i} = 0$ (explicit) or $a_{i,j\geq i} = 0$ (implicit).

For the considered tableaus, the system of Eqs. (A.2) and (A.3) can be solved explicitly by using the following recurrence identity:

$$|\kappa_i^{(n)}\rangle = \hat{\Pi}_i^{(n)} |\psi^{(n)}\rangle + h_n \sum_{j=1}^{i-1} \hat{\Pi}_i^{(n)} a_{ij} |\kappa_j^{(n)}\rangle, \tag{A.5}$$

with

$$\hat{\Pi}_i^{(n)} = \Big[\hat{\mathbb{1}} - a_{ii} h_n \hat{\Phi}_{t_{n,i}}\Big]^{-1} \hat{\Phi}_{t_{n,i}}. \tag{A.6}$$

Indeed, one has:

$$|\kappa_i^{(n)}\rangle = \left\{\hat{\Pi}_i^{(n)} + h_n\hat{\Pi}_i^{(n)}a^{ij}\hat{\Pi}_j^{(n)} + h_n^2\hat{\Pi}_i^{(n)}a^{ij}\hat{\Pi}_j^{(n)}a^{jk}\hat{\Pi}_k^{(n)} + \dots\right\}|\psi^{(n)}\rangle, \tag{A.7}$$

and thus

$$|\psi^{(n+1)}\rangle = \hat{T}_s^{(n)}|\psi^{(n)}\rangle, \tag{A.8}$$

with

$$\hat{T}_s^{(n)} = \hat{\mathbb{1}} + \sum_{i=1}^{s} b_i\left(h_n\hat{\Pi}_i^{(n)} + h_n^2\hat{\Pi}_i^{(n)}a^{ij}\hat{\Pi}_j^{(n)} + h_n^3\hat{\Pi}_i^{(n)}a^{ij}\hat{\Pi}_j^{(n)}a^{jk}\hat{\Pi}_k^{(n)} + \dots\right). \tag{A.9}$$

Representing the wavefunctions above with a variational ansatz, namely $\left\langle\boldsymbol{\sigma}\middle|\psi_{\boldsymbol{\eta}_n}\right\rangle = \psi_{\boldsymbol{\eta}_n}(\boldsymbol{\sigma})$, the generic update can finally be recast into the following optimization process:

$$|\psi_{\boldsymbol{\eta}_{n+1}}\rangle = \arg\min_{|\psi_{\boldsymbol{\eta}}\rangle} \mathrm{dist}\left(|\psi_{\boldsymbol{\eta}}\rangle, \hat{T}_s^{(n)}|\psi_{\boldsymbol{\eta}_n}\rangle\right). \tag{A.10}$$

Time-independent explicit case

The equations above considerably simplify when considering time-independent Hamiltonians and an explicit integration method. Indeed, to any order s, we have

$$\hat{T}_s^{(n)} = \sum_{m=0}^{s-1}\lambda_m h_n^m\hat{\Pi}^{(n)m} = \sum_{m=0}^{s-1}\lambda_m(-ih_n\hat{H})^m\,, \tag{A.11}$$

with

$$\lambda_m := \begin{cases}1, & m < 2,\\ \boldsymbol{b}^T\mathbf{a}^{m-2}\boldsymbol{c}, & \text{else.}\end{cases} \tag{A.12}$$

Two common Butcher tableaus corresponding to the fourth-order Runge-Kutta method are:

0				
1/2	1/2			
1/2	0	1/2		
1	0	0	1	
	1/6	1/3	1/3	1/6

0				
1/3	1/3			
2/3	−1/3	1		
1	1	−1	1	
	1/8	3/8	3/8	1/8

(A.13)

For these, we have:

λ_0	λ_1	λ_2	λ_3
1	1	1/2	1/6

This corresponds to the usual factor $\lambda_m = 1/m!$ of the truncated Taylor expansion of the propagator. Note, however, that this is generally no longer the case in a time-dependent scenario, as will appear below.

Time-dependent implicit midpoint method

The implicit midpoint method is characterized by the tableau

$$\begin{array}{c|c} 1/2 & 1/2 \\ \hline & 1 \end{array} \tag{A.14}$$

and yields

$$\hat{T}_1^{(n)} = \hat{\mathbb{1}} + h_n b_1 \hat{\Phi}_{t_{n,1}} + h_n^2 a_{11}^2 \hat{\Phi}_{t_{n,1}}^2 = \hat{\mathbb{1}} - i h_n \hat{H}(t_n + h_n/2) - \frac{h_n^2}{4} \hat{H}^2(t_n + h_n/2). \tag{A.15}$$

Time-dependent Heun method

The Butcher tableau of this explicit second-order Runge-Kutta method reads

$$\begin{array}{c|cc} 0 & & \\ 1 & 1 & \\ \hline & 1/2 & 1/2 \end{array} \tag{A.16}$$

and yields the following propagator:

$$\begin{aligned} \hat{T}_2^{(n)} &= \hat{\mathbb{1}} + h_n b_1 \hat{\Phi}_{t_{n,1}} + h_n b_2 \hat{\Phi}_{t_{n,2}} + b_2 a_{21} h_n^2 \hat{\Phi}_{t_{n,2}} \hat{\Phi}_{t_{n,1}} \\ &= \hat{\mathbb{1}} - i h_n \frac{\hat{H}(t_n) + \hat{H}(t_n + h_n)}{2} - \frac{h_n^2}{2} \hat{H}(t_n + h_n) \hat{H}(t_n). \end{aligned} \tag{A.17}$$

Time-dependent Ralston method

The Butcher tableau of this explicit second-order Runge-Kutta method reads

$$\begin{array}{c|cc} 0 & & \\ 2/3 & 2/3 & \\ \hline & 1/4 & 3/4 \end{array} \tag{A.18}$$

and yields the following propagator:

$$\begin{aligned} \hat{T}_2^{(n)} &= \hat{\mathbb{1}} + h_n b_1 \hat{\Phi}_{t_{n,1}} + h_n b_2 \hat{\Phi}_{t_{n,2}} + b_2 a_{21} h_n^2 \hat{\Phi}_{t_{n,2}} \hat{\Phi}_{t_{n,1}} \\ &= \hat{\mathbb{1}} - i h_n \frac{\hat{H}(t_n) + 3\hat{H}(t_n + h_n)}{4} - \frac{h_n^2}{2} \hat{H}(t_n + h_n) \hat{H}(t_n). \end{aligned} \tag{A.19}$$

B Résumé substantiel

Cette thèse est consacrée à l'étude de la dynamique hors équilibre dans les systèmes quantiques idéaux (fermés, isolés de leur environnement direct) et dissipatifs (ouverts, couplés à leur environnement direct). Après une présentation générale de questions liées à la dynamique de systèmes physique ainsi que de leur Histoire, une introduction détaillée du sujet ainsi qu'une exposition des diverses difficultés apparaissant lors de son étude sont faites dans le chapitre 1. Ensuite, un premier objet d'étude, la dynamique de l'intrication dans des systèmes photoniques dissipatifs fortement corrélés, est exposé dans le chapitre 2 ainsi que les divers nouveaux résultats physiques obtenus par une approche de réseaux de tenseurs. Puis, la méthode du sous-espace dynamique, méthode générale pour la dynamique de systèmes ouverts à entropie faible est présentée dans le chapitre 3. Son application à l'algorithme de la transformée de Fourier quantique est présentée, encore une fois accompagné de nouveaux résultats obtenus grâce à cette nouvelle méthode. Enfin, la simulation de la dynamique de systèmes isolés grâce à une représentation en états quantiques neuronaux est détaillée, ainsi qu'une nouvelle méthode développée pour obtenir des résultats plus fiables pour les systèmes avec un Hamiltonien dépendant du temps sont présentés dans le chapitre 4. Divers résultats concernant le mécanisme de Kibble-Zurek dans des systèmes d'Ising avec champ transverse dans la région critique ont ainsi pu être obtenus.

Un système photonique fortement corrélé composé de L résonateurs électromagnétiques couplés peut être décrit par l'Hamiltonien suivant

$$\hat{H} = \sum_{i=1}^{L} \left(\omega_c b_i^\dagger b_i + \frac{U}{2} b_i^\dagger b_i^\dagger b_i b_i \right) - J \sum_{i=1}^{L-1} (b_i^\dagger b_{i+1} + h.c.), \tag{B.1}$$

avec ω_c la fréquence du mode de la cavité, U la force d'interaction photon-photon (Kerr), J le couplage entre sites voisins, et b_i ($b_i^\dagger$) l'opérateur d'annihiliation (création) sur chaque site. Un tel système présente une transition de Mott entre un système fortement corrélé, lorsque $U \gg J$ et une phase superfluide dans le régime à interactions faibles $U \ll J$). Lorsque les résonateurs sont dissipatifs, le système peut être décrit par l'équation de Lindblad suivante, sous les approximations de Born et Markov :

$$\frac{d\rho}{dt} = -\mathrm{i}[\hat{H}, \rho] + \frac{1}{2} \sum_{i=1}^{L} \sum_{\mathcal{C}} 2 J_i^{(\mathcal{C})} \rho J_i^{(\mathcal{C})\dagger} - \{ J_i^{(\mathcal{C})\dagger} J_i^{(\mathcal{C})}, \rho \}, \tag{B.2}$$

avec $J_i^{(\mathcal{C})}$ l'opérateur de saut pour le site i et le canal de dissipation $\mathcal{C}$. Nous considérons deux canaux : le canal de dissipation d'énergie, et le canal de déphasage. En utilisant une

approche par réseaux de tenseurs (MPO), nous avons obtenus de nombreux résultats sur la dynamique de l'intrication. Nous considérons deux cas : dans le premier, nous ajoutons une particule au site central, au dessus d'un état type Mott fortement corrélé. Dans le deuxième cas, nous considérons le retrait d'une particule (donc l'ajout d'un trou) sur le site central. Nous pouvons estimer l'intrication grâce à la négativité, qui est une mesure de l'intrication bipartite. Nous pouvons donc quantifier l'intrication entre deux sites i et j. En faisant une trace partielle sur l'ensemble du système privé des sites i, j, nous obtenons une matrice densité réduite, pour laquelle nous calculons la négativité. Ainsi, nous obtenons que la propagation de l'intrication est ballistique aux temps courts, et est correctement prédite par les calculs pour le système fermé. Cela indique que l'image de quasiparticules reste pertinente même lorsque la dissipation est forte. Remarquablement, la dissipation affecte beaucoup plus la propagation de l'intrication dans le cas du retrait d'une particule au temps initial. Cela s'explique par une image de trajectoires quantiques, qui rend compte de la fragilité de la propagation du trou.

Le chapitre 3 est dédié à la simulation numérique de systèmes quantiques ouverts en général, et en particulier des systèmes ayant une entropie limitée. Ce type de systèmes est celui généralement construit pour les tâches de traitement de l'information quantique, ou les systèmes sont isolés le plus possible de leur environnement direct. Cette technique est basée sur la décomposition de la solution de l'équation maîtresse $\hat{\rho}(t)$ sous la forme, pour tout instant t :

$$\hat{\rho}(t) \simeq \sum_{k=1}^{M(t)} p_k(t) \left|\phi_k(t)\right\rangle\!\left\langle\phi_k(t)\right|, \quad p_k(t) \geq p_{k+1}(t),\ \forall k, \tag{B.3}$$

où les probabilités $\{p_k(t)\}_k$ sont les $M(t)$ valeurs propres les plus élevées à l'instant t et $\{|\phi(t)\rangle_k\}_k$ sont leurs vecteurs propres associés. Lorsque l'entropie du système reste faible au cours de son évolution temporelle, cette décomposition peut être tronquée à un nombre de composantes M bien plus faible que la dimension de l'espace de Hilbert associé au système. La méthode du sous-espace dynamique permet ainsi de limiter l'évolution temporelle à ce nombre très limité de composantes sans jamais reconstruire la matrice densité complète. S'ensuit un énorme avantage sur le plan de la complexité numérique de la simulation. À titre d'exemple, pour un ensemble de L spins 1/2, la complexité est réduite de $O(2^L \times 2^L)$ à $O(M \times 2^L)$. Cette méthode est appliquée à un algorithme réalisant la transformée de Fourier quantique en présence de dissipation ou de déphasage. Cet algorithme est au centre de nombreux algorithmes quantiques, comme l'algorithme de Shor, et en comprendre les mécanismes d'erreur avec une modélisation réaliste est ainsi de grande importance. Nos simulations montrent que l'erreur introduite par le couplage du circuit à un environnement extérieur croît quadratiquement avec la taille du système L. Nous montrons aussi que l'erreur dépend sensiblement du type d'état à partir duquel l'opération est menée. Nous envisageons aussi d'étendre ce type de méthode afin de pouvoir utiliser des états variationnels comme états composant le sous-espace, ce qui en augmentera considérablement sa compléxité numérique.

Enfin, nous présentons le cadre théorique de la représentation d'états quantiques par des réseaux de neurones. Pour une fonction d'onde ψ dans un espace de Hilbert discret,

l'amplitude de chacune des configurations s peut être représenté par un réseau de neurone, auquel l'on donne en entrée la configuration, et en sortie nous obtenons $\psi(s)$. L'application canonique d'une telle méthode est trouver l'état fondamental d'un système décrit par un Hamiltonien $\hat{H}$. En minimisant la quantité

$$\frac{\langle\psi_\theta|\hat{H}|\psi_\theta\rangle}{\langle\psi_\theta|\psi_\theta\rangle} \tag{B.4}$$

sur l'ensemble des valeurs possible de θ, nous pouvons obtenir une estimation de l'état fondamental et de son énergie, en vertu du principe variationnel. En utilisant des méthodes standard d'apprentissage profond, cette quantité peut être minimisée et efficacement, et ainsi ces méthodes deviennent très utiles pour les problèmes fortement corrélés, ou les autres méthodes numériques à notre disposition sont très limitées. En particulier, nous étudions le mécanisme de Kibble-Zurek pour des systèmes d'Ising avec champ transverse. Ce système présente une transition de phase entre une phase paramagnétique et ferromagnétique (ou antiferromagnétique). Proche de la transition de phase, il est notoirement difficile d'étudier ce système, notamment lorsque l'Hamiltonien dépend du temps comme lors de dynamique type Kibble-Zurek. En considérant une intégration numérique alternative avec des états quantiques neuronaux, nous sommes capable d'obtenir de résultats pour ce mécanisme cohérent avec d'autres méthodes numériques, pour ainsi corroborer de tels résultats. Nous montrons aussi les limites d'autres méthodes d'intégration numériques pour les états neuronaux, comme les méthodes de Monte Carlo variationnelles en temps (t-VMC). Ces méthodes pourront donc être utilisées à l'avenir pour investiguer ce mécanisme dans des systèmes plus complexes, voire dissipatifs.

Bibliography

[1] K. Donatella, A. Biella, A. Le Boité and C. Ciuti, "Entanglement dynamics in dissipative photonic Mott insulators", Phys. Rev. Research **2**, 043232 (2020).

[2] K. Donatella, Z. Denis, A. Le Boité and C. Ciuti, "Continuous-time dynamics and error scaling of noisy highly entangling quantum circuits", Phys. Rev. A **104**, 062407, 10.1103/PhysRevA.104.062407 (2021).

[3] K. Donatella, Z. Denis, A. Le Boité and C. Ciuti, "Dynamics with autoregressive neural quantum states: application to critical quench dynamics", arXiv:2209.03241 (2022).

[4] Z. Li, V. Heyraud, K. Donatella, Z. Denis and C. Ciuti, "Machine learning via relativity-inspired quantum dynamics", Phys. Rev. A **106**, 032413 (2022).

[5] G. Galilei, *Systema cosmicum* (Elsevier, 1635).

[6] M. Gardner, "Mathematical games", Scientific American **223**, 120–123 (1970).

[7] P. W. Anderson, "More is different", Science **177**, 393–396 (1972).

[8] S. P. Klevansky, "The Nambu-Jona-Lasinio model of quantum chromodynamics", Rev. Mod. Phys. **64**, 649–708 (1992).

[9] D. R. Durran, *Numerical methods for fluid dynamics* (Springer New York, 2010).

[10] R. Orús, "A practical introduction to tensor networks: matrix product states and projected entangled pair states", Annals of Physics **349**, 117–158 (2014).

[11] J. Eisert, M. Cramer and M. B. Plenio, "Colloquium: area laws for the entanglement entropy", Reviews of Modern Physics **82**, 277–306 (2010).

[12] R. O. Jones, "Density functional theory: its origins, rise to prominence, and future", Rev. Mod. Phys. **87**, 897–923 (2015).

[13] F. Minganti, A. Biella, N. Bartolo and C. Ciuti, "Spectral theory of Liouvillians for dissipative phase transitions", Phys. Rev. A **98**, 042118 (2018).

[14] M. Cheneau, "Experimental tests of Lieb-Robinson bounds", arXiv:2206.15126 (2022).

[15] E. Guardado-Sanchez, P. T. Brown, D. Mitra, T. Devakul, D. A. Huse, P. Schauß and W. S. Bakr, "Probing the quench dynamics of antiferromagnetic correlations in a 2D quantum Ising spin system", Phys. Rev. X **8**, 021069 (2018).

[16] A. Keesling, A. Omran, H. Levine, H. Bernien, H. Pichler, S. Choi, R. Samajdar, S. Schwartz, P. Silvi, S. Sachdev, P. Zoller, M. Endres, M. Greiner, V. Vuletić and M. D. Lukin, "Quantum Kibble–Zurek mechanism and critical dynamics on a programmable rydberg simulator", Nature **568**, 207–211 (2019).

[17] H. P. Breuer and F. Petruccione, *The theory of open quantum systems* (Oxford University Press, Great Clarendon Street, 2002).

[18] L. S. Bishop, "Circuit quantum electrodynamics", PhD thesis (Yale University, Connecticut, Jan. 2010).

[19] M. Cheneau, P. Barmettler, D. Poletti, M. Endres, P. Schauß, T. Fukuhara, C. Gross, I. Bloch, C. Kollath and S. Kuhr, "Light-cone-like spreading of correlations in a quantum many-body system", Nature **481**, 484–487 (2012).

[20] V. Alba and F. Carollo, "Spreading of correlations in Markovian open quantum systems", Phys. Rev. B **103**, L020302, L020302 (2021).

[21] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto and L. Zdeborová, "Machine learning and the physical sciences", Rev. Mod. Phys. **91**, 045002 (2019).

[22] A. Dawid, J. Arnold, B. Requena, A. Gresch, M. Płodzień, K. Donatella, K. A. Nicoli, P. Stornati, R. Koch, M. Büttner, R. Okuła, G. Muñoz-Gil, R. A. Vargas-Hernández, A. Cervera-Lierta, J. Carrasquilla, V. Dunjko, M. Gabrié, P. Huembeli, E. van Nieuwenburg, F. Vicentini, L. Wang, S. J. Wetzel, G. Carleo, E. Greplová, R. Krems, F. Marquardt, M. Tomza, M. Lewenstein and A. Dauphin, "Modern applications of machine learning in quantum sciences", arXiv:2204.04198 (2022).

[23] P. A. M. Dirac and R. H. Fowler, "Quantum mechanics of many-electron systems", Proc. R. Soc. A: Math. Phys. Eng. Sci. **123**, 714–733 (1929).

[24] M. Srednicki, "Chaos and quantum thermalization", Phys. Rev. E **50**, 888–901 (1994).

[25] T. W. B. Kibble, "Topology of cosmic domains and strings", Journal of Physics A: Mathematical and General **9**, 1387–1398 (1976).

[26] W. H. Zurek, "Cosmological experiments in superfluid helium?", Nature **317**, 505–508 (1985).

[27] A. Mitra, "Quantum quench dynamics", Annual Review of Condensed Matter Physics **9**, 245–259 (2018).

[28] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, 2010).

[29] M. Carrega, A. Crescente, D. Ferraro and M. Sassetti, "Dissipative dynamics of an open quantum battery", New Journal of Physics **22**, 083085 (2020).

[30] J. Preskill, "Quantum Computing in the NISQ era and beyond", Quantum **2**, 79 (2018).

[31] R. Kubo, "The fluctuation-dissipation theorem", Reports on Progress in Physics **29**, 255–284 (1966).

[32] E. H. Lieb and D. W. Robinson, "The finite group velocity of quantum spin systems", Communications in Mathematical Physics **28**, 251–257 (1972).

[33] M. Cheneau, P. Barmettler, D. Poletti, M. Endres, P. Schauss, T. Fukuhara, C. Gross, I. Bloch, C. Kollath and S. Kuhr, "Light-cone-like spreading of correlations in a quantum many-body system", Nature **481**, 484–487 (2012).

[34] P. Calabrese and J. Cardy, "Time dependence of correlation functions following a quantum quench", Phys. Rev. Lett. **96**, 136801 (2006).

[35] H. W. J. Blöte and Y. Deng, "Cluster Monte-Carlo simulation of the transverse ising model", Phys. Rev. E **66**, 066110 (2002).

[36] M. C. Gutzwiller, "Effect of correlation on the ferromagnetism of transition metals", Phys. Rev. Lett. **10**, 159–162 (1963).

[37] J. Hubbard, "Electron correlations in narrow energy bands", Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **276**, 238–257 (1963).

[38] J. Kanamori, "Electron correlation and ferromagnetism of transition metals", Progress of Theoretical Physics **30**, 275–289 (1963).

[39] G. Borghi, M. Fabrizio and E. Tosatti, "Self-consistent Gutzwiller study of bcc–Fe: interplay of ferromagnetic order and kinetic energy", arXiv:1307.5738 (2013).

[40] D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner and P. Zoller, "Cold bosonic atoms in optical lattices", Phys. Rev. Lett. **81**, 3108–3111 (1998).

[41] T. McIntosh, P. Pisarski, R. J. Gooding and E. Zaremba, "Multisite mean-field theory for cold bosonic atoms in optical lattices", Phys. Rev. A **86**, 013623, 013623 (2012).

[42] J. Eisert, M. Cramer and M. B. Plenio, "Icolloquium/i: area laws for the entanglement entropy", Reviews of Modern Physics **82**, 277–306 (2010).

[43] I. M. Georgescu, S. Ashhab and F. Nori, "Quantum simulation", Rev. Mod. Phys. **86**, 153–185 (2014).

[44] R. P. Feynman, "Simulating physics with computers", International Journal of Theoretical Physics **21**, 467–488 (1982).

[45] A. Browaeys and T. Lahaye, "Many-body physics with individually controlled rydberg atoms", Nature Physics **16**, 132–142 (2020).

[46] V. Lienhard, S. de Léséleuc, D. Barredo, T. Lahaye, A. Browaeys, M. Schuler, L.-P. Henry and A. M. Läuchli, "Observing the space- and time-dependent growth of correlations in dynamically tuned synthetic ising models with antiferromagnetic interactions", Phys. Rev. X **8**, 021070 (2018).

[47] J.-M. Cui, F. J. Gómez-Ruiz, Y.-F. Huang, C.-F. Li, G.-C. Guo and A. del Campo, "Experimentally testing quantum critical dynamics beyond the Kibble–Zurek mechanism", Communications Physics **3** (2020).

[48] B. Damski and W. H. Zurek, "Dynamics of a quantum phase transition in a ferromagnetic Bose-Einstein condensate", Phys. Rev. Lett. **99**, 130402 (2007).

[49] S. Higuera-Quintero, F. J. Rodriguez, L. Quiroga and F. J. Gómez-Ruiz, "Experimental validation of the Kibble-Zurek Mechanism on a Digital Quantum Computer", arXiv:2208.01050 (2022).

[50] K. Kraus, A. Böhm, J. D. Dollard and W. H. Wootters, eds., *States, effects, and operations fundamental notions of quantum theory* (Springer Berlin Heidelberg, 1983).

[51] H. Carmichael, *An open systems approach to quantum optics*, Vol. 18, Lecture notes in physics (Springer, Berlin, 1993).

[52] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2009).

[53] K. Mølmer, Y. Castin and J. Dalibard, "Monte-Carlo wave-function method in quantum optics", J. Opt. Soc. Am. B **10**, 524–538 (1993).

[54] S. Krämer, D. Plankensteiner, L. Ostermann and H. Ritsch, "QuantumOptics.jl: a julia framework for simulating open quantum systems", Computer Physics Communications **227**, 109–116 (2018).

[55] A. J. Daley, "Quantum trajectories and open many-body quantum systems", Advances in Physics **63**, 77–149 (2014).

[56] J. Jin, A. Biella, O. Viyuela, L. Mazza, J. Keeling, R. Fazio and D. Rossini, "Cluster mean-field approach to the steady-state phase diagram of dissipative spin systems", Phys. Rev. X **6**, 031011 (2016).

[57] A. Le Boité, G. Orso and C. Ciuti, "Steady-state phases and tunneling-induced instabilities in the driven dissipative Bose-Hubbard model", Phys. Rev. Lett. **110**, 233601 (2013).

[58] M.-D. Choi, "Completely positive linear maps on complex matrices", Linear Algebra and its Applications **10**, 285–290 (1975).

[59] M. Kliesch, D. Gross and J. Eisert, "Matrix-product operators and states: np-hardness and undecidability", Phys. Rev. Lett. **113**, 160503 (2014).

[60] A. H. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert and S. Montangero, "Positive tensor network approach for simulating open quantum many-body systems", Phys. Rev. Lett. **116**, 237201 (2016).

[61] G. D. las Cuevas, N. Schuch, D. Pé rez-Garcia and J. I. Cirac, "Purifications of multipartite states: limitations and constructive methods", New Journal of Physics **15**, 123021 (2013).

[62] G. D. las Cuevas, T. S. Cubitt, J. I. Cirac, M. M. Wolf and D. Pé rez-Garcia, "Fundamental limitations in the purifications of tensor networks", Journal of Mathematical Physics **57**, 071902 (2016).

[63] A. Biella, L. Mazza, I. Carusotto, D. Rossini and R. Fazio, "Photon transport in a dissipative chain of nonlinear cavities", Phys. Rev. A **91** (2015).

[64] J. Cui, J. I. Cirac and M. C. Bañuls, "Variational matrix product operators for the steady state of dissipative quantum systems", Phys. Rev. Lett. **114**, 220601 (2015).

[65] S. Finazzi, A. Le Boité, F. Storme, A. Baksic and C. Ciuti, "Corner-space renormalization method for driven-dissipative two-dimensional correlated systems", Phys. Rev. Lett. **115**, 080604 (2015).

[66] G. Wendin, "Quantum information processing with superconducting circuits: a review", Reports on Progress in Physics **80**, 106001 (2017).

[67] A. Blais, S. M. Girvin and W. D. Oliver, "Quantum information processing and quantum optics with circuit quantum electrodynamics", Nature Physics **16**, 247–256 (2020).

[68] B. Josephson, "Possible new effects in superconductive tunnelling", Physics Letters **1**, 251–253 (1962).

[69] R. Ma, B. Saxberg, C. Owens, N. Leung, Y. Lu, J. Simon and D. I. Schuster, "A dissipatively stabilized Mott insulator of photons", Nature **566**, 51–57 (2019).

[70] E. Schrödinger, "Die gegenwärtige situation in der quantenmechanik", Naturwiss. **23**, 807 (1935).

[71] A. Einstein, B. Podolsky and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?", Phys. Rev **47**, 777 (1935).

[72] R. Horodecki, P. Horodecki, M. Horodecki and K. Horodecki, "Quantum entanglement", Rev. Mod. Phys. **81**, 865 (2009).

[73] O. Gühne and G. Tóth, "Entanglement detection", Phys. Rep. **474**, 1–75 (2009).

[74] M. B. Plenio and S. Virmani, "An introduction to entanglement measures", Quant. Inf. Comp. **7**, 1–51 (2007).

[75] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin and W. K. Wootters, "Purification of noisy entanglement and faithful teleportation via noisy channels", Phys. Rev. Lett. **76**, 722–725 (1996).

[76] V. Giovannetti, S. Lloyd and L. Maccone, "Advances in quantum metrology", Nat. Phot. **5**, 222–229 (2011).

[77] N. Gisin, G. Ribordy, W. Tittel and H. Zbinden, "Quantum cryptography", Rev. Mod. Phys. **74**, 145 (2002).

[78] J. Preskill, "Quantum information and physics: some future directions", J. Mod. Opt. **47**, 127 (2000).

[79] L. Amico, R. Fazio, A. Osterloh and V. Vedral, "Entanglement in many-body systems", Rev. Mod. Phys. **80**, 517 (2008).

[80] G. De Chiara and A. Sanpera, "Genuine quantum correlations in quantum many-body systems: a review of recent progress", Rep. Prog. Phys. **81**, 074002 (2018).

[81] R. Islam, R. Ma, P. M. Preiss, M. E. Tai, A. Lukin, M. Rispoli and M. Greiner, "Measuring entanglement entropy in a quantum many-body system", Nature **528**, 77 (2015).

[82] L. Amico, A. Osterloh, F. Plastina, R. Fazio and G. Massimo Palma, "Dynamics of entanglement in one-dimensional spin systems", Phys. Rev. A **69**, 022304 (2004).

[83] P. Jurcevic, B. P. Lanyon, P. Hauke, C. Hempel, P. Zoller, R. Blatt and C. F. Roos, "Quasiparticle engineering and entanglement propagation in a quantum many-body system", Nature **511**, 202 (2014).

[84] M. Kastner, "Entanglement-enhanced spreading of correlations", New J. Phys. **17**, 123024 (2015).

[85] A. M. Läuchli and C. Kollath, "Spreading of correlations and entanglement after a quench in the one-dimensional Bose–Hubbard model", J. Stat. Mech. **2008**, P05018 (2008).

[86] A. Daley, H. Pichler, J. Schachenmayer and P. Zoller, "Measuring entanglement growth in quench dynamics of bosons in an optical lattice", Phys. Rev. Lett. **109**, 020505 (2012).

[87] J.-S. Bernier, R. Tan, L. Bonnes, C. Guo, D. Poletti and C. Kollath, "Light-Cone and Diffusive Propagation of Correlations in a Many-Body Dissipative System", Phys. Rev. Lett. **120**, 020401 (2018).

[88] L. Aolita, F. de Melo and L. Davidovich, "Open-system dynamics of entanglement:a key issues review", Reports on Progress in Physics **78**, 042001 (2015).

[89] Z. Leghtas, S. Touzard, I. M. Pop, A. Kou, B. Vlastakis, A. Petrenko, K. M. Sliwa, A. Narla, S. Shankar, M. J. Hatridge, M. Reagor, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi and M. H. Devoret, "Confining the state of light to a quantum manifold by engineered two-photon loss", Science **347**, 853–857 (2015).

[90] I. Carusotto and C. Ciuti, "Quantum fluids of light", Rev. Mod. Phys. **85**, 299 (2013).

[91] S. Schmidt and J. Koch, "Circuit qed lattices: towards quantum simulation with superconducting circuits", Ann. Phys. **525**, 395 (2013).

[92] D. Roy, C. M. Wilson and O. Firstenberg, "Colloquium: strongly interacting photons in one-dimensional continuum", Rev. Mod. Phys. **89**, 021001 (2017).

[93] A. Biella, L. Mazza, I. Carusotto, D. Rossini and R. Fazio, "Photon transport in a dissipative chain of nonlinear cavities", Phys. Rev. A **91**, 053815 (2015).

[94] C. Lee, C. Noh, N. Schetakis and D. G. Angelakis, "Few-photon transport in many-body photonic systems: a scattering approach", Phys. Rev. A **92**, 063817 (2015).

[95] T. Mertz, I. Vasić, M. J. Hartmann and W. Hofstetter, "Photonic currents in driven and dissipative resonator lattices", Phys. Rev. A **94**, 013809 (2016).

[96] K. Debnath, E. Mascarenhas and V. Savona, "Nonequilibrium photonic transport and phase transition in an array of optical cavities", New Journal of Physics **19**, 115006 (2017).

[97] S. Maity, S. Bandyopadhyay, S. Bhattacharjee and A. Dutta, "Growth of mutual information in a quenched one-dimensional open quantum many-body system", Phys. Rev. B **101**, 180301 (2020).

[98] A. A. Houck, H. E. Türeci and J. Koch, "On-chip quantum simulation with superconducting circuits", Nature Physics **8**, 292–299 (2012).

[99] I. Carusotto and C. Ciuti, "Quantum fluids of light", Rev. Mod. Phys. **85**, 299–366 (2013).

[100] M. Fitzpatrick, N. M. Sundaresan, A. C. Y. Li, J. Koch and A. A. Houck, "Observation of a dissipative phase transition in a one-dimensional circuit qed lattice", Phys. Rev. X **7**, 011016 (2017).

[101] M. J. Hartmann, "Quantum simulation with interacting photons", Journal of Optics **18**, 104005 (2016).

[102] C. Noh and D. G. Angelakis, "Quantum simulations and many-body physics with light", Reports on Progress in Physics **80**, 016401 (2016).

[103] I. Carusotto, A. A. Houck, A. J. Kollár, P. Roushan, D. I. Schuster and J. Simon, "Photonic materials in circuit quantum electrodynamics", Nature Physics **16**, 268–279 (2020).

[104] V. Goblot, B. Rauer, F. Vicentini, A. Le Boité, E. Galopin, A. Lemaitre, L. Le Gratiet, A. Harouri, I. Sagnes, S. Ravets, C. Ciuti, A. Amo and J. Bloch, "Nonlinear polariton fluids in a flatband reveal discrete gap solitons", Phys. Rev. Lett. **123**, 113901 (2019).

[105] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch and I. Bloch, "Quantum phase transition from a superfluid to a mott insulator in a gas of ultracold atoms", Nature **415**, 39–44 (2002).

[106] D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner and P. Zoller, "Cold bosonic atoms in optical lattices", Phys. Rev. Lett. **81**, 3108–3111 (1998).

[107] H. Pichler, A. J. Daley and P. Zoller, "Nonequilibrium dynamics of bosonic atoms in optical lattices: decoherence of many-body states due to spontaneous emission", Phys. Rev. A **82**, 063605 (2010).

[108] M. J. Hartmann, "Quantum simulation with interacting photons", Journal of Optics **18**, 104005 (2016).

[109] M. Girardeau, "Relationship between systems of impenetrable bosons and fermions in one dimension", Journal of Mathematical Physics **1**, 516–523 (1960).

[110] B. Paredes, A. Widera, V. Murg, O. Mandel, S. Fölling, I. Cirac, G. V. Shlyapnikov, T. W. Hänsch and I. Bloch, "Tonks–girardeau gas of ultracold atoms in an optical lattice", Nature **429**, 277–281 (2004).

[111] K. Życzkowski, P. Horodecki, A. Sanpera and M. Lewenstein, "Volume of the set of separable states", Phys. Rev. A **58**, 883–892 (1998).

[112] R. T. Thew, K. Nemoto, A. G. White and W. J. Munro, "Qudit quantum-state tomography", Phys. Rev. A **66**, 012303 (2002).

[113] J.-C. Besse, S. Gasparinetti, M. C. Collodo, T. Walter, A. Remm, J. Krause, C. Eichler and A. Wallraff, "Parity detection of propagating microwave fields", Phys. Rev. X **10**, 011046 (2020).

[114] M. Zwolak and G. Vidal, "Mixed-state dynamics in one-dimensional quantum lattice systems: a time-dependent superoperator renormalization algorithm", Phys. Rev. Lett. **93**, 207205 (2004).

[115] F. Verstraete, J. J. Garcia-Ripoll and J. I. Cirac, "Matrix product density operators: simulation of finite-temperature and dissipative systems", Phys. Rev. Lett. **93**, 207204 (2004).

[116] B. Bellomo, R. Lo Franco and G. Compagno, "Non-markovian effects on the dynamics of entanglement", Phys. Rev. Lett. **99**, 160502 (2007).

[117] Z. Ficek and R. Tana ś, "Dark periods and revivals of entanglement in a two-qubit system", Phys. Rev. A **74**, 024304 (2006).

[118] F. Galve, G. L. Giorgi and R. Zambrini, "Entanglement dynamics of nonidentical oscillators under decohering environments", Phys. Rev. A **81**, 062117 (2010).

[119] R. Horodecki, P. Horodecki, M. Horodecki and K. Horodecki, "Quantum entanglement", Rev. Mod. Phys. **81**, 865–942 (2009).

[120] G. L. Giorgi, B. Bellomo, F. Galve and R. Zambrini, "Genuine quantum and classical correlations in multipartite systems", Phys. Rev. Lett. **107**, 190501 (2011).

[121] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen et al., "Quantum supremacy using a programmable superconducting processor", Nature **574**, 505–510 (2019).

[122] A. J. Daley, "Quantum trajectories and open many-body quantum systems", Advances in Physics **63**, 77–149 (2014).

[123] S. Finazzi, A. Le Boité, F. Storme, A. Baksic and C. Ciuti, "Corner-space renormalization method for driven-dissipative two-dimensional correlated systems", Phys. Rev. Lett. **115**, 080604 (2015).

[124] C. Le Bris and P. Rouchon, "Low-rank numerical approximations for high-dimensional lindblad equations", Phys. Rev. A **87**, 022125 (2013).

[125] Y.-T. Chen, C. Farquhar and R. M. Parrish, "Low-rank density-matrix evolution for noisy quantum circuits", npj Quantum Information **7**, 1–12 (2021).

[126] G. McCaul, K. Jacobs and D. I. Bondar, "Fast computation of dissipative quantum systems with ensemble rank truncation", Phys. Rev. Research **3**, 013017 (2021).

[127] C. D. Bruzewicz, J. Chiaverini, R. McConnell and J. M. Sage, "Trapped-ion quantum computing: Progress and challenges", Applied Physics Reviews **6**, 021314 (2019).

[128] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao,

P. Yeh, A. Zalcman, H. Neven and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor", Nature **574**, 505–510 (2019).

[129] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu and J.-W. Pan, "Quantum computational advantage using photons", Science **370**, 1460–1463 (2020).

[130] S. J. Devitt, W. J. Munro and K. Nemoto, "Quantum error correction for beginners", Reports on Progress in Physics **76**, 076001 (2013).

[131] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik and A. G. White, "Towards quantum chemistry on a quantum computer", Nature Chemistry **2**, 106–111 (2010).

[132] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe and S. Lloyd, "Quantum machine learning", Nature **549**, 195–202 (2017).

[133] R. Orús, S. Mugel and E. Lizaso, "Quantum computing for finance: Overview and prospects", Reviews in Physics **4**, 100028 (2019).

[134] J. M. Martinis, "Qubit metrology for building a fault-tolerant quantum computer", npj Quantum Information **1**, 1–3 (2015).

[135] R. Harper, S. T. Flammia and J. J. Wallman, "Efficient learning of quantum noise", Nature Physics **16**, 1184–1188 (2020).

[136] I. H. Deutsch, "Harnessing the Power of the Second Quantum Revolution", PRX Quantum **1**, 020101 (2020).

[137] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation and J. M. Gambetta, "Validating quantum computers using randomized model circuits", Phys. Rev. A **100**, 032328 (2019).

[138] D. Aharonov and M. Ben-Or, "Polynomial simulations of decohered quantum computers", in Proceedings of 37th Conference on Foundations of Computer Science (Oct. 1996), pp. 46–55.

[139] S. Aaronson and D. J. Brod, "Boson sampling with lost photons", Phys. Rev. A **93**, 012335 (2016).

[140] A. W. Harrow and M. A. Nielsen, "Robustness of quantum gates in the presence of noise", Phys. Rev. A **68**, 012308 (2003).

[141] D. Koch, B. Martin, S. Patel, L. Wessing and P. M. Alsing, "Demonstrating NISQ era challenges in algorithm design on IBM's 20 qubit quantum computer", AIP Advances **10**, 095101 (2020).

[142] R. LaRose, "Overview and Comparison of Gate Level Quantum Software Platforms", Quantum **3**, 130 (2019).

[143] G. Vidal, "Efficient Classical Simulation of Slightly Entangled Quantum Computations", Phys. Rev. Lett. **91**, 147902 (2003).

[144] Y. Zhou, E. M. Stoudenmire and X. Waintal, "What Limits the Simulation of Quantum Computers?", Phys. Rev. X **10**, 041038 (2020).

[145] K. Noh, L. Jiang and B. Fefferman, "Efficient classical simulation of noisy random quantum circuits in one dimension", Quantum **4**, 318 (2020).

[146] J. Napp, R. L. La Placa, A. M. Dalzell, F. G. S. L. Brandao and A. W. Harrow, "Efficient classical simulation of random shallow 2D quantum circuits", arXiv: 2001.00021 (2020).

[147] C. Oh, K. Noh, B. Fefferman and L. Jiang, "Classical simulation of lossy boson sampling using matrix product operators", arXiv: 2101.11234 (2021).

[148] T. Jones, A. Brown, I. Bush and S. C. Benjamin, "QuEST and high performance simulation of quantum computers", Scientific Reports **9**, 10736 (2019).

[149] D. Gottesman, "Fault-tolerant quantum computation with constant overhead", Quantum Info. Comput. **14**, 1338–1372 (2014).

[150] Y. S. Weinstein, T. F. Havel, J. Emerson, N. Boulant, M. Saraceno, S. Lloyd and D. G. Cory, "Quantum process tomography of the quantum Fourier transform", The Journal of Chemical Physics **121**, Publisher: American Institute of Physics, 6117–6133 (2004).

[151] R. Klesse and S. Frank, "Quantum error correction in spatially correlated quantum noise", Phys. Rev. Lett. **95**, 230503 (2005).

[152] H.-P. Breuer and F. Petruccione, *The theory of open quantum systems*, Publication Title: The Theory of Open Quantum Systems (Oxford University Press, 2007).

[153] J. Lebreuilly, A. Biella, F. Storme, D. Rossini, R. Fazio, C. Ciuti and I. Carusotto, "Stabilizing strongly correlated photon fluids with non-Markovian reservoirs", Phys. Rev. A **96**, 033828 (2017).

[154] J. Guillaud and M. Mirrahimi, "Repetition cat qubits for fault-tolerant quantum computation", Phys. Rev. X **9**, 041053 (2019).

[155] J. Preskill, "Quantum computing in the NISQ era and beyond", Quantum **2**, 79 (2018).

[156] C. Beck and F. Schögl, *Thermodynamics of Chaotic Systems: An Introduction*, Cambridge Nonlinear Science Series (Cambridge University Press, Cambridge, 1993).

[157] J. E. Gentle, *Matrix Algebra: Theory, Computations, and Applications in Statistics*, Springer Texts in Statistics (Springer-Verlag, New York, 2007).

[158] M. B. Plenio and P. L. Knight, "The quantum-jump approach to dissipative dynamics in quantum optics", Rev. Mod. Phys. **70**, 101–144 (1998).

[159] A. Grimm, N. E. Frattini, S. Puri, S. O. Mundhada, S. Touzard, M. Mirrahimi, S. M. Girvin, S. Shankar and M. H. Devoret, "Stabilization and operation of a kerr-cat qubit", Nature **584**, 205–209 (2020).

[160] D. Deutsch and R. Penrose, "Quantum theory, the Church-Turing principle and the universal quantum computer", Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **400**, 97–117 (1985).

[161] F. Marquardt and A. Püttmann, "Introduction to dissipation and decoherence in quantum systems", arXiv:0809.4403, arXiv:0809.4403 (2008).

[162] D. M. Greenberger, M. A. Horne and A. Zeilinger, "Going Beyond Bell's Theorem", arXiv: 0712.0921 (2007).

[163] D. Bouwmeester, J.-W. Pan, M. Daniell, H. Weinfurter and A. Zeilinger, "Observation of Three-Photon Greenberger-Horne-Zeilinger Entanglement", Phys. Rev. Lett. **82**, 1345–1349 (1999).

[164] J.-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter and A. Zeilinger, "Experimental test of quantum nonlocality in three-photon Greenberger–Horne–Zeilinger entanglement", Nature **403**, 515–519 (2000).

[165] R. Horodecki, P. Horodecki, M. Horodecki and K. Horodecki, "Quantum entanglement", Reviews of Modern Physics **81**, 865–942 (2009).

[166] M. Mohseni, A. T. Rezakhani and D. A. Lidar, "Quantum-process tomography: resource analysis of different strategies", Phys. Rev. A **77** (2008).

[167] J. L. O'Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph and A. G. White, "Quantum process tomography of a controlled-not gate", Phys. Rev. Lett. **93**, 080502 (2004).

[168] F. Nissen, J. M. Fink, J. A. Mlynek, A. Wallraff and J. Keeling, "Collective suppression of linewidths in circuit QED", Phys. Rev. Lett. **110**, 203602 (2013).

[169] N. Shammah, S. Ahmed, N. Lambert, S. De Liberato and F. Nori, "Open quantum systems with local and collective incoherent processes: efficient numerical simulations using permutational invariance", Phys. Rev. A **98** (2018).

[170] G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial neural networks", Science **355**, 602–606 (2017).

[171] G. Carleo, Y. Nomura and M. Imada, "Constructing exact representations of quantum many-body systems with deep neural networks", Nature Communications **9**, 5322 (2018).

[172] M. Schmitt and M. Heyl, "Quantum Many-Body Dynamics in Two Dimensions with Artificial Neural Networks", Phys. Rev. Lett. **125**, 100503 (2020).

[173] O. Sharir, Y. Levine, N. Wies, G. Carleo and A. Shashua, "Deep Autoregressive Models for the Efficient Variational Simulation of Many-Body Quantum Systems", Phys. Rev. Lett. **124**, 020503 (2020).

[174] T. Fösel, P. Tighineanu, T. Weiss and F. Marquardt, "Reinforcement learning with neural networks for quantum feedback", Phys. Rev. X **8**, 031084 (2018).

[175] Y. Zeng, J. Shen, S. Hou, T. Gebremariam and C. Li, "Quantum control based on machine learning in an open quantum system", Physics Letters A **384**, 126886 (2020).

[176] J. Sun, X. Yuan, T. Tsunoda, V. Vedral, S. C. Benjamin and S. Endo, "Mitigating Realistic Noise in Practical Noisy Intermediate-Scale Quantum Devices", Phys. Rev. Applied **15**, 034026 (2021).

[177] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition", Proc. IEEE **86**, 2278–2324 (1998).

[178] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel and D. Hassabis, "Mastering the game of go with deep neural networks and tree search", Nature **529**, 484–489 (2016).

[179] A. Brock, J. Donahue and K. Simonyan, "Large Scale GAN Training for High Fidelity Natural Image Synthesis", arXiv:1809.11096 (2018).

[180] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever and D. Amodei, "Language Models are Few-Shot Learners", arXiv:2005.14165 (2020).

[181] G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial neural networks", Science **355**, 602–606 (2017).

[182] O. Sharir, Y. Levine, N. Wies, G. Carleo and A. Shashua, "Deep autoregressive models for the efficient variational simulation of many-body quantum systems", Phys. Rev. Lett. **124**, 020503 (2020).

[183] K. Choo, T. Neupert and G. Carleo, "Two-dimensional frustrated J_1-J_2 model studied with neural network quantum states", Phys. Rev. B **100**, 125124 (2019).

[184] M. Schmitt and M. Heyl, "Quantum many-body dynamics in two dimensions with artificial neural networks", Phys. Rev. Lett. **125**, 100503 (2020).

[185] I. López Gutiérrez and C. B. Mendl, "Real time evolution with neural-network quantum states", arXiv:1912.08831 (2019).

[186] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko and G. Carleo, "Neural-network quantum state tomography", Nature Physics **14**, 447–450 (2018).

[187] F. Vicentini, A. Biella, N. Regnault and C. Ciuti, "Variational neural-network ansatz for steady states in open quantum systems", Phys. Rev. Lett. **122**, 250503 (2019).

[188] M. J. Hartmann and G. Carleo, "Neural-network approach to dissipative quantum many-body dynamics", Phys. Rev. Lett. **122**, 250502 (2019).

[189] N. Yoshioka and R. Hamazaki, "Constructing neural stationary states for open quantum many-body systems", Phys. Rev. B **99**, 214306 (2019).

[190] A. Nagy and V. Savona, "Variational quantum Monte Carlo method with a neural-network ansatz for open quantum systems", Phys. Rev. Lett. **122**, 250501 (2019).

[191] B. Jónsson, B. Bauer and G. Carleo, "Neural-network states for the classical simulation of quantum computing", arXiv preprint arXiv:1808.05232 (2018).

[192] M. Medvidović and G. Carleo, "Classical variational simulation of the quantum approximate optimization algorithm", npj Quantum Information **7**, 10.1038/s41534-021-00440-z (2021).

[193] Y. Levine, O. Sharir, N. Cohen and A. Shashua, "Quantum entanglement in deep learning architectures", Phys. Rev. Lett. **122**, 065301 (2019).

[194] O. Sharir, A. Shashua and G. Carleo, "Neural tensor contractions and the expressive power of deep neural quantum states", arXiv:2103.10293 (2021).

[195] D. Wu, R. Rossi, F. Vicentini and G. Carleo, "From Tensor Network Quantum States to Tensorial Recurrent Neural Networks", arXiv:2206.12363 (2022).

[196] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning* (The MIT Press, 2016).

[197] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, "Equation of state calculations by fast computing machines", The Journal of Chemical Physics **21**, 1087–1092 (1953).

[198] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications", Biometrika **57**, 97–109 (1970).

[199] R. Jastrow, "Many-body problem with strong forces", Phys. Rev. **98**, 1479–1484 (1955).

[200] E. Manousakis, "The spin-½ Heisenberg antiferromagnet on a square lattice and its application to the cuprous oxides", Rev. Mod. Phys. **63**, 1–62 (1991).

[201] A. N. Kolmogorov, "On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition", in Doklady Akademii Nauk, Vol. 114 (Russian Academy of Sciences, 1957), pp. 953–956.

[202] G. Cybenko, "Approximation by superpositions of a sigmoidal function", Math. Control Signals Syst. **2**, 303–314 (1989).

[203] K. Hornik, "Approximation capabilities of multilayer feedforward networks", Neural Netw. **4**, 251–257 (1991).

[204] M. Caron, H. Touvron, I. Misra, H. Jégou, J. Mairal, P. Bojanowski and A. Joulin, "Emerging Properties in Self-Supervised Vision Transformers", arXiv:2104.14294 (2021).

[205] A. Nichol, P. Dhariwal, A. Ramesh, P. Shyam, P. Mishkin, B. McGrew, I. Sutskever and M. Chen, "GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models", arXiv:2112.10741 (2021).

[206] M. Ruggeri, S. Moroni and M. Holzmann, "Nonlinear network description for many-body quantum systems in continuous space", Phys. Rev. Lett. **120**, 205302 (2018).

[207] J. Hermann, Z. Schätzle and F. Noé, "Deep-neural-network solution of the electronic schrödinger equation", Nature Chemistry **12**, 891–897 (2020).

[208] H. Saito, "Solving the Bose-Hubbard Model with Machine Learning", Journal of the Physical Society of Japan **86**, 093001, 093001 (2017).

[209] G. Pescia, J. Han, A. Lovato, J. Lu and G. Carleo, "Neural-network quantum states for periodic systems in continuous space", Phys. Rev. Research **4**, 023138, 023138 (2022).

[210] A. Barra, A. Bernacchia, E. Santucci and P. Contucci, "On the equivalence of Hopfield networks and Boltzmann machines", Neural Netw. **34**, 1–9 (2012).

[211] D.-L. Deng, X. Li and S. Das Sarma, "Quantum entanglement in neural network states", Phys. Rev. X **7**, 021021 (2017).

[212] J. Chen, S. Cheng, H. Xie, L. Wang and T. Xiang, "Equivalence of restricted Boltzmann machines and tensor network states", Phys. Rev. B **97**, 085104 (2018).

[213] X. Gao and L.-M. Duan, "Efficient representation of quantum many-body states with deep neural networks", Nature Communications **8**, 662 (2017).

[214] D. Luo, G. Carleo, B. K. Clark and J. Stokes, "Gauge equivariant neural networks for quantum lattice gauge theories", Phys. Rev. Lett. **127**, 276402 (2021).

[215] D. Wu, L. Wang and P. Zhang, "Solving statistical mechanics using variational autoregressive networks", Phys. Rev. Lett. **122**, 080602 (2019).

[216] C. Roth and A. H. MacDonald, "Group Convolutional Neural Networks Improve Quantum State Accuracy", arXiv:2104.05085 (2021).

[217] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez and J. I. Cirac, "Neural-network quantum states, string-bond states, and chiral topological states", Phys. Rev. X **8**, 011006 (2018).

[218] M. Hibat-Allah, E. M. Inack, R. Wiersema, R. G. Melko and J. Carrasquilla, "Variational neural annealing", Nat. Mach. Intell. **3**, 952–961 (2021).

[219] M. Schmitt, M. M. Rams, J. Dziarmaga, M. Heyl and W. H. Zurek, "Quantum phase transition dynamics in the two-dimensional transverse-field Ising model", Science Advances **8** (2022).

[220] M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko and J. Carrasquilla, "Recurrent neural network wave functions", Phys. Rev. Res. **2**, 023358 (2020).

[221] F. Becca and S. Sorella, *Quantum Monte Carlo approaches for correlated systems* (Cambridge University Press, 2017).

[222] D. Hofmann, G. Fabiani, J. H. Mentink, G. Carleo and M. A. Sentef, "Role of stochastic noise and generalization error in the time propagation of neural-network quantum states", SciPost Phys. **12**, 165 (2022).

[223] C.-Y. Park and M. J. Kastoryano, "Geometry of learning neural quantum states", Phys. Rev. Res. **2**, 023232 (2020).

[224] F. Vicentini, D. Hofmann, A. Szabó, D. Wu, C. Roth, C. Giuliani, G. Pescia, J. Nys, V. Vargas-Calderon, N. Astrakhantsev and G. Carleo, "NetKet 3: Machine Learning Toolbox for Many-Body Quantum Systems", arXiv:2112.10526, arXiv:2112.10526 (2021).

[225] X. Yuan, S. Endo, Q. Zhao, Y. Li and S. C. Benjamin, "Theory of variational quantum simulation", Quantum **3**, 191 (2019).

[226] R. Puebla, O. Marty and M. B. Plenio, "Quantum kibble-zurek physics in long-range transverse-field ising models", Phys. Rev. A **100**, 032115 (2019).

[227] S.-H. Lin and F. Pollmann, "Scaling of Neural-Network Quantum States for Time Evolution", Physica Status Solidi B Basic Research **259**, 2100172 (2022).

[228] A. Nagy and V. Savona, "Variational quantum Monte Carlo method with a neural-network ansatz for open quantum systems", Phys. Rev. Lett. **122**, 250501 (2019).

[229] D. Luo, Z. Chen, J. Carrasquilla and B. K. Clark, "Autoregressive neural network for simulating open quantum systems via a probabilistic formulation", Phys. Rev. Lett. **128**, 090501 (2022).

[230] M. Reh, M. Schmitt and M. Gärttner, "Time-dependent variational principle for open quantum systems with artificial neural networks", Phys. Rev. Lett. **127**, 230501 (2021).

[231] F. Vicentini, R. Rossi and G. Carleo, "Positive-definite parametrization of mixed quantum states with deep neural networks", arXiv:2206.13488 (2022).

[232] G. Carleo, L. Cevolani, L. Sanchez-Palencia and M. Holzmann, "Unitary dynamics of strongly interacting bose gases with the time-dependent variational monte carlo method in continuous space", Phys. Rev. X **7**, 031026 (2017).

[233] A. Fawzi, M. Balog, A. Huang, T. Hubert, B. Romera-Paredes, M. Barekatain, A. Novikov, F. J. R. Ruiz, J. Schrittwieser, G. Swirszcz, D. Silver, D. Hassabis and P. Kohli, "Discovering faster matrix multiplication algorithms with reinforcement learning", Nature **610**, 47–53 (2022).

[234] M. Krenn, R. Pollice, S. Y. Guo, M. Aldeghi, A. Cervera-Lierta, P. Friederich, G. dos Passos Gomes, F. Häse, A. Jinich, A. Nigam, Z. Yao and A. Aspuru-Guzik, "On scientific understanding with artificial intelligence", arXiv:2204.01467 (2022).