Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

École doctorale Mathématiques, sciences de l'information et de l'ingénieur | ED 269 Université de Strasbourg

ECOLE DOCTORALE MATHEMATIQUE, SCIENCE DE L'INFORMATION ET DE L'INGENIEUR- ED269

Laboratoire ICube (UMR CNRS 7357)

THÈSE

présentée par :

Tam'si LEY

Soutenue le : 11/07/2025

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/Spécialité : INFORMATIQUE

Optimizing gate pulses and circuits of Rydberg quantum computers with artificial evolution

Composition du jury

Présidente et rapporteur:	
Pr. Arghavan Safavi-Naini	Professeur, Université d'Amsterdam, Pays-Bas
Second rapporteur:	
Pr. Lee Spector	Professeur, Amherst College, États-Unis
Autres membres du jury:	
Dr. Aline Deruyver, directrice	Maîtresse de Conférences HDR, Université de Strabourg
Dr. Johannes Schachenmayer, codirecteur	Chargé de recherche HDR, Université de Strasbourg
Pr. Pierre Collet, encadrant	Professeur, Université Andrés Bello, Chili
Dr. Anna Ouskova Leonteva, encadrante	QPerfect, Centre Européen des Sciences Quantiques

Tam'si Ley

Optimizing gate pulses and circuits of Rydberg quantum computers with artificial evolution

Résumé

Cette thèse explore la synthèse de portes quantiques par évolution artificielle (EA), une classe d'heuristiques d'optimisation. Elle est motivée par la sensibilité au bruit du matériel quantique, ici des atomes neutres en interaction de Rydberg optiquement piégés, requérant de moduler le contrôle en conséquence. Ayant introduit l'EA, en particulier CMA-ES et NSGA-III, nous les appliquons à Pareto-minimiser l'infidélité et la dégradation de portes sujettes à un bruit de contrôle stochastique (BCS), y obtenant de raisonnables taux d'erreur. Des atomes restant en état de Rydberg après une porte bruitée, ils sont désexcités dans un état binaire ou un mélange statistique de ceux-ci. Nous calculons les canaux d'erreurs de ces projections, les incluons dans la simulation de modèle d'Ising transverse, comparons l'évolution de la fidélité de circuit et de l'intrication d'opérateur dans chaque cas, concluons que la projection optimale dépend du circuit. La structure du contrôle de porte quantique est conservée sous BCS, nous formulons deux algorithmes en atteignant l'optimum global, tirés des fonctions de valeur et de complexité polynomiale en le nombre d'atomes.

Mots clés: Quantique, bit, porte, logique, atome, Rydberg, optique, contrôle, Pareto, optimisation, évolution artificielle, simulation, canal, bruit, erreur

Asbtract

This thesis explores the synthesis of quantum logic gates by artificial evolution (AE), a class of heuristic biology-inspired optimisation algorithms. It is motivated by the noise sensitivity of quantum hardware, Rydberg-interacting neutral atoms in optical lattices here, needing to modulate control pulses into robust gates. After introducing AE, focusing on CMA-ES and NSGA-III, we apply them to Pareto-minimise the infidelity and decay risk of pulses subject to stochastic control noise (SCN), getting reasonable error rates in this context. As noise leaves atoms in Rydberg states after the gate, they are de-excited to a qubit state or a statistical mixture thereof. We then derive the error channels of these projections, plug them in the emulation of a transverse-field Ising model, compare the evolution of circuit fidelity and operator entanglement in each scenario, and conclude that the best projection is a circuit-dependent question. As that derivation proves that the structure of gate control problems is conserved under SCN, we formulate two globally-optimal algorithms solving them, based on value functions and scaling favorably in the number of atoms.

Keywords: Quantum, bit, logic, gate, error, channel, noise, atom, Rydberg, optical, control, Pareto, optimisation, artificial evolution, simulation

Remerciements

Je remercie mes directeurs de thèse pour leur nombreuses indications de recherche, conseils et avis éclairés; mon encadrante pour son assistance scientifique, sa ténacité dans le travail et sa volonté de me voir réussir; mes collègues de C STB pour leur aide en informatique et tous les moments légers passés au laboratoire et en dehors; mes collègues du CESQ (doctorants et permanents) pour les discussions et idées de recherche, et l'organisation d'évènements scientifiques; mes proches pour leur aide psychologique et matérielle; et surtout ma compagne pour son assistance émotionnelle et encouragements de tous les jours.

Résumé

0.1 Introduction générale

L'informatique quantique (IQ) constitue l'un des domaines les plus dynamiques de la recherche contemporaine, à l'intersection de la physique, de l'informatique théorique et de l'ingénierie expérimentale. Son ambition fondamentale est de tirer parti des principes de la mécanique quantique — superposition, intrication, non-localité — afin de concevoir des dispositifs de calcul capables de résoudre efficacement des classes de problèmes réputés intraitables pour les ordinateurs classiques. Cette promesse repose notamment sur des résultats fondateurs tels que l'algorithme de Shor pour la factorisation ou l'algorithme de Grover pour la recherche dans des bases non triées, qui présentent un avantage asymptotique démontré.

Historiquement, l'idée d'un calcul quantique émerge dès les années 1980, lorsque Feynman puis Deutsch proposent un modèle computationnel tirant parti de la dynamique quantique elle-même. Feynman met en évidence l'impossibilité fondamentale de simuler efficacement des systèmes quantiques généraux sur des machines classiques, en raison de l'explosion combinatoire de l'espace des états. Deutsch va plus loin en formulant le concept d'ordinateur quantique universel, posant les bases d'une théorie computationnelle compatible avec la mécanique quantique.

Depuis ces travaux initiaux, de nombreuses avancées ont été réalisées sur le plan théorique (algorithmes, complexité, tolérance aux fautes) comme sur le plan expérimental (implémentations matérielles, contrôle des qubits, lecture et manipulation d'état). Plusieurs technologies concurrentes ont émergé : ions piégés, qubits supraconducteurs, photons intégrés, boîtes quantiques, ou encore atomes neutres manipulés par pinces optiques. Cette diversité reflète une dynamique de recherche intense et interdisciplinaire, mais elle souligne également l'absence actuelle d'un standard technologique unique.

Parmi les défis majeurs à surmonter pour espérer la réalisation d'un ordinateur quantique universel tolérant aux fautes, la maîtrise du bruit et de la décohérence est centrale. Les qubits sont des objets physiques fragiles, sensibles à leur environnement, aux imperfections de contrôle, à l'interaction résiduelle entre eux ou encore à des fluctuations techniques (température, stabilité laser, perturbations électromagnétiques). Ces perturbations entraînent des erreurs dans l'exécution des portes logiques quantiques, qui doivent être réduites ou corrigées afin de maintenir une dynamique computationnelle fiable. La correction d'erreurs quantiques (QEC) offre des solutions

structurelles, mais elle repose sur la capacité à implémenter des portes élémentaires avec une fidélité dépassant un certain seuil (le théorème du seuil de la tolérance aux fautes).

La présente thèse s'inscrit dans cette problématique de contrôle quantique en environnement bruité. Elle se focalise sur une architecture matérielle spécifique — les atomes neutres en interaction via l'effet de blocage de Rydberg — et sur une stratégie d'optimisation des portes logiques à deux qubits opérant dans ce cadre. L'objectif est double : d'une part, améliorer la fidélité et la robustesse des opérations quantiques en présence de bruit réaliste ; d'autre part, évaluer la pertinence d'approches dites d'« évolution artificielle » (telles que les algorithmes CMA-ES ou NSGA-III) comme alternatives ou compléments aux méthodes traditionnelles de contrôle optimal.

L'intérêt d'une telle démarche repose sur plusieurs constats. D'abord, les systèmes physiques impliqués sont complexes, non linéaires, et souvent mal connus, ce qui rend l'optimisation analytique difficile. Ensuite, les algorithmes évolutionnaires sont capables de traiter des paysages de coût non convexes, bruités, voire discontinus, en évitant les pièges classiques liés aux minima locaux. Enfin, ces méthodes sont facilement parallélisables, ce qui les rend particulièrement adaptées aux environnements de simulation haute performance nécessaires à l'évaluation de la fidélité d'une porte quantique sous bruit.

L'organisation du manuscrit suit une progression logique : une première partie présente les fondements physiques et informatiques du calcul quantique avec atomes froids ; une seconde explore les concepts d'optimisation, avec un accent mis sur les méthodes multi-objectifs et évolutionnaires ; une troisième partie propose des applications concrètes sur des portes à deux qubits soumises à des bruits réalistes ; enfin, une dernière partie esquisse des perspectives théoriques autour de la commande robuste et de la programmation dynamique stochastique.

Ce chapitre d'introduction pose ainsi le cadre général de la thèse : entre modélisation physique, optimisation numérique et expérimentation simulée, il s'agit de contribuer à la conception de portes quantiques plus fiables, en s'appuyant sur des outils computationnels modernes et adaptatifs.

0.2 Ordinateurs quantiques à atomes froids

Parmi les diverses plateformes matérielles développées pour l'implémentation physique de l'informatique quantique, les systèmes à atomes neutres en interaction via l'effet de

blocage de Rydberg ont émergé comme l'un des candidats les plus prometteurs. Cette architecture combine plusieurs avantages notables : un fort degré d'extensibilité (scalabilité), une manipulation optique très flexible, un couplage contrôlable entre qubits, ainsi qu'une relative isolation de l'environnement. Ce chapitre explore en détail les fondements physiques et expérimentaux de cette approche, en posant les bases nécessaires à l'étude des protocoles de contrôle étudiés dans la suite du manuscrit.

La base de cette technologie repose sur des atomes neutres — typiquement des atomes alcalins comme le rubidium ou le césium — piégés individuellement dans des potentiels optiques créés par focalisation de faisceaux lasers. Grâce au refroidissement laser, ces atomes sont amenés à des températures de l'ordre du microkelvin, permettant leur confinement dans des pièges de type « pinces optiques » (optical tweezers) avec une précision nanométrique. Des dispositifs d'optique adaptative et d'adressage local permettent de reconfigurer dynamiquement la géométrie du réseau, de détecter et de repositionner les atomes, assurant ainsi un contrôle très fin sur l'état initial du système.

Le codage de l'information quantique s'effectue à l'aide de deux niveaux atomiques stables, souvent choisis dans les états hyperfins de la couche fondamentale. Les opérations logiques sont réalisées en appliquant des impulsions laser qui permettent d'exciter un ou plusieurs atomes vers des états de Rydberg — des états très excités caractérisés par un nombre quantique principal élevé. Ces états présentent des propriétés remarquables : leur taille électronique est grande, leur polarisabilité est forte, et surtout, ils interagissent de manière intense via des forces dipolaires ou van der Waals, même à plusieurs micromètres de distance.

L'effet de blocage de Rydberg constitue le cœur fonctionnel de cette architecture. Lorsqu'un atome est excité dans un état de Rydberg, il modifie le spectre d'énergie de ses voisins immédiats, empêchant leur excitation simultanée : on parle de blocage dipolaire. Ce phénomène permet de créer une interaction conditionnelle entre qubits, indispensable pour implémenter des portes à deux qubits telles que la porte contrôlée-Z (CZ) ou la porte CNOT. En pratique, cette interaction est exploitée via des séquences d'impulsions laser soigneusement calibrées, permettant une opération logique cohérente avec des durées typiques de quelques centaines de nanosecondes.

Les expériences récentes ont montré des avancées significatives : génération de réseaux bidimensionnels de plusieurs centaines d'atomes (jusqu'à 300 ou 400), contrôle individuel des qubits, lecture par fluorescence quasi sans erreur, et fidélités de porte dépassant les 99 % pour certains protocoles. Des entreprises comme Pasqal ou QuEra, issues de collaborations académiques, développent actuellement des processeurs quantiques à base de réseaux Rydberg exploitant ces principes, avec des objectifs à court terme dans la simulation quantique et la résolution de problèmes combinatoires.

Cependant, de nombreux défis subsistent. Le premier concerne la stabilité des faisceaux lasers utilisés pour l'adressage et l'excitation Rydberg : fluctuations d'intensité, de phase et de fréquence peuvent perturber les séquences de contrôle. Le second tient

à la gestion du bruit spatio-temporel dans les réseaux de pièges, qui peut induire une variabilité significative dans les temps de gate d'un atome à l'autre. D'autres limitations concernent le couplage aux états parasites (hors sous-espace de qubit), les effets de chauffage résiduel, et l'alignement temporel entre les impulsions d'excitation et de détection.

Une caractéristique importante de cette plateforme est le nombre relativement élevé de degrés de liberté accessibles pour le contrôle. En effet, les expériences permettent de moduler dynamiquement l'amplitude et la fréquence des impulsions laser ($\Omega(t)$, $\Delta(t)$), de contrôler la durée de chaque impulsion, la forme temporelle (façonnage d'impulsion), voire la polarisation lumineuse et l'angle d'incidence. Cette richesse ouvre la voie à des stratégies d'optimisation avancées, capables d'adapter précisément le protocole à l'environnement réel, au-delà des modèles idéalisés.

Dans ce contexte, l'objectif de la thèse est de tirer parti de cette flexibilité en explorant des méthodes d'optimisation inspirées de l'intelligence computationnelle — en particulier les algorithmes évolutionnaires. Ceux-ci permettront de définir des séquences d'impulsions robustes, efficaces et adaptées aux contraintes expérimentales. En somme, le chapitre met en évidence la combinaison unique qu'offrent les atomes neutres bloqués Rydberg : un système quantique hautement contrôlable, à forte connectivité, et doté de ressources de manipulation particulièrement riches pour le contrôle quantique.

0.3 Contrôle quantique et optimisation

Contrôle quantique et optimisation

Dans un système quantique contrôlé, l'objectif fondamental est de piloter dynamiquement l'évolution d'un état quantique ou d'un canal unitaire via des actions physiques accessibles. Cette tâche, connue sous le nom de contrôle quantique, repose sur la modulation précise de paramètres externes (champs laser, impulsions micro-ondes, gradients magnétiques, etc.) dans le but d'atteindre un objectif cible défini : réaliser une opération logique, préparer un état, supprimer une excitation indésirable, etc.

La description dynamique d'un système quantique ouvert est généralement donnée par l'équation maîtresse de Lindblad :

$$\frac{d\rho(t)}{dt} = -i[H(t), \rho(t)] + \sum_k \left(L_k \rho(t) L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho(t)\} \right),$$

où $\rho(t)$ est la matrice de densité du système, $H(t)$ le Hamiltonien contrôlé, et L_k des opérateurs de Lindblad représentant les couplages dissipatifs au bain environnemental. L'ensemble du contrôle repose alors sur le choix d'un ensemble de fonctions de contrôle $u_i(t)$ apparaissant comme coefficients dans l'écriture de $H(t)$:

$$H(t) = H_0 + \sum_i u_i(t) H_i,$$

avec H_0 le Hamiltonien libre et H_i les opérateurs de contrôle.

Le problème du contrôle quantique optimal consiste alors à déterminer les fonctions $u_i(t)$ qui maximisent une fonction objectif $J[u]$ — par exemple, la fidélité par rapport à une opération cible, ou la minimisation du temps de porte, sous des contraintes physiques et technologiques (énergie maximale, bande passante limitée, bruit instrumental, etc.).

Dans le cadre de l'informatique quantique à atomes de Rydberg, les variables de contrôle incluent typiquement : l'amplitude Rabi $\Omega(t)$, la détonation $\Delta(t)$, la polarisation du champ laser, et parfois même le profil spatial ou l'angle d'incidence. Ces variables peuvent être modulées via des acousto-optic modulators (AOM), des phase-locked loops (PLL) ou des générateurs arbitraires. En pratique, on discrétise le contrôle en un vecteur de paramètres, de dimension typiquement comprise entre 20 et 100, représentant les valeurs de $\Omega(t)$, $\Delta(t)$, etc., sur une grille temporelle fine.

Deux grandes classes de méthodes d'optimisation sont employées :

Méthodes déterministes telles que GRAPE (Gradient Ascent Pulse Engineering) ou KROTOV, qui exploitent le calcul de gradient du fonctionnel $J[u]$ via des méthodes d'adjoint. Elles permettent une convergence rapide sur des modèles différentiables et bien conditionnés. Cependant, elles sont sensibles aux minima locaux, exigent une connaissance analytique du système (modèle exact, dérivabilité) et sont souvent instables en présence de bruit, de non-linéarités ou de discontinuités (comme le passage par des états non contrôlés).

Méthodes évolutionnaires , notamment CMA-ES (Covariance Matrix Adaptation Evolution Strategy), qui échantillonnent l'espace de contrôle selon une loi gaussienne adaptative, et construisent progressivement une meilleure approximation des régions optimales. Ces méthodes ne nécessitent pas de dérivée, sont robustes aux modèles bruités et tolèrent les paysages de coût non convexes. Leur convergence est plus lente mais leur exploration est plus globale. Dans le cas de systèmes à dynamique stochastique, ou à dépendances non linéaires complexes, elles surpassent souvent les approches déterministes.

Un cas particulièrement important en physique expérimentale est celui du contrôle multi-objectif. En effet, dans un système réel, plusieurs objectifs sont simultanément souhaités : une fidélité élevée, un temps de passage minimal dans les états excités (pour réduire la décohérence), une robustesse aux erreurs expérimentales, ou encore une faible sensibilité aux variations spatiales du champ laser. Les optimiser simultanément revient à traiter un problème de Pareto, où l'on cherche à générer un ensemble de solutions non dominées (aucune n'est meilleure dans tous les critères), plutôt qu'un unique optimum.

Dans cette thèse, nous mobilisons l'algorithme NSGA-III (Non-dominated Sorting

Genetic Algorithm III), particulièrement adapté à l'optimisation multi-objectif à plus de deux critères. Il repose sur un tri par domination et l'échantillonnage d'un simplexe de référence pour couvrir uniformément le front de Pareto. Cette approche permet au physicien de choisir parmi plusieurs compromis pertinents, selon les conditions expérimentales du jour.

Un point fondamental du contrôle quantique est sa robustesse : une séquence d'impulsions doit produire une bonne opération même en présence d'incertitudes expérimentales. Pour cela, la fonction objectif doit être évaluée non pas sur un seul jeu de paramètres physiques (laser parfait, géométrie idéale), mais sur un ensemble d'instances bruitées. Cela implique une phase de « sampling » : on tire aléatoirement des réalisations du bruit (de phase, d'intensité, de fréquence), on évalue la fidélité dans chacun des cas, et on optimise l'espérance de cette performance. Cela rend l'optimisation plus coûteuse (car elle implique des dizaines ou centaines de simulations par candidat), mais bien plus réaliste.

La fidélité d'une porte quantique est mesurée à l'aide d'indicateurs tels que :

$$F(U,V) = \left|\frac{1}{d}\mathrm{Tr}(U^\dagger V)\right|^2,$$

où U est la porte cible idéale, V l'opération réellement implémentée, et d la dimension du sous-espace logique (généralement $d = 4$ pour deux qubits). La moyenne sur les réalisations du bruit donne une mesure de robustesse. D'autres métriques comme la distance trace, la distance de diamond, ou encore l'entropie résiduelle, peuvent également être utilisées dans un cadre plus général.

Enfin, nous soulignons que le contrôle optimal n'est pas une fin en soi : il doit être intégré dans un pipeline expérimental complet, incluant la caractérisation du bruit (noise spectroscopy), la calibration de l'appareil, la validation par tomographie partielle, et le suivi en temps réel des dérives. L'interfaçage entre modélisation numérique et données expérimentales constitue un enjeu crucial de cette discipline, à la frontière entre la physique théorique, l'ingénierie et l'informatique.

Évolution artificielle et optimisation multi-objectif

Les approches évolutionnaires, inspirées du fonctionnement des processus naturels d'évolution, ont montré un potentiel remarquable dans des contextes où les méthodes analytiques classiques ou les algorithmes à base de gradients atteignent leurs limites. Dans le cas du contrôle quantique de systèmes complexes, soumis à des perturbations aléatoires, à des interactions non linéaires ou à une connaissance incomplète du modèle, ces méthodes offrent une alternative robuste, flexible et globalement convergente.

L'optimisation évolutionnaire consiste à faire évoluer, génération après génération, une population d'individus (dans notre cas, des vecteurs de contrôle) soumis à des

opérations de mutation, recombinaison et sélection. Chaque individu encode un protocole de contrôle — typiquement, une discrétisation temporelle d'une impulsion laser, une suite de valeurs d'intensité, de fréquence ou de phase — que l'on évalue en simulant sa performance sur un modèle physique réaliste du système quantique.

Le premier algorithme mobilisé dans cette thèse est le **CMA-ES** (Covariance Matrix Adaptation Evolution Strategy), une méthode stochastique sans gradient, particulièrement efficace pour l'optimisation continue de fonctions bruitées ou non différentiables. L'idée centrale du CMA-ES est de maintenir une gaussienne de génération définie par un vecteur moyen $\mu \in \mathbb{R}^n$ et une matrice de covariance $\Sigma \in \mathbb{R}^{n\times n}$, qui s'adaptent dynamiquement au paysage du coût.

À chaque itération, une population de λ candidats $x_i \sim \mathcal{N}(\mu, \Sigma)$ est générée, évaluée, puis les μ meilleurs (selon une métrique donnée, par exemple la fidélité moyenne) sont utilisés pour actualiser les paramètres statistiques :

$$\mu \leftarrow \sum_{i=1}^{\mu} w_i x_i, \quad \Sigma \leftarrow \sum_{i=1}^{\mu} w_i (x_i - \mu)(x_i - \mu)^T,$$

avec des poids w_i décroissants. Ce mécanisme encode à la fois une direction d'évolution (via μ) et une exploration anisotrope de l'espace (via Σ), ce qui permet de capter des corrélations physiques entre paramètres — par exemple, une synergie entre amplitude et détonation dans une impulsion.

L'un des atouts majeurs du CMA-ES réside dans son adaptabilité : il peut évoluer même sur des fonctions bruitées, discontinues ou mal conditionnées, sans nécessiter d'information de gradient. Cette propriété le rend particulièrement adapté à la simulation de portes quantiques en présence de bruit, où chaque évaluation implique une moyenne stochastique sur des réalisations bruitées du système. De plus, sa capacité à explorer globalement l'espace des solutions permet d'éviter les minima locaux typiques des méthodes basées sur des gradients.

Cependant, dans un cadre expérimental réaliste, les critères de performance sont souvent multiples et parfois conflictuels : fidélité maximale, durée minimale, énergie déposée réduite, robustesse vis-à-vis de certains types de bruit, etc. Cela appelle à une extension naturelle vers des méthodes **multi-objectives**, où l'on cherche non pas une unique solution optimale, mais un *ensemble de compromis optimaux*, appelé front de Pareto.

Pour cela, la thèse mobilise un second algorithme : **NSGA-III** (Non-dominated Sorting Genetic Algorithm III), conçu pour le traitement de problèmes multi-objectifs avec un nombre arbitraire de critères. Cet algorithme génétique repose sur les principes suivants : (1) construction de fronts de non-domination successifs parmi les individus ; (2) insertion guidée par un ensemble de points de référence bien répartis dans l'espace objectif, assurant une couverture homogène du front ; (3) sélection par élitisme et diversité.

Plus précisément, NSGA-III introduit un hyperplan de normalisation qui permet de projeter les solutions dans un espace réduit, où les points de référence servent à répartir équitablement la population le long du front. Chaque individu est ainsi rattaché à un point de référence, et un mécanisme de "niching" permet de maintenir la diversité de la population même en présence de régions denses.

Dans nos applications, les critères typiques incluent : (i) 1 – Fidélité, (ii) temps de séjour dans les états de Rydberg (proxy pour la décohérence), (iii) sensibilité aux fluctuations de phase laser. Le front de Pareto obtenu offre à l'expérimentateur une liberté nouvelle : choisir, selon ses contraintes du moment, une solution optimisant un critère spécifique, ou une solution intermédiaire équilibrant les compromis.

L'évaluation d'un individu dans ce contexte repose sur une chaîne de simulation complète : génération du vecteur de contrôle ; résolution de l'équation de Schrödinger (ou équation maître) associée ; calcul des observables (fidelité, pertes, durée) ; moyennage sur plusieurs réalisations du bruit. Ces étapes sont parallélisées via un moteur de tâches distribuées (par exemple `Dask` en Python), permettant de tester des centaines d'individus par génération.

Enfin, ces approches ne sont pas limitées à la recherche de contrôles "statistiques" : elles peuvent également être couplées à des schémas adaptatifs hybrides, où un pré-apprentissage basé sur des données expérimentales guide le choix initial de la population, ou encore à des algorithmes bayésiens, pour un raffinement local plus précis. L'intégration des méthodes évolutionnaires dans une boucle d'optimisation expérimentale en temps réel constitue d'ailleurs une des pistes les plus prometteuses pour les prochaines générations de processeurs quantiques.

0.4 Application : Optimisation de portes logiques en présence de bruit

Ce chapitre présente l'application concrète des méthodes d'optimisation évolutionnaires à la conception de portes logiques quantiques robustes dans une architecture à atomes neutres interagissant via l'effet de blocage de Rydberg. Il constitue le cœur technique de la thèse, croisant modélisation physique, simulation numérique stochastique et optimisation multi-objectif. L'enjeu est de concevoir des séquences de contrôle réalisant une porte à deux qubits (typiquement CZ ou CNOT) avec une fidélité élevée, en présence de bruit réaliste, dans des conditions proches de celles rencontrées en laboratoire.

Modèle physique et hypothèses de bruit

Le système étudié comprend deux atomes neutres piégés individuellement, chacun codant un qubit dans deux états hyperfins stables de la couche fondamentale. L'interaction

entre les atomes est activée via une excitation conditionnelle vers un état de Rydberg, conformément au protocole de gate CZ par blocage dipolaire. Le Hamiltonien effectif est de la forme :

$$H(t) = \sum_{i=1}^{2} \left[\frac{\Omega_i(t)}{2} \left(|0\rangle_i\langle r| + |r\rangle_i\langle 0| \right) + \Delta_i(t) |r\rangle_i\langle r| \right] + V_{rr} |rr\rangle\langle rr|,$$

où $\Omega_i(t)$ est l'amplitude Rabi du laser sur l'atome i, $\Delta_i(t)$ la détonation, $|r\rangle$ l'état de Rydberg, et V_{rr} l'énergie d'interaction de van der Waals.

Le bruit est introduit selon deux modalités principales : (1) un bruit de phase global, modélisé comme une détonation fluctuante $\Delta_i(t) + \delta_\phi(t)$, avec $\delta_\phi(t)$ tiré d'un bruit de type $1/f^\alpha$ (bruit rose) ; (2) un bruit d'intensité, modélisé par un facteur multiplicatif $\Omega_i(t) \to \Omega_i(t) \cdot (1 + \delta_I(t))$, où $\delta_I(t)$ est un bruit gaussien centré. Ces sources d'erreur traduisent les imperfections expérimentales telles que les dérives laser, les inhomogénéités spatiales ou la gigue temporelle des modulateurs.

Protocole d'optimisation et métriques d'évaluation

Le protocole consiste à discrétiser les fonctions $\Omega_i(t)$ et $\Delta_i(t)$ en vecteurs de paramètres, typiquement sur 20 à 30 points temporels. Chaque candidat dans la population évolutionnaire encode alors une séquence complète d'impulsions, qui est testée numériquement par intégration de l'équation de Schrödinger (ou de l'équation maîtresse dans certains cas), sur un ensemble de réalisations bruitées.

Les métriques d'évaluation sont :

- la **fidélité moyenne** $F = \langle |\mathrm{Tr}(U^\dagger V_{\text{eff}})/d|^2 \rangle$, où U est la porte cible, V_{eff} l'opération effective induite par l'impulsion optimisée et $d = 4$ la dimension du sous-espace logique ;
- la **durée d'occupation de l'état de Rydberg**, mesurée comme $T_r = \int_0^T \sum_i |\langle r \rangle_i(t)|^2 dt$, critère proxy de décohérence ;
- l'**écart-type de la fidélité** sur les différentes réalisations du bruit, mesurant la robustesse stochastique.

Optimisation mono-objectif avec CMA-ES

Dans un premier temps, l'algorithme CMA-ES est utilisé pour maximiser la fidélité moyenne F seule, en présence d'un bruit réaliste de phase. L'algorithme converge, après environ 150 générations, vers une solution atteignant $F > 0.996$ dans un scénario simulé contenant du bruit *pink* ($\alpha = 1$), ce qui dépasse les seuils typiques de tolérance aux fautes.

On observe que les impulsions optimisées présentent une structure non triviale : au lieu de simples impulsions carrées ou gaussiennes, les profils sont asymétriques, avec des plateaux, des rampes lentes, voire des oscillations correctives en fin de gate. Ces formes sont interprétées comme des compensations actives du bruit, intégrées dans la dynamique unitaire par l'algorithme d'optimisation.

Optimisation multi-objectif avec NSGA-III

On passe ensuite à une optimisation multi-objectif, en considérant simultanément la fidélité F et la durée d'occupation T_r. L'algorithme NSGA-III est paramétré avec une population de 100 individus sur 300 générations. Le front de Pareto obtenu révèle une diversité de compromis : certaines solutions privilégient une fidélité très haute ($F > 0.998$) au prix d'une occupation Rydberg significative ; d'autres réduisent fortement T_r tout en maintenant $F \approx 0.990$.

Ce front est d'autant plus précieux pour l'expérimentateur qu'il permet d'adapter la stratégie de contrôle aux priorités du moment : réduction du bruit de phase dans un environnement instable, ou préservation des états Rydberg dans des conditions de haute décohérence. On note également que la robustesse (définie comme l'écart-type de F) est corrélée positivement à la régularité des impulsions : les séquences les plus robustes sont souvent celles présentant un changement adiabatique progressif des paramètres.

Validation croisée et sensibilité au bruit

Une validation croisée est effectuée sur des instances de bruit générées de manière indépendante. On observe que les solutions optimisées conservent une performance élevée hors échantillon, confirmant leur caractère générique et non surajusté. La sensibilité à la variance du bruit (σ_ϕ^2, σ_I^2) est également cartographiée, montrant une décroissance lente de la fidélité avec le bruit, preuve de robustesse effective.

Par ailleurs, on constate que l'optimisation de la durée T_r contribue à stabiliser la fidélité vis-à-vis du bruit : un protocole qui minimise le séjour dans les états de Rydberg est moins affecté par les fluctuations de phase, ce qui rejoint des intuitions expérimentales.

Comparaison avec protocoles standards

Les séquences obtenues sont comparées à des impulsions classiques (carrées, chirp linéaire, séquences 'π-pulse'). On constate une amélioration de la fidélité de l'ordre de 1 à 2 points de pourcentage, ce qui est significatif dans une perspective de correction d'erreur. Les impulsions évolutionnaires présentent aussi une plus grande tolérance aux dérives lentes, et une meilleure immunité aux désaccords entre atomes (inhomogénéités de position). Enfin, la structure temporelle des impulsions optimisées peut servir de

base à des schémas analytiques : en ajustant des fonctions de base (splines, gaussiennes modulées) aux impulsions numériques, on peut produire des interpolations paramétrées prêtes à être testées en laboratoire.

Conclusion

L'ensemble des résultats confirme l'intérêt de l'approche évolutionnaire pour l'optimisation des portes quantiques. Contrairement aux approches fondées sur des modèles idéalisés ou des heuristiques manuelles, les méthodes CMA-ES et NSGA-III permettent de découvrir des solutions non intuitives, souvent plus performantes, et adaptables à différentes contraintes. Elles s'inscrivent pleinement dans une démarche moderne de co-conception entre modélisation, simulation et expérimentation, et constituent un levier concret pour améliorer la fiabilité des architectures à Rydberg.

0.5 Émulation classique de systèmes quantiques

L'émulation classique constitue un outil central dans cette thèse, en tant que passerelle entre le développement théorique de protocoles de contrôle et leur validation par des moyens computationnels. L'objectif est de pouvoir évaluer, avec une fidélité numérique élevée, les performances attendues de portes quantiques optimisées dans des architectures à atomes de Rydberg, en tenant compte des contraintes physiques du système.

Deux approches distinctes ont été mobilisées pour émuler des systèmes quantiques de tailles variables : la simulation à vecteur d'état complet (*Full Statevector*) et l'approche par opérateurs produits de matrices (*Matrix Product Operators*, MPO). Ces deux outils répondent à des besoins complémentaires, dictés par la taille du système simulé et la nature du bruit ou des pertes considérés.

La méthode Full Statevector repose sur une représentation explicite de l'état quantique global sous la forme d'un vecteur dans l'espace de Hilbert complet de dimension 2^n, où n est le nombre de qubits. Cette approche permet une description exacte de l'état du système et autorise l'accès à toutes les observables. Elle est particulièrement adaptée pour l'étude de systèmes de petite taille, et a été utilisée ici jusqu'à un maximum de **10 qubits**.

Dans cette thèse, l'approche Full Statevector a permis d'évaluer précisément les performances des portes à deux qubits, en présence ou non de bruit, sur des sous-ensembles restreints d'un registre quantique. Elle est également utilisée dans les boucles d'optimisation pour évaluer chaque candidat proposé par les algorithmes évolutionnaires. Sa simplicité conceptuelle et sa précision en font une référence pour les cas où l'espace d'état reste manipulable directement.

Lorsque la taille du système devient plus importante, la simulation exacte devient inenvisageable. Pour dépasser cette barrière, la thèse fait appel à la représentation par

Matrix Product Operators (MPO), qui offre une compression efficace des états mixtes et des super-opérateurs sous l'hypothèse que les corrélations quantiques restent localisées ou peu étendues.

Les MPO permettent de simuler des registres allant jusqu'à **20 qubits** dans cette thèse, en particulier pour étudier l'effet de pertes, de bruit corrélé ou de couplages localisés. L'approche repose sur une factorisation structurée de l'objet quantique global (état ou opérateur) en une chaîne de tenseurs, chacun associé à un qubit et relié à ses voisins via des indices de liaison de dimension bornée. Cette compression est rendue possible par l'existence d'une *décomposition de Schmidt* efficace : lorsque l'intrication bipartite entre les sous-systèmes reste modérée, la représentation MPO conserve une bonne précision avec un rang de lien réduit.

L'*entropie d'intrication* entre deux partitions successives du système est ainsi un indicateur clé de la complexité de simulation. Dans les situations explorées ici — bruit local, interactions limitées — cette entropie reste suffisamment faible pour maintenir la taille des tenseurs à un niveau maîtrisé. Cela permet la manipulation d'espaces de Hilbert qui seraient sinon inaccessibles en représentation pleine, tout en respectant les contraintes de mémoire et de calcul.

Dans le contexte de cette thèse, les MPO sont utilisés non pas pour simuler des dynamiques complètes, mais pour émuler des effets structurels liés à l'extension spatiale ou à la présence de dissipation. Cela permet, entre autres, d'évaluer la robustesse des protocoles optimisés dans des régimes réalistes à l'échelle du processeur quantique.

Le choix entre Full Statevector et MPO dépend directement du compromis entre précision et échelle du système. Les simulations en vecteur d'état sont utilisées pour les tâches nécessitant une très haute résolution, comme la reconstruction fidèle des opérations logiques sur deux qubits. À l'inverse, les MPO sont mobilisés pour explorer la scalabilité des protocoles ou leur comportement en présence de bruit distribué.

Les deux approches coexistent ainsi dans une architecture logicielle qui permet d'interfacer ces outils à des boucles d'optimisation, des analyses statistiques et des comparaisons systématiques entre familles d'impulsions.

Conclusion — L'émulation classique des systèmes quantiques constitue une composante essentielle de l'approche développée dans cette thèse. En s'appuyant sur les méthodes Full Statevector et MPO, il a été possible de valider les protocoles de contrôle dans des régimes représentatifs des expérimentations futures. Ces outils permettent de faire le lien entre l'abstraction algorithmique et la complexité physique réelle des dispositifs à atomes de Rydberg, et seront appelés à jouer un rôle central dans la co-conception des architectures quantiques à venir.

0.6 Application : l'impact des mécanismes post-portes de compensation du bruit sur les métriques standard de simulation quantique

Ce chapitre explore le rôle des mécanismes passifs de compensation du bruit appliqués à l'issue d'une porte quantique à deux qubits, dans le contexte de la simulation numérique de modèles à plusieurs corps. Il s'appuie sur une analyse en deux temps : (1) le calcul analytique des canaux quantiques bruités associés à une séquence de contrôle fixée réalisant une porte CZ, et (2) leur intégration dans des circuits de simulation numérique d'Hamiltoniens de type Ising transverse, afin d'en évaluer l'impact sur des métriques physiques pertinentes.

Calcul analytique des canaux quantiques

On considère une séquence de pulsations optiques déterminée, conçue pour implémenter une porte CZ idéale, définie par l'unitaire :

$$U_{\mathrm{CZ}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Plutôt que d'optimiser cette séquence, on part d'un protocole donné, sur lequel on impose des hypothèses statistiques réalistes sur le bruit (fluctuations lentes, bruit de phase, pertes dans les états de Rydberg, etc.). Le système est alors modélisé comme une dynamique ouverte, et la transformation induite est décrite par un **canal quantique effectif** $\mathcal{E}$, représentant l'effet agrégé du bruit sur l'espace des deux qubits logiques.

Le calcul de ce canal est mené **analytiquement**, sur la base des hypothèses bruitées : il ne s'agit pas d'une tomographie numérique ni expérimentale, mais bien d'une dérivation théorique des termes dissipatifs dominants. La carte $\mathcal{E}$ est exprimée sous forme de super-opérateur dans une base de Pauli, ou alternativement via la représentation de Choi ou de Kraus, selon les besoins de la simulation en aval.

L'étude inclut également une stratégie passive de mitigation du bruit : le **choix de l'état "qubnit"** — c'est-à-dire l'état de base dans lequel est projeté l'atome à l'issue de la porte — peut influencer la structure du canal obtenu. En modifiant cette base logique finale, on affecte la nature des erreurs induites (déphasage, bruit incohérent, perte de contraste). Cette projection ne nécessite ni correction active ni lecture intermédiaire : elle s'intègre naturellement à l'encodage logique du système et peut être sélectionnée pour minimiser certains effets perturbateurs spécifiques.

Intégration dans une simulation de hamiltoniens bruités

Les canaux quantiques $\mathcal{E}$ ainsi calculés sont ensuite intégrés dans des circuits quantiques digitaux simulant la dynamique d'Hamiltoniens de spin. On considère ici deux orientations différentes du modèle d'Ising transverse, qui sert de plateforme pour l'étude des dynamiques quantiques à plusieurs corps. Ces modèles sont discrétisés via une décomposition trotterisée, chaque pas de temps étant implémenté par un circuit comprenant des rotations locales et des portes CZ.

Dans cette simulation numérique, chaque porte CZ idéale est remplacée par le **canal bruité $\mathcal{E}$** correspondant, introduisant une altération réaliste de la dynamique cible. Ce montage permet d'évaluer l'effet cumulé du bruit sur la qualité de la simulation, et notamment sur les **métriques physiques d'intérêt** : énergie moyenne finale, observables locales, fidélité avec l'état idéalement attendu, structure de corrélations et d'intrication.

Les circuits sont évalués numériquement selon deux méthodes :

- **Full Statevector (FS)** jusqu'à 10 qubits : permettant une description exacte de l'état du système et l'accès à l'ensemble des observables ;
- **Matrix Product Operators (MPO)** jusqu'à 20 qubits : méthode de compression tensorielle adaptée aux systèmes faiblement intriqués. Elle s'appuie sur la décomposition de Schmidt et le contrôle de l'entropie d'intrication bipartite pour limiter le rang de lien et garantir une simulation efficace.

L'analyse met en évidence l'influence combinée de l'orientation du champ transverse dans le modèle d'Ising, du choix du canal bruité appliqué aux portes, et de la stratégie de projection choisie en sortie. Certaines combinaisons conduisent à une amplification du bruit par résonance dynamique, tandis que d'autres montrent une robustesse accrue. La structure spectrale du modèle simulé joue un rôle clé dans la propagation ou la neutralisation des erreurs.

Ce cadre d'étude permet ainsi d'évaluer, de manière systématique, la robustesse d'un protocole de simulation quantique face à un bruit structurel réaliste, et de proposer des ajustements simples — tels que le choix de la base logique — pour en atténuer l'impact à grande échelle.

0.7 Contrôle optimal déterministe et stochastique pour la synthèse robuste de portes quantiques

Ce chapitre s'intéresse à la formulation rigoureuse du problème de synthèse de portes quantiques robustes, en présence de bruit, comme un problème de contrôle optimal. On y distingue deux approches : l'une déterministe, traitant de perturbations

paramétriques connues ou discrètes ; l'autre stochastique, adaptée aux bruits aléatoires typiques des systèmes quantiques physiques. Les méthodes explorées ici visent à surmonter les limites des techniques traditionnelles de gradient ou de recherche évolutionnaire, en mobilisant des résultats issus du contrôle optimal, de l'analyse numérique et de la finance quantitative.

Difficulté fondamentale du problème de contrôle

La synthèse robuste d'une porte quantique peut se formuler comme un problème de contrôle optimal sur une dynamique de Schrödinger (ou de Lindblad) paramétrée. Le critère à optimiser — souvent la fidélité moyenne ou une norme de processus — dépend de la commande, typiquement la pulsation optique $\Omega(t)$, que l'on cherche à ajuster. L'outil classique pour résoudre ce type de problème est le *principe du maximum de Pontryagin*, qui ramène l'optimisation à la résolution d'un système différentiel non-linéaire, combinant équation d'état, équation adjointe et condition de stationnarité.

Cependant, dans ce contexte, ces équations se révèlent particulièrement instables : elles requièrent une bonne initialisation (c'est-à-dire une solution "devinée") pour converger, et leur résolution est rendue difficile par la non-linéarité forte des conditions au bord. Une alternative simple mais coûteuse est une méthode de bisection, cherchant la commande optimale en balayant l'espace des paramètres. Toutefois, cette approche souffre d'une complexité exponentielle en la dimension des trajectoires contrôlées, ce qui confirme la nature intrinsèquement difficile du problème.

Approche semi-linéaire via l'équation de Hamilton-Jacobi-Bellman

Une piste prometteuse est d'interpréter le problème sous-jacent comme la résolution d'une *équation de Hamilton-Jacobi-Bellman* (HJB), une équation aux dérivées partielles semi-linéaire dont la solution encode la commande optimale en tout point de l'espace-temps. Bien que la solution exacte de cette EDP soit hors de portée en dimension élevée, il est possible d'en proposer une approximation numérique par des méthodes de Monte-Carlo, fondées sur une reformulation variationnelle du problème.

Dans la thèse, une telle méthode est proposée pour approximer la solution de l'équation HJB, en échantillonnant des trajectoires d'état et en estimant numériquement les gradients nécessaires. Cette méthode produit une *solution candidate* pour l'équation de Pontryagin, qui peut ensuite servir de point de départ pour une méthode de Newton. L'intérêt majeur de cette stratégie est que la complexité en temps de calcul est ici indépendante de la dimension de l'espace de contrôle — une rupture avec les méthodes classiques.

Ce cadre s'applique particulièrement bien à des bruits modulés par un spectre discret (plusieurs fréquences cohérentes ou bruit synthétique), où l'on cherche à synthé-

tiser une commande robuste à un ensemble d'échantillons paramétriques. Le processus d'optimisation tire ainsi parti de l'information statistique globale sans recourir à des grilles de discrétisation exhaustives.

Contrôle optimal stochastique via la programmation dynamique

Lorsque le bruit affectant le système est véritablement aléatoire — typiquement modélisé comme un processus stochastique continu, tel que le bruit "rose" $1/f$ —, le problème de synthèse robuste se formule naturellement comme un *contrôle optimal stochastique*. Dans ce cas, le principe de Pontryagin n'est plus directement applicable. En revanche, les outils de la programmation dynamique stochastique, bien développés en finance quantitative, permettent une résolution rigoureuse.

Le cœur de cette approche repose sur l'équation de Hamilton-Jacobi-Bellman stochastique, équivalente à une formulation variationnelle du problème, dont la solution peut être approchée à l'aide de formules de type Feynman-Kac non-linéaires. Ces expressions lient la solution de l'EDP à des moyennes d'observables le long de trajectoires bruitées, et sont accessibles via des simulations Monte-Carlo. L'intérêt majeur est que la complexité numérique reste, ici encore, indépendante de la dimension de l'espace d'états ou du contrôle, ce qui rend ces méthodes adaptées aux systèmes à haute dimension.

Une hypothèse structurante ici est la modélisation du bruit quantique (notamment du bruit $1/f$) par une équation différentielle stochastique. Cette modélisation permet de capturer les corrélations temporelles du bruit tout en évitant une modélisation spectrale trop lourde, reposant sur de nombreux modes discrets. Le problème de contrôle ainsi obtenu peut être résolu de manière globale, produisant une commande optimale robuste à la réalisation du bruit, sans nécessiter d'ajustement a posteriori.

Conclusion

Ce chapitre propose une double avancée conceptuelle et numérique dans la synthèse robuste de portes quantiques. D'un côté, il met en lumière la difficulté intrinsèque du problème — à la fois par l'instabilité des équations de Pontryagin et par la complexité exponentielle des méthodes classiques. De l'autre, il offre deux cadres alternatifs : l'un semi-linéaire, utilisant une approximation variationnelle de l'EDP de Hamilton-Jacobi-Bellman ; l'autre stochastique, exploitant la dynamique du bruit quantique pour construire un algorithme robuste, global et scalable. Ces outils ouvrent la voie à une co-conception plus fine entre les modèles physiques du bruit quantique et les méthodes numériques d'optimisation.

Conclusion générale

Cette thèse s'inscrit dans le contexte de la simulation quantique et de la conception de portes robustes pour les ordinateurs quantiques à atomes neutres. Elle combine des outils d'analyse théorique, de contrôle optimal et de simulation numérique pour explorer des questions fondamentales et pratiques relatives à l'implémentation fiable de calculs quantiques en présence de bruit.

Le manuscrit commence par une revue critique de l'état de l'art, mettant en évidence les défis spécifiques aux architectures à atomes de Rydberg, notamment la sensibilité aux imperfections expérimentales, aux pertes et aux fluctuations lentes. Les chapitres suivants présentent la modélisation des portes quantiques, la conception de protocoles de contrôle robustes, et l'évaluation systématique des effets du bruit.

Une attention particulière a été portée à l'étude de séquences de contrôle réalistes. À partir de l'analyse d'un protocole de type CZ, un canal quantique bruité est obtenu de manière analytique, à partir d'hypothèses physiques sur les sources de perturbation. Ce canal est ensuite utilisé pour quantifier, via simulation numérique, l'impact du bruit dans la dynamique d'Hamiltoniens à plusieurs corps, dans une optique de simulation digitale. Les simulations, réalisées via les méthodes full statevector (jusqu'à 10 qubits) et MPO (jusqu'à 20 qubits), mettent en lumière la robustesse ou la fragilité de certaines configurations physiques.

Sur le plan méthodologique, la thèse explore différentes formulations du problème de synthèse de portes quantiques robustes. La difficulté intrinsèque du problème est illustrée par les limitations des approches de Pontryagin ou de bisection, en raison de leur instabilité ou de leur complexité exponentielle. Deux axes alternatifs sont alors proposés : d'une part, une reformulation variationnelle autour de l'équation de Hamilton-Jacobi-Bellman, approximée par des méthodes Monte-Carlo et des techniques de Newton ; d'autre part, un cadre stochastique rigoureux fondé sur des équations de Feynman-Kac non-linéaires, adaptées à la présence de bruit de type $1/f$.

En réunissant modélisation physique, techniques d'optimisation modernes et simulations à large échelle, ce travail contribue à l'effort de co-conception matériel-logiciel nécessaire à l'essor du calcul quantique. Il ouvre plusieurs perspectives, notamment en direction de stratégies hybrides de correction et de mitigation, intégrant les contraintes expérimentales dès la conception algorithmique.

Contenu

List of Figures

List of Tables

Chapter 1

Introduction

Sommaire

1.1 Historical perspective on quantum computing

1.1.1 Generalities

Quantum Computing (QC) is a recent paradigm leveraging quantum mechanics (QM) to process information in fundamentally novel programmable ways, ideally faster than with classical means in some areas of interest.It was originally proposed as an abstract thought experiment to understand the relationship between the computational power of a given device and its underlying physical theory. In 1981, Richard Feynman famously argued [1] that classical computers are fundamentally limited in their ability to simulate quantum systems efficiently. He argued that the measurement outcomes of a simple correlated-two-photons experiment could not be reproduced by any equally simple classical probabilistic procedure, because the shared amplitudes of the joint quantum state in the experiment intefere both 'constructively' and 'destructively', a tendency dubbed 'negative probability' by the author and unreproduceable by classical stochastic means. Indeed, this argument can be generalized to a typical wavefunction ψ of n interacting d-level quantum particles as:

1. The exact description of ψ fundamentally requires d^n coordinates, forbidding for large n an efficient classical determinstic calculation of quantum observables, i.e quantities of the form $\langle\psi|\hat{O}|\psi\rangle$ (this notation is explained later).

2. The synthesis of ψ features rapid oscillations and sign-flippings in additive amplitudes that prevent an efficient classical sampling of $|\psi|^2$, so any fast classical calculation of $\langle\psi|\hat{O}|\psi\rangle$ with Monte Carlo methods is typically infeasible. More information about this hardness can be found in [7]

To cope with this issue, Feynman proposed a universal quantum simulator: a controllable quantum device reproducing the dynamics of any fellow quantum system in a tunable way, collecting the relevant informations upon measurements. In 1985, David Deutsch expanded on this and other previous ideas (including Benioff's 1980 microscopic characterization of a computer as a quantum system [3] and his 1982 formulation of a quantum-mechanical Turing machine [2], as well Albert's 1983 quantum automata) by formulating a fully programmable, more general and robust model of a universal quantum computer [4]. Beyond individual exclusive bit values (0 or 1) and multivariate boolean functions met in classical computing, QC treats information as *qubits* (short for *quantum bits*), consisting in two-level quantum systems in linear superpositions of $|0\rangle$ and $|1\rangle$, and through logic gates resulting from the controlled physical evolution of various sites of the qubit hardware, and thus represented as unitary matrices.

1.1.2 Some quantum circuits

Building on Feynman's original insight that quantum systems should be efficiently simulatable by quantum machines, [28] provided the first formal proof that a universal quantum computer could indeed simulate any *local Hamiltonian* efficiently. More precisely, the goal of this quantum *Hamiltonian Simulation* (HS) algorithm is to approximate the time evolution of a quantum system governed by a Hamiltonian $\hat{H}$, as dictated by the Schrödinger equation $\imath\,|\dot{\psi}\rangle = \hat{H}\,|\psi\rangle$. This is achieved by concatenating a series of quantum gates such that the total unitary matrix of the circuit is $e^{\imath t\hat{H}}$ (in the time-dependent hamiltonian case, the operator exponent must be replaced by Magnus expansion terms), allowing a quantum processor to *reproduce the wavefunction of a given physical system*. Llyod's method relied on the *Trotter-Suzuki decomposition* to approximately split the time evolution operator into a sequence of easily implementable terms and applying them sequentially: $\hat{H} = \sum_j \hat{H}_j \implies e^{-\imath t\sum_j \hat{H}_j} \approx \left(\prod_j e^{-\imath(t/n)\hat{H}_j}\right)^n$, which is effective as long as the $e^{-\imath t\hat{H}_j}$ are naturally converted to simple quantum gates, e.g when the $\hat{H}_j$ are tensor products of Pauli matrices. However, this method introduces *Trotter error*, scaling in the step size and the number of such Pauli terms (which can theoretically be exponential). Two decades of improvements in hamiltonian encoding and approximate exponentiation, the major coneptual shift being the introduction of *quantum walks*, have culminated in the optimal HS algorithm based on *quantum signal processing (QSP)* and *qubitization* [13]. Based on these insights, quantum algorithms have also been proposed to prepare thermal states and compute thermal observables efficiently [30, 31], evading the fermionic sign problem [29], through *imaginary time evolution*. Another key application of HS is ground state preparation, attained through the *adiabatic algorithm*[14]: by running the HS algorihm along a certain path (generally slow) between two hamiltonians ($\hat{H}_0$, $\hat{H}_1$), starting from the ground state of $\hat{H}_0$ leads to ending in the ground state of $\hat{H}_1$. However, the choice of hamiltonian path must be careful in order to avoid small spectral gaps incurring a non-neglectible deviation from the target. Ground states can also be prepared through imaginary time evolution.

Beyond these rigorous algorithms, heuristic approaches based on *parameterized quantum circuits* have been explored to either approximate the evolution under a given hamiltonian through *variational simulation* in both real or imaginary time, or to approximately prepare the ground state of a hamiltonian, the main proposition there being the *Variational Quantum Eigensolver* (VQE). These methods attempt to parameterize quantum states using shallow quantum circuits, which are optimized through classical feedback. While promising for near-term devices, **these methods lack formal guarantees of accuracy and efficiency**, making them heuristic in nature. Nonetheless, variational simulation techniques offer a practical way to explore quantum dynamics with current hardware, especially in cases where deep circuits or

long coherence times are prohibitive.

Many fundamental NP-complete combinatorial and graph-theoretic problems can be formulated as instances of Ising hamiltonians [12], so that computing the ground state thereof yields the solution to the problem. This has motivated to tackle these problems using adiabatic quantum computing, imaginary time evolution (more marginally) and quantum annealing: a hardware-specific approximation to adiabatic computing. Much like the fixed time evolution and ground state preparation circuits have variational counterparts, there also exists a category of heuristic parameterized circuits to tackle combinatorial problems, in the form of the Quantum Approximate Optimization Algorithm (QAOA). Though it was originally tied to the adiabatic theory, it has since evolved as a specific topic with its own results.

Besides physical simulations, QC has found applications in other fields such as applied number theory, we refer to [6] for a rigorous and up-to-date survey of existing quantum algorithms. The main key to this success is that the algebraic structure of those number-theoretic problems is efficiently captured by the Quantum Phase Estimation (QPE) circuit, initially designed to compute eigenphases of unitary operators and that can be tweaked to compute the period of arithmetic functions, among other instances of the Hidden Subgroup Problem. Theoretically those algorithms may break many number-theoretic-cryptosystems as those are often based on algebraic equations. That is an important difference between applications of quantum computing: cryptosystems can still evolve to new standards (temporarily) unattainable for quantum computers; however the simulation of strongly-interacting quantum physical system is fundamentally hard no matter what classical algorithm is used and in this category of problems, QC is avoidable.

1.1.3 Limitations

The rising interest into QC has also led scientists to question the ability of quantum chips to actually perform those computations. An important counterpoint to the physical realization of viable quantum computers, is that the hardware is intrisically hard to control in a way that preserves the very quantum characteristics essential to its performance:

1. The theory behind the hardware control is based on a few assumptions with limited physical range of validity (e.g: qubit hardware is actually never a genuine two-level system, operations rely on intermediate excited states that can spontaneously decay to ground levels, exactly zero temperatures may not be attained, etc...)

2. Individual qubits are not entirely isolated from each other: in other words, actual physical single qubit gates can spread out to fellow qubits. Plus, by definition two-qubit gates entangle neighboring qubits (or even further qubits in flexible platforms!), yet those very interactions cannot be switched on and off with integral accuracy, so they may mess up logic gates by staying in force after the desired gate. Those pehnomenons are gathered under the term 'cross-talk'.

3. The quantum hardware cannot be an entirely closed physical system as controlling and measuring it introduces a small coupling with the environment, therefore it gradually loses its coherence throughout the logic gates.

4. The experimental setup can itself be slightly faulty (e.g external electromagnetic fields driving the qubit transitions can suffer from unwanted flucutations in frequency offset and electric power, or they can miss their target by a bit, thereby affecting other qubits)

Therefore, quantum errors are unavoidable and dramatically degrade the performance of the computations if left untreated. The first solution is to use a better and more precise hardware to physically enforce the theoretic assumptions, though this path suffers from the very physical limitations of the hardware (e.g in neutral-atomic qubits cannot be arbitrarily close to each other to enforce an integral blockade regime, lasers cannot be arbitrarily strong to engineer fast gates, etc). The second solution is to modulate the control signal to reduce the quantum noise, which **is the main focus on this thesis**. Indeed, through mathematical theories such as Optimal Control or Harmonic Analysis combined with perturbation theory, it is possible to strongly reduce the terms coming from violations of physical assumptions made on the qubit hardware, as long as those depend on the control. However, those two methods cannot integrally cancel the points raised below, so another approach is necessary. Worse, the redundancy protocols usually featured in classical denoising are not applicable to quantum states, due to the impossibility of duplicating quantum information implied by the no-cloning theorem [9]. This has led to formulate the specific field of *quantum error correction* [2] in which each logical qubit is encoded into an array of entangled physical qubits, errors are detected by subtle non-destructive partial measurements and corrected by local bit flips and phase flips. The field of clean quantum logic gates applied on encoded qubits is called *quantum fault-tolerance* (FT) theory. A cornerstone result of FT and QEC theory is the *threshold theorem*, which states that assuming the qubit errors to be distributed spatially locally and weakly correlated in time, if the quality of each quantum gate is above a constant threshold particular to the QEC code, then a repeated embedding of this codes allows to implement the quantum logic gates with an arbitrarily low failure rate.

1.2 Neutral atoms as quantum hardware

The first quantum computing hardware setup made of neutral atoms in optical lattice was formally proposed in [17, 18]. Indeed, arrays of neutral atoms can be trapped in optical tweezers in a process called 'laser cooling' [18, 19], which has undercome tremendous progress in the last decades, up to enabling precise spatial arrangements of hundreds or even thousands of atoms. These arrays are highly flexible, supporting a wide range of geometries [22], including 1D, 2D, 3D configurations. Additionally, beyond the trapping fields, individual atoms can be addressed using tightly focused laser beams [18] or collective microwave fields [19], allowing for transitions and superposition of optical atomic levels, selective initialization, manipulation, and readout of qubits through fluorescence spectroscopy [19, 29]. It is common to use hydrogen-like atoms with simple structure of alkali or alkali-earth types, drawn from the first two columns of the periodic table (e.g., rubidium or cesium); as their theoretical treatment is close to the analytically known case of hydrogen.

Leveraging the Rydberg blockade effect was first hinted in [5]. Indeed, Rydberg-interacting neutral atoms represent a promising platform for quantum computing, combining scalability, high coherence times, and strong, tunable interactions [19]. The **Rydberg blockade** is an essential mechanism for implementing entangling two-qubit operations in neutral atom quantum computing. A key feature of this platform is the use of Rydberg states to mediate strong, long-range interactions between qubits. When a neutral atom is excited to a high-energy Rydberg state, the valence electron's distance from the nucleus generates a large dipole moment [23] This creates strong dipole-dipole or van der Waals interactions that can span several micrometers—ideal for coupling qubits separated in optical traps. These interactions are highly tunable, as their strength can be controlled by adjusting the Rydberg state's principal quantum number, the distance between atoms, or external electromagnetic fields. This tunability enables precise control over the interaction dynamics, which is crucial for implementing entangling quantum gates and generating entangled states.

Rydberg interactions form the basis of entangling two-qubit gates such as the controlled phase flip, which are essential for universal quantum computing; while single qubit gates are carried by engineering multi-photon Rabi oscillations with lower excited states). These gates rely on the Rydberg blockade effect, where the excitation of one atom to a Rydberg state shifts the energy levels of neighboring atoms, preventing their simultaneous excitation [5]. This mechanism enables deterministic and fast two-qubit operations, with gate fidelities exceeding 99% in experimental setups. Combined with robust single-qubit gates and efficient qubit initialization, this platform demonstrates significant progress toward fault-tolerant quantum computation.

A decisive standing advantage of the Rydberg-interacting neutral-atomic platform is its inherent stability, declined in three characteristics. First, neutral atoms enjoy a large inherent stability and long coherence times: computational basis states are typically encoded in the hyperfine electronic ground states of neutral atoms [19], which are less susceptible to environmental noise than the hardware platforms storing at least one logical bit state in a physical excited state, such as superconducting qubits. The shorter-lived intermediate states necessary for gate operations are only occupied transiently during gate operations, leaving the qubits' overall coherence dominated by the long-lived ground states. Second, a tangible advantange of atomic platforms is their ability to work at room temperature, since the atoms are already 'cooled' by the optical apparatus, thereby showcasing their quantum powers. Thus, in principle they do not need the industrial cryogenic apparatus common in other quantum hardware platforms, though it is still sometimes used to enforce vacuum around the atoms to minimize noise sources. Third, despite being highly excited, the Rydberg orbitals intervening in multi-qubit entangling gates, have relatively high lifetimes [23, 21, 23] compared to lower excited other orbitals. A recent trend in this field has been to excite atoms to *circular* Rydberg states which showcase even longer lifetimes [22]. Another advantage of this paltform is that the atoms can be moved by the optical apparatus, e.g to perform entangling gates between far-away qubits: this motional flexibility promises an effective all-to-all connectivity (common to trapped-ions system too) as opposed to 'solid-state' hardware platforms marked by a physically local qubit connectivity such as semi-conductor dots or superconducting loops.

Recent experimental advances have further demonstrated the platform's viability. Researchers have achieved defect-free 2D arrays of neutral atoms using optical tweezers and rearrangement algorithms, producing systems with nearly perfect filling. High-fidelity entanglement generation between atom pairs and clusters has also been demonstrated, leveraging the strong interactions provided by Rydberg states. Large arrays of atoms have been used to simulate many-body hamiltonians [29]. These developments are supported by ongoing progress in laser systems, cooling techniques, and error correction strategies, which continue to enhance gate fidelities and reduce noise. Quantum gates with high fidelities have been physically implemented [25], and mechanisms required for fault-tolerance such as magic states distillation have also been implemented [26].

Despite its promise, the platform faces challenges that remain active areas of research. Rydberg transitions require high-power, narrow-linewidth lasers, which are technologically demanding. Residual atomic motion, even in ultracold systems, introduces errors in gate operations, motivating the development of advanced cooling meth-

ods such as sub-recoil cooling. Another major drawback of this platform is that the current gate times are quite slow compared to other quantum hardware platforms such as superconducting loops. Fortunately, along with the development in hardware precision technology, those shortcomings can be answered with quantum error correction, and QEC codes specific to the neutral-atomic platform have been proposed. [27, 28, 25]

1.3 Problem and relevance

This thesis raises some important questions :

1. How does the design of multi-qubit gates require optimization in the first place ?
2. Can one identify controllable elements in this task ?
3. How exclusive is our optimization framework to the underlying mathematical model of Rydberg atoms ? Can it be generalized to other hardware platforms ?
4. What are the general specificities of artificial evolution algorithms ?
5. Are they relevant to optimize the design of quantum gates ?
6. Does this task really require the subtelties of those algorithms ?
7. In other words, can artificial evolution lead to better control of quantum technologies ?

1. Both the technical progress in the precise manipulation and control of quantum hardware and the theoretical discovery of quantum algorithms asymptotically solving classically intractable problems, have made QC a highly promising paradigm of computation. However, as already said, nanoscopic quantum systems are notoriously hard to control faithfully, chiefly because they do not genuinely constitute closed two-level quantum systems: their small coupling to the environment (due to small fluctuations in lab conditions, nonzero temperature, faulty control, or the unavoidable entanglement between the computing hardware and the control field) progressively break their quantum nature, a process known as *decohence*, hampering their computational power. Another source of disturbances is that the reliable application of quantum logic operations rests on physical assumptions that are verified only asymptotically, and whose violations in actual conditions ultimately induce computational errors if left unchecked. This is why quantum hardware needs first a physical optimization in topics such as the choice and quality of control apparatus, and the spatial arrangement of the hardware, in order to minimize the unwanted coupling between the hardware and its environment and the violation of physical assumptions. However, these adjustments are made before starting any computation cycle, and are obviously fundamentally limited by physical

or even cost constarints: one only has lasers with fixed maximal power and bandwidth, and qubits cannot be magically rearranged to somehow cancel crosstalk noise, etc. These limitations would be answered if another adjustable characteristic of the quantum hardware gives more degrees of freedom to the experimentalist. This leads to the first question raised by this thesis: **How does the design of multi-qubit gates require a dynamic optimization of the control setup in the first place ?**

2. As already mentioned, QC hardware is feasible only in particular lab conditions and is the result of a very careful human intervention on physical phenomena. Therefore, the physical theories underlying the QC hardware make room for a reasonable collection of adjustable parameters whose modifications sensibly changes the quantum states of the hardware, as is visible upon measurement. Yet not all of those parameters can be considered as dynamic variables. This immediately leads to the second question: *Can one identify dynamical controllable elements in the task of optimizing quantum gates?*

3. Though we apply optimization algorithms on quantum gates adapated to the hardware platform of neutral atoms, this present work is not concerned with their deepest physical subtelties and we adopt simplistic physical models that would be prone to generalization to other setups. But since the field is still in its infancy, the QC hardware landscape is full of vastly differing technologies derived from various physics and harboring distinct advantages and inconvenients; contrary to the case of classical computing where all the chips are built in similar materials along the qsame physics. This raises the questions: **How exclusive is our optimization framework to the underlying mathematical model of Rydberg atoms ? Can it be generalized to other hardware platforms ?**

4. As the name suggests, artificial evolutions is a class of algorithms seeking to optimise functions by leveraging the biological theories of evolution and natural selection. Yet those are not their only characteristics, as optimization algorithms come in all shapes and colors and leverage various features of the functions sought to optimize, differing accross algorithms. This raises the question: **What are the general specificities of artificial evolution algorithms ?**

5. As hinted by the previous arguments, the sensitivity of quantum hardware to noise, faulty setups and decoherence has laid the theory of *optimal control* (OC) to be a chief concern in quantum technologies globally, finding applications e.g from quantum sensing to quantum computing [1] and widely believed to play a leading role towards reliable quantum gates. However, this deterministic theory of OC makes certain assumptions differing from the setups of the more versatile artificial evolution algorithms. This leads to questioning whether AE algorithms are adapted to the structure of quantum control problems. **Are they relevant to optimize the design of quantum gates ?**

6. AE algorithms do not standout only because of their original structure or parallel mechanisms, but also because of their computational and memory requirements that

differ from standard optimal control. This raises the questions : *Does this task of optimizing quantum gates really require the subtelties of AE algorithms ?* A related problem is, given the context in which AE algorithms are sought to be applied, can one somehow relax the conditions of optimal control to narrow the applicability gap between the theories ?

7. Summarizing those questions, an investigation of control problems in quantum hardware is necessary to assess wether AE can help there. Beyond that, it is also important to question the problem structures and the assumptions of various algorithms seeking to solve them, as well as their computational hurdle, so in shorter terms: **can artificial evolution lead to better control of quantum technologies** ? This thesis is the journey accross those problems, from applying AE to a noisy quantum control problem to a questioning of AE's applicability in the field by assessing the assumptions of previously overlooked theories.

1.4 Proposition/contribution

The spirit of this thesis is to explore non-conventional numerical methods to optimize existing gate protocols.

- Chapter 2 introduces the reader to a historical perspective on quantum computing (QC) and its applications, an brief overview of the neutral-atomic hardware platform, basic concepts in QC and an elementary derivation of simple quantum gates with Rydberg atoms.
- Chapter 3 introduces the important optimization concepts of artificial evolution and multi-objectoive optimization, which are featured in the next chapter.
- In Chapter 4, we propose to apply artificial evolution to minimize the effects of stochastic control noise on the synthesis of quantum gates.
    1. We start by applying CMA-ES to maximize the fidelity at a fixed level of intensity and frequency noise.
    2. Then, we use NSGA-III to minimize simultaneously the gate infidelity and the time spent in the Rydberg state, yielding a Pareto Front of candidate solutions.
- Chapter 5 introduces the reader to the classical emulation of quantum systems through Matrix Product States.
- In Chapter 6, we consider a fixed gate setup evolved through a stochastically-noisy hamiltonian. Three physical noise mitigation schemes are described and compared accross several generic metrics computed through the noisy simulation

of two quantum Ising-models, having derived the relevant local quantum noise channel for doing so. For each model we identify the best noise compensation scheme and relate the performance to the noise nature.

- Chapter 7 is a collection of outlook perspectives for the analytical framework of value functions and dynamic programming in the design of robust quantum gates.

    1. First, we propose a method based on the vanishing viscosity limit of Hamilton-Jacobi equations to approximately initialize the dynamic Lagrange multipliers in optimal control, through Monte Carlo approximations.
    2. Second,we propose to apply a recent dynamic programming algorithm based on second-order HJB equations, Poisson jump processes and stochastic differential equiations of pink noise to derive robust gates without going through any further optimization, while minimizing any further root-finding and outputting the global minimum of the control problem.

- Chapter 8 concludes the thesis, giving our final word about the usage of artificial evolution in quantum computing.

Bibliography

[1] Feynman, R.P. Simulating physics with computers. Int J Theor Phys 21, 467–488 (1982). https://doi.org/10.1007/BF02650179, https://s2.smu.edu/~mitch/class/5395/papers/feynman-quantum-1981.pdf (Cited on page 2.)

[2] Benioff, P. Quantum mechanical hamiltonian models of turing machines. J Stat Phys 29, 515–546 (1982). https://doi.org/10.1007/BF01342185 (Cited on page 2.)

[3] Benioff, P. The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. J Stat Phys 22, 563–591 (1980). https://doi.org/10.1007/BF01011339 (Cited on page 2.)

[4] Deutsch, D., 1985. Quantum theory, the Church–Turing principle and the universal quantum computer. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 400(1818), pp.97-117. https://doi.org/10.1098/rspa.1985.0070 (Cited on page 2.)

[5] Lloyd, S., 1996. Universal quantum simulators. Science, 273(5278), pp.1073-1078. (Cited on page 3.)

[6] Jordan, S. (2025, March 5). Quantum Algorithm Zoo. https://quantumalgorithmzoo.org/ (Cited on page 4.)

[7] Hangleiter, D., 2021. Sampling and the complexity of nature: Assessing, testing, and challenging the computational power of quantum devices (Doctoral dissertation). https://arxiv.org/abs/2012.07905 (Cited on page 2.)

[8] Di Meglio, A., Jansen, K., Tavernelli, I., Alexandrou, C., Arunachalam, S., Bauer, C.W., Borras, K., Carrazza, S., Crippa, A., Croft, V. and De Putter, R., 2024. Quantum computing for high-energy physics: State of the art and challenges. PRX Quantum, 5(3), p.037001. (Not cited.)

[9] Tan, K.C., Bhowmick, D. & Sengupta, P. Sign-problem free quantum stochastic series expansion algorithm on a quantum computer. npj Quantum Inf 8, 44 (2022). https://doi.org/10.1038/s41534-022-00555-x (Cited on page 3.)

[10] Motta, M., Sun, C., Tan, A.T.K. et al. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. Nat. Phys. 16, 205–210 (2020). https://doi.org/10.1038/s41567-019-0704-4 (Cited on page 3.)

[11] Jouzdani, P., Johnson, C.W., Mucciolo, E.R. and Stetcu, I., 2022. Alternative approach to quantum imaginary time evolution. Physical Review A, 106(6), p.062435. (Cited on page 3.)

[12] Lucas, A. (2014). Ising formulations of many NP problems. Frontiers in physics, 2, 5. https://doi.org/10.3389/fphy.2014.00005 (Cited on page 4.)

[13] Low, G. H., & Chuang, I. L. (2019). Hamiltonian simulation by qubitization. Quantum, 3, 163. https://doi.org/10.22331/q-2019-07-12-163 (Cited on page 3.)

[14] Albash, T. and Lidar, D.A., 2018. Adiabatic quantum computation. Reviews of Modern Physics, 90(1), p.015002. https://link.aps.org/accepted/10.1103/RevModPhys.90.015002 (Cited on page 3.)

[15] "Numerically optimized bosonic code", The Error Correction Zoo (V. V. Albert and P. Faist, eds.), 2022. https://errorcorrectionzoo.org/ (Cited on page 133.)

[16] Lidar, D.A. and Brun, T.A. eds., 2013. Quantum error correction. Cambridge university press. (Cited on pages 5 and 133.)

[17] Brennen, G. K., Caves, C. M., Jessen, P. S., & Deutsch, I. H. (1999). Quantum logic gates in optical lattices. Physical Review Letters, 82(5), 1060. (Cited on page 6.)

[18] Deutsch, I. H., Brennen, G. K., & Jessen, P. S. (2000). Quantum computing with neutral atoms in an optical lattice. Fortschritte der Physik: Progress of Physics, 48(9-11), 925-943. (Cited on page 6.)

[19] M. Morgado, S. Whitlock; Quantum simulation and computing with Rydberg-interacting qubits. AVS Quantum Sci. 1 June 2021; 3 (2): 023501. https://doi.org/10.1116/5.0036562 (Cited on pages 6, 7, 21, 23, 24 and 26.)

[20] Browaeys, A., Lahaye, T. Many-body physics with individually controlled Rydberg atoms. Nat. Phys. 16, 132–142 (2020). https://doi.org/10.1038/s41567-019-0733-z (Not cited.)

[21] Henriet, L., Beguin, L., Signoles, A., Lahaye, T., Browaeys, A., Reymond, G.O. and Jurczak, C., 2020. Quantum computing with neutral atoms. Quantum, 4, p.327. https://doi.org/10.22331/q-2020-09-21-327 (Cited on pages 6, 7 and 26.)

[22] Kim, M., Ahn, J., Song, Y. et al. Quantum computing with Rydberg atom graphs. J. Korean Phys. Soc. 82, 827–840 (2023). https://doi.org/10.1007/s40042-023-00774-1 (Cited on page 6.)

[23] Gallagher, T.F. (1994) Rydberg Atoms. Cambridge University Press. https://doi.org/10.1017/cbo9780511524530 (Cited on pages 6, 7 and 25.)

[24] Labuhn, H., Barredo, D., Ravets, S. et al. Tunable two-dimensional arrays of single Rydberg atoms for realizing quantum Ising models. Nature 534, 667–670 (2016). https://doi.org/10.1038/nature18274 (Not cited.)

[25] Jandura, S. and Pupillo, G., 2024. Surface code stabilizer measurements for Rydberg atoms. arXiv preprint arXiv:2405.16621. (Cited on page 8.)

[26] Koch, C.P., Boscain, U., Calarco, T. et al. Quantum optimal control in quantum technologies. Strategic report on current status, visions and goals for research in Europe. EPJ Quantum Technol. 9, 19 (2022). https://doi.org/10.1140/epjqt/s40507-022-00138-x (Cited on pages 9 and 106.)

[27] Di Meglio, A., Jansen, K., Tavernelli, I., Alexandrou, C., Arunachalam, S., Bauer, C.W., Borras, K., Carrazza, S., Crippa, A., Croft, V. and De Putter, R., 2024. Quantum computing for high-energy physics: State of the art and challenges. PRX Quantum, 5(3), p.037001. (Not cited.)

[28] Lloyd, S., 1996. Universal quantum simulators. Science, 273(5278), pp.1073-1078. (Cited on page 3.)

[29] Tan, K.C., Bhowmick, D. & Sengupta, P. Sign-problem free quantum stochastic series expansion algorithm on a quantum computer. npj Quantum Inf 8, 44 (2022). https://doi.org/10.1038/s41534-022-00555-x (Cited on page 3.)

[30] Motta, M., Sun, C., Tan, A.T.K. et al. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. Nat. Phys. 16, 205–210 (2020). https://doi.org/10.1038/s41567-019-0704-4 (Cited on page 3.)

[31] Jouzdani, P., Johnson, C.W., Mucciolo, E.R. and Stetcu, I., 2022. Alternative approach to quantum imaginary time evolution. Physical Review A, 106(6), p.062435. (Cited on page 3.)

Chapter 2

Overview of quantum computing with cold atoms

Sommaire

This chapter briefly describes the usage of neutral atoms in optical lattices as quantum computing hardware, in order to understand the rest of the thesis. Section 2.1 gives a more general introduction to quantum computing and its motivation. Section 2.2 presents the arguments vindicating the neutral-atomic hardware platform. Section 2.3 gives a brief account of atom optics and section 2.4 explains the Rydberg interactions, ending with the description of two-qubit gates on this platform. Section 2.5 lists the optimization challenges related to the design

2.1 Some fundamental concepts in quantum mechanics

The mathematical formulation of Quantum Mechanics (QM) sits on four postulates, which provide the theoretical framework for quantum computing defining the nature of quantum states, their evolution, measurement, and the behavior of composite systems. We refer to [9] for intuitive explanations of these postulates.

2.1.1 Postulate 1: vector description of quantum physical systems

2.1.1.1 Closed system

Let $\mathcal{S}$ be a quantum physical system. According to QM, any physical quantity characteristic of $\mathcal{S}$ can be observed in at most m distinct configurations, where m is either a finite natural integer or the countable infinity $+\infty$. The first postulate of QM states that all the informations relevant to the analysis of $\mathcal{S}$ derive from objects of a *separable complex Hilbert space* $\mathcal{H}$, equipped with a complex-valued inner product $\langle\cdot|\cdot\rangle$ (so that $\langle\psi|\psi\rangle = 1$), and a countable orthonormal basis $\{|n\rangle \mid n \in \mathbb{N}\}$. This simply means that, up to vector space isomorphisms:

- If $m < \infty$ then $\mathcal{H} = \mathbb{C}^m$
- If $m = \infty$ then $\mathcal{H} = \{u \in \mathbb{C}^{\mathbb{N}} \mid \sum_{n\geq 0} |u_n|^2 < +\infty\}$

Expressions in this formalism are simplified by the *Dirac notation*:

- The *ket* $|\psi\rangle = \sum_{n\geq 0} \psi_n |n\rangle$ refers to a column vector.
- The *bra* $\langle\phi| = \sum_{n\geq 0} \psi_n^* \langle n|$ refers to a row vector.
- The *braket* $\langle\phi|\psi\rangle = \sum_{n\geq 0} \phi_n^* \psi_n$ refers to the standard complex inner product.
- The *ketbra* $|\phi\rangle\langle\psi| : |u\rangle \to \langle\phi|u\rangle |\psi\rangle$ refers to the standard outer product.

When $\mathcal{S}$ is a *closed* system (i.e not interacting with its environment) then $\mathcal{S}$ is said to be in a *pure state*, completely described by a unit vector $|\psi\rangle \in \mathcal{H}$ (i.e $\langle\psi|\psi\rangle = 1$).

In realistic settings, $\mathcal{S}$ is not fully isolated and is rather an *open* quantum system interacting with its environment, called a *bath* and noted $\mathcal{B}$. Then $\mathcal{S}$ is said to be in a *mixed state*, completely described by a *density operator*, i.e an element $\hat{\rho} \in \mathcal{L}(\mathcal{H})$ that is hermitian (i.e $\rho^\dagger = \rho$), semi-definite positive ($\forall \psi \in \mathcal{H}$, $\langle\psi|\rho|\psi\rangle \geq 0$) with unit trace (i.e $\mathrm{Tr}(\rho) = 1$). Owing to the spectral theorem, it can be decomposed as a convex combination of orthogonal pure states:

$$\hat{\rho} = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \quad p_{i+1} \geq p_i \geq 0, \; \sum_i p_i = 1$$

The degree of 'mixedness' of $\mathcal{S}$ is then expressed by the *purity* $\mathrm{Tr}\left(\hat{\rho}^2\right) = \sum_i p_i^2 \in [0,1]$: if $\mathrm{Tr}\left(\hat{\rho}^2\right) = 1$ then $p_1 = 1$, $\hat{\rho} = |\psi_1\rangle\langle\psi_1|$.

If $\mathcal{S}$ is a closed system with stochastic internal dynamics, then its statevector ψ evolves randomly, and its statistical properties are captured by the average state $\hat{\rho} := \mathbb{E}(|\psi\rangle\langle\psi|)$ that is a density matrix too. Thus the formalism of open quantum systems allows to study and control the average properties of stochastically evolving closed systems, e.g when the control signal is perturbed as in chapters 5 and 6. Conversely, the deterministic evolution of an quantum system can be emulated as the evolution of a stochastic closed system, from where many paths are sampled: this is the method of *quantum trajectories* studied in [2].

2.1.2 Postulate 2: evolution of quantum systems

The energy interactions in a closed system $\mathcal{S}$ are described by a hermitian operator $\hat{H}_\mathcal{S} \in \mathcal{L}(\mathcal{H})$ called the *hamiltonian*. The second QM postulate states that $\mathcal{S}$ evolves through the Schrödinger equation $\imath|\dot{\psi}\rangle = \hat{H}_\mathcal{S}|\psi\rangle$. This can be reformulated through a unitary operator $U : \mathbb{R} \to \mathcal{U}(\mathcal{H})$, called the 'propagator' of $\hat{H}$, such that $|\psi(t)\rangle = U(t)|\psi(0)\rangle$, and the Schrödinger equation is restated as $U(0) = I \; \imath\dot{U}(t) = \hat{H}_\mathcal{S} U(t)$.

The evolution of an open system $\mathcal{S}$ interacting with an external bath $\mathcal{B}$ is generically described by by a *master equation* of the form:

$$\dot{\hat{\rho}}(t) = -\imath[\hat{H}(t), \hat{\rho}(t)] + \mathcal{D}'((\rho(s))_{s\in[0,t]})$$

Where $[\hat{A}, \hat{B}] :\mapsto \hat{A}\hat{B} - \hat{B}\hat{A}$ is the *commutator*, and $\mathcal{D}'$ is the *dissipator*, a linear superoperator such that $\mathrm{Tr}\left(\mathcal{D}'((\rho(s))_{s\in[0,t]})\right) = 0$, depending on the internal bath hamiltonian $\hat{H}_\mathcal{B}$ and the system-bath hamiltonian $\hat{H}_{\mathcal{S}-\mathcal{B}}$. The dissipator introduces both non-unitarity and memory effects as the instantaneous evolution depends on the system's past states, which are absorbed by the bath and reinjected into the system. For this reason, the dissipator often involves a *memory kernel* K through an integral term

$\int_0^t K(s,t)\hat{\rho}(s)ds$, making the master equation a linear integro-differential equation. The past memory dependence usually makes the master equation intractable at first sight and requires a few approximations to construct a tractable memory kernel, as in the projection method of the Nakajima-Zwanzig equation [4, 5].
One gets a much simpler approximation when the system-bath dynamics satisfy the following conditions:

- Weak system-bath coupling
- Bath being much larger than the system
- Internal bath dynamics much slower than the internal system evolution
- Bath being in thermal equilibrium

Then [3] the memory effects can be neglected, the dissipator depends only the current system state and can be reduced to a *Lindbladian* $\mathcal{D}'((\rho(s))_{s\in[0,t]} = \mathcal{D}(\rho(t))$ of the form:

$$\mathcal{D}(\rho) = \sum_k L_k \rho L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho\}$$

Where $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$ is the anticommutator and L_k are Lindblad operators modelling dissipation and decoherence. This reduces the master equation to a tractable ODE:

$$\dot{\hat{\rho}}(t) = -\imath[\hat{H}(t), \hat{\rho}(t)] + \sum_k L_k \rho(t) L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho(t)\}$$

More generally, beyond unitary maps (acting on density matrices as $\hat{\rho} \mapsto U\hat{\rho}U^\dagger$), the evolution of open quantum systems is expressed by *completely positive trace-preserving* (CPTP) linear maps (e.g linear maps $\mathcal{G}$ that preserve the property of density matrices), also known as *quantum channels*. Through an extension of the spectral theorem, for any such $\mathcal{G}$ there exists *Kraus operators* (K_m) such that

$$\forall \hat{\rho},\ \mathcal{G}(\hat{\rho}) = \sum_m K_m \hat{\rho} K_m^\dagger \quad \text{where} \sum_m K_m^\dagger K_m = I$$

2.1.3 Postulate 3: Composition of quantum systems

If $\mathcal{S} = \mathcal{S}_0 \cup \mathcal{S}_1$ where $\mathcal{S}_0, \mathcal{S}_1$ are two disjoint ($\mathcal{S}_0 \cap \mathcal{S}_1 = \varnothing$) systems with respective state spaces $\mathcal{H}_0, \mathcal{H}_1$, then the third QM postulate states that $\mathcal{H} = \mathcal{H}_0 \otimes \mathcal{H}_1$ where $\otimes$ is the *tensor product*, known in finite dimension as the *Kronecker product*. *Entanglement* is the propension of the state of $\mathcal{S}$ to not be decomposable as a simple tensor product of vectors in each component space: mathematically, $\mathcal{S}$ is in an entangled state $|\psi\rangle \in \mathcal{H}$ if $\forall\, |\psi^b\rangle \in \mathcal{H}_b\ (b \in \{0,1\})$, $|\psi\rangle \neq |\psi^0\rangle \otimes |\psi^1\rangle$. The *Schmidt decomposition* theorem states

the existence of decreasing 'Schmidt values' $\lambda_\alpha \in \mathbb{R}^+$ and orthogonal *Schmidt vectors* $|u^b_\alpha\rangle \in \mathcal{H}_b$ ($\langle u^b_\alpha | u^b_\beta \rangle = \mathbb{1}_{\alpha=\beta}$) such that:

$$|\psi\rangle = \sum_\alpha \lambda_\alpha |u^{[0]}_\alpha\rangle \otimes |u^{[1]}_\alpha\rangle$$

By the orthonormality of the Schmidt vectors, the squared Schmidt values are normalized: $\sum_\alpha \lambda^2_\alpha = 1$, relating them to a probability distribution, whose Shannon information entropy is the *Von Neumann entropy*:

$$S^{\mathrm{V}} := -\sum_\alpha \lambda^2_\alpha \log_2(\lambda^2_\alpha)$$

S^{V} formally quantifies the entanglement building in $\mathcal{S}$ between $\mathcal{S}_0$ and $\mathcal{S}_1$. A relatively rich category of physically relevant quantum systems satisfy a polynomial relationship between their entanglement entropy and the surface of their boundary, a phenomenon known as area scaling laws [2]. This motivates an emulation procedure based on the Schmidt decomposition to take advantage of such entanglement-related bounds. The 'Schmidt rank' χ is the minimal number of nonzero Schmidt values and satisfies $S^{\mathrm{V}} \leq \log_2(\chi)$, $2^S \leq \chi$, making it a fundamental upper bound on the entanglement entropy, so the evolution of S^{V} roughly quantifies the Schmidt rank necessary to efficiently represent ψ.

Similarly, density operators of joint open systems admit a Schmidt decomposition too, but with orthonormal 'Schmidt vectors' density matrices $\rho^b_\alpha \in \mathcal{H}_b$ ($b \in \{0,1\}$, $\mathrm{Tr}\big(\rho^b_\alpha \rho^b_\beta\big) = \mathbb{1}_{\alpha=\beta}$) i.e decreasing positive 'Schmidt values' (λ_α):

$$\frac{\hat{\rho}}{\mathrm{Tr}(\hat{\rho}^2)} = \sum_\alpha \lambda_\alpha \rho^0_\alpha \otimes \rho^1_\alpha$$

The *Operator Space Entanglement Entropy* of $\mathcal{S}$ between $\mathcal{S}_0$ and $\mathcal{S}_1$ is the Shannon entropy of those Schmidt values:

$$S^{\mathrm{O}}(\hat{\rho}) = -\sum_\alpha \lambda^2_\alpha \log_2(\lambda^2_\alpha)$$

In the pure state case, it is related to the Von Neumann entropy through $S^{\mathrm{V}}(|\psi\rangle) = \frac{1}{2} S^{\mathrm{O}}(|\psi\rangle\langle\psi|)$. Generally in mixed states S^{O} does not quantify the actual entanglement per se and is simply an upper bound on it. For reference, a genuine measure of entanglement of mixed states is the *entanglement of formation* described in [8]:

$$S^{\mathrm{F}}(\hat{\rho}) := \inf\left(\sum_i p_i S^{\mathrm{V}}(|\psi_i\rangle) \,\middle|\, \hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|,\ p_i \geq 0,\ \sum_i p_i = 1 \right)$$

2.1.4 Postulate 4: Measurement in quantum systems

Measurement in quantum mechanics is described by a set of operators $\{M_k\}$ such that $\sum_k M_k^\dagger M_k = I$, and consists in applying them on a quantum state $|\psi\rangle$ with some probability expressed by the *Born rule*:

$$|\psi\rangle \to \frac{M_k|\psi\rangle}{\langle\psi|M_k^\dagger M_k|\psi\rangle} \quad \text{with probability} \quad \langle\psi|M_k^\dagger M_k|\psi\rangle$$

In particular, the full measurement in the computational basis is represented by the operators $M_k = |k\rangle\langle k|$, yielding $|k\rangle$ (up to a meaningless global phase) with probability $|\langle\psi|k\rangle|^2$.
The measurement of an open quantum system works on similar grounds:

$$\hat{\rho} \mapsto \frac{M_k\hat{\rho}M_k^\dagger}{\mathrm{Tr}\Big(M_k\hat{\rho}M_k^\dagger\Big)} \quad \text{with probability} \quad \mathrm{Tr}\Big(M_k\hat{\rho}M_k^\dagger\Big)$$

In particular, the full computational basis measurement is represented by the operators $M_k = |k\rangle\langle k|$, yielding $|k\rangle\langle k|$ with probability $\langle k|\rho|k\rangle$.

2.2 Atomic levels and two-level atom-light interactions

2.2.1 Free atom hamiltonian

The quantized free orbit of the valence electron orbiting the nucleus in a hydrogen-like atom is summarized by the *free atom hamiltonian* $\hat{H}_0 = \hat{H}_p + \hat{H}_{\mathrm{fs}} + \hat{H}_{\mathrm{hfs}}$ where:

- $\hat{H}_p$ is the principal part accounting for the electronic kinetic energy and is the nucleus-electron Coulomb electrostatic potential [8].
- $\hat{H}_{\mathrm{fs}}$ is the *fine structure* hamiltonian, whose magnitude is much lower than the previous term, made of:
  - The first-order kinetic energy relativistic correction
  - Spin-orbit coupling
  - Darwin term
  - Lamb shift
- $\hat{H}_{\mathrm{hfs}}$ is the *hyperfine structure* hamiltonian, made of:
  - The interaction between the *nuclear magnetic dipole* and respectively the *electron orbital* and *electron spin* magnetic fields
  - Nuclear electric quadruopole

The approximative diagonalization results in a free atom hamiltonian of the form

$$\hat{H}_0 = \sum_{n=1}^{+\infty} \sum_{l=0}^{n-1} \sum_{m_l=-l}^{l} \sum_{m_s=-s}^{s} E_{n,l,m_l,m_s} \, |n,l,m_l,m_s\rangle\langle n,l,m_l,m_s| \quad E_{n,l,m_l,m_s} = \frac{R}{n-\delta} + \frac{Q(l,m_l,m_s)}{(n-\delta)^3}$$

Where R is the Rydberg constant, δ is the *quantum defect*, Q is a multivariate second-degree rational fraction, and the eigenelements are indexed by the *quantum numbers* (QN):

- $n \in \mathbb{N}^*$ is the *radial/principal* QN.
- $l \in [\![0, n-1]\!]$ is the *azimuthal* QN.
- $m_l \in [\![-l, l]\!]$ is the *magnetic* QN.
- $m_s \in [\![-s, s]\!]$ is the *magnetic* QN. For electrons: $s = 1/2$.

The atomic spectral rays can be further split by external magnetic fields through the Zeeman effect, or external electric fields through the Stark effect [8]. Since $\hat{H}_p \gg \hat{H}_{\text{fs}} \gg \hat{H}_{\text{hfs}}$, the eigenfunctions and the dominant part in n of the energy levels are derived by properly diagonalising $\hat{H}_p$, while the other terms in the energy levels are derived from the approximately additive contribution of $\hat{H}_{\text{fs}}, \hat{H}_{\text{hfs}}$ in perturbation theory. Numerical values for all those constants are given in [8]. The key point is that $\hat{H}_0$ is *anharmonic*: the level transition frequencies are not linearly distributed, and this means that with the correct hardware calibration and pulse engineering, it is possible to reduce it to a two-level system as is refered below, enabling it to serve as quantum computing hardware. In the rest of this work, to simplify the numerical analysis we assume that qubit states are stored in twin hyperfine electronic ground states, preferably $(|n=1,l=0,m_l=0,m_s=-1/2\rangle, |n=1,l=0,m_l=0,m_s=1/2\rangle)$, the most stable orbitals. However, in heavy atoms those orbitals might be saturated, so the logical bit states will instead be stored in the next lowest energy states possible, from where they won't decay to the already occupied lower-energy states (which cannot be occupied by two distinct electrons due to Pauli's exclusion principle), so the arguments for hyperfine ground states still hold. Also, in some criteria it can be a better idea to instead encode one or both logical bit states in Rydberg orbitals [19].

2.2.2 Electric field and Atom-light hamiltonian and level-transitions

In classical physics, a particle of position $\boldsymbol{r}$ and electric charge q admits an electric dipole moment $\boldsymbol{\mu} = q\boldsymbol{r}$, that when interacting with a nearby electric field $\boldsymbol{E}(t,\boldsymbol{r})$, induces an electric energy $\boldsymbol{\mu} \cdot \boldsymbol{E}(t,\boldsymbol{r})$. In the quantum setting, the EDM is $\hat{\boldsymbol{\mu}} = q\hat{\boldsymbol{r}}$ where $\hat{\boldsymbol{r}}$ is the position operator. The field's spatial dependence is neglected as its wavelegenth is assumed much larger than the size of the atoms considered, that

is the *dipole approximation* [6, 8] and tight quantitative bounds about it in the quantum optical context are given in [9], the field polarization direction is assumed to be a constant unit vector $\boldsymbol{u} \in \mathbb{R}^3$, so the field variations are solely described by its signal $E(t) \in \mathbb{R}$. In the simplified semi-classical picture, one omits the intrisic quantum nature of the field, and the electric energy is simply quantized as the electric field hamiltonian: $\hat{H}_1(t) = E(t)\boldsymbol{u} \cdot \hat{\boldsymbol{\mu}}$. Let's apply this to the electron: as explained, the electric field induces a transition between eigenstates of $\hat{H}_0$ if and only if their associated matrix element in $\boldsymbol{u} \cdot \hat{\boldsymbol{\mu}}$ is nonzero. Arbitary transitions are restricted by the *selection rules*, which roughly state that they are forbidden by default, unless the principal numbers differ and the magnetic and azimuthal numbers are somehow neighbors. In particular, the orbitals $|0\rangle, |1\rangle$ as previously defined are dipole-forbidden. We refer to [8] for a rigorous account of the selection rules.

$\hat{H}_0$ features a countable infinity of eigenstates that can be coupled through the electric field, so that even if the optical hardware is calibrated to a fixed transition between two particular energy levels, there is still a possibility of residual leakage to other levels. This issue is asymptotically solved by switching to a frame in which the Rabi cycle explained below is tweaked to perturbatively eliminate the unwanted excitations at any order: this is the Derivative Removal by Adiabatic Gate (DRAG) pulses method [11]. That being said, up to a desired accuracy $\hat{H}_0$ is reasonably truncated down to a reduced eigenbasis $\mathcal{B} = \{|1\rangle, |2\rangle\}$ (respectively ground and excited states), with interlevel energy difference ω_0 and transition dipole moment $a := \langle 1|\hat{\boldsymbol{\mu}} \cdot \boldsymbol{u}|2\rangle \neq 0$. Consider the electric signal monochromatic, so $E(t) = (E_0 e^{+\imath\omega t} + E_0^* e^{-\imath\omega t})/2$ for some frequency $\omega \in \mathbb{R}$ and amplitude $E_0 \in \mathbb{C}$. Since the spherical harmonics are odd functions, $\hat{\boldsymbol{r}}$ is off-diagonal and $\langle 1|\hat{\boldsymbol{\mu}}|1\rangle = \langle 2|\hat{\boldsymbol{\mu}}|2\rangle = \boldsymbol{0}$. Let us define:

$$\sigma^+ := |2\rangle\langle 1| \quad \sigma^- := |1\rangle\langle 2| \quad \sigma^z := |2\rangle\langle 2| - |1\rangle\langle 1|$$

Therefore the field-atom interaction hamiltonian in $\mathcal{B}$ is (up to a meaningless global phase shift):

$$\hat{H}_2 = \hat{H}_1(t) + \hat{H}_0 = \frac{1}{2}((E_0 e^{+\imath\omega t} + E_0^* e^{-\imath\omega t})(a\sigma^+ + a^*\sigma^-) + \omega_0\sigma^z)$$

Let U be the propagator of $\hat{H}_2$, and $V(t) := e^{-\imath\omega_0 t\sigma^z/2}U(t)$, then:

$$\begin{aligned}
\imath\dot{V}(t)V(t)^\dagger &= e^{\imath\omega_0 t\sigma^z/2}\hat{H}_1(t)e^{-\imath\omega_0 t\sigma^z/2} = \frac{1}{2}(E_0 e^{+\imath\omega t} + E_0^* e^{-\imath\omega t})(e^{+\imath\omega_0 t}a\sigma^+ + e^{-\imath\omega_0 t}a^*\sigma^-) \\
&= \hat{H}_+(t) + \hat{H}_-(t) \\
\hat{H}_\pm(t) &:= \frac{1}{2}(\Omega e^{+\imath(\omega\pm\omega_0)t}\sigma^+ + \Omega^* e^{-\imath(\omega\pm\omega_0)t}\sigma^-)
\end{aligned}$$

Where the *Rabi frequency* is $\Omega := aE_0$. Let $\lambda := \omega + \omega_0$ and W_λ be the propagator of $\hat{H}_+$. Applying the Dyson expansion:

$$\begin{aligned} I - W_\lambda(\tau) &= \imath \int_0^\tau \hat{H}_+(t)W(t)dt = \imath\lambda \int_0^\tau \hat{H}_+(t)dt + \int_0^\tau \hat{H}_+(t)\int_0^t \hat{H}_+(s)W(s)dsdt \\ &= \imath\lambda^{-1}\sin(\tau\lambda/2)(e^{\imath\tau\lambda/2}\Omega\sigma^+ - e^{-\imath\tau\lambda/2}\Omega^*\sigma^-) + O(\lambda^{-2}) \end{aligned}$$

So the effect of $\hat{H}_+$ vanishes as $\lambda \to \infty$ at a computable rate, while similarly the effect of H_- scales in $|\omega-\omega_0|^{-1}$. From now on we adopt the *Rotating wave approximation* (RWA), which consists in considering $|\omega-\omega_0| \ll \omega+\omega_0$ to remove $\hat{H}_+$ from the picture, and keep only $\hat{H}_-$. Let's redefine V as the propagator of $\hat{H}_-$, and enter the *rotating frame* by the basis change $|b\rangle \leftarrow e^{-\imath b\omega t}|b\rangle$ (which does not change anything to the observer). Letting $\Delta := \omega - \omega_0$ be the *detuning*, the rotating frame propagator is $U = e^{\imath\Delta t/2}V$, generated by the rotating frame hamiltonian:

$$\imath\dot{U}(t)U(t)^\dagger = e^{-\imath\Delta t\sigma^z/2}\hat{H}_-(t)e^{\imath\Delta t\sigma^z/2} + \frac{\Delta}{2}\sigma_z = \frac{1}{2}(\Omega\sigma^+ + \Omega^*\sigma^- + \Delta\sigma^z)$$

Defining a further effective Rabi frequency $\theta := (|\Omega|^2 + \Delta^2)^{1/2}$, the analytical expression of the Rabi cycle is:

$$U(t) = \cos\left(\frac{\theta t}{2}\right)\hat{I} + \sin\left(\frac{\theta t}{2}\right)\frac{1}{\theta}(\Omega\sigma^+ + \Omega^*\sigma^- + \Delta\sigma^z)$$

It is possible to consider a further time dependent Rabi frequency $\Omega(t)$ by modulating the source electric signal, i.e $E_0(t)$ instead of E_0. To stay compatible with the previously outlined framework, the field must evolve much slower than the reference frequency, i.e $\dot{E}_0 \ll \omega$. Beyond the two-level picture, this dynamic control setup is displayed in three-level schemes such as STImulated Raman Adiabatic Passage (STIRAP), used to transfer state populations between dipole-forbidden states [10, 8] or to engineer atomic single-qubit gates [19]. Deeper details about light-atom interactions can be found in [6].

A downside effect of light-atom processes is that

2.3 Rydberg-blockade multi-qubit gates

2.3.1 Interatomic dipolar interactions

Consider two charged particles with position vectors $\boldsymbol{r}_1, \boldsymbol{r}_2$, transition dipole moments $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$, distant by $R := \|\boldsymbol{r}_1 - \boldsymbol{r}_2\|$, connected by the unit vector $\boldsymbol{w} := (\boldsymbol{r}_1 - \boldsymbol{r}_2)/R$. The classical electric dipolar interaction energy between them is:

$$V = \frac{\boldsymbol{\mu}_1\cdot\boldsymbol{\mu}_2 - 3(\boldsymbol{\mu}_1\cdot\boldsymbol{w})(\boldsymbol{\mu}_2\cdot\boldsymbol{w})}{4\pi\varepsilon_0 R^3}$$

In the simplified biatomic setup, the angular dependence appearing through the inner products on the right is neglected [19], the distance R is treated classically and the electric dipole moments are quantized but only considering eigenstates of $\hat{H}_0$, so $|a\rangle, |c\rangle$ on the first atom and $|b\rangle, |d\rangle$ on the second one. Thus V is approximately quantized as:

$$\hat{V} = \frac{\langle a|\hat{\mu}|c\rangle \langle b|\hat{\mu}|d\rangle}{4\pi\varepsilon_0 R^3}(|ab\rangle\langle cd| + |cd\rangle\langle ab|) + \Delta_F |cd\rangle\langle cd|$$

Where $\Delta_F := E_c + E_d - E_a - E_b$ is the *Förster defect* with $E^a = \langle a|\hat{H}_0|a\rangle$ and so on. For the rest of this work, let's assume that through a careful choice of orbitals and laser frequencies (not detailed here), $\Delta_F \gg |\langle a|\hat{\mu}|c\rangle \langle b|\hat{\mu}|d\rangle|/4\pi\varepsilon_0 R^3$, resulting in the *Van der Waals* regime [19], where the dipolar interaction becomes:

$$\hat{V} = -\frac{C_6}{R^6}|cd\rangle\langle cd| \quad C_6 = \frac{|\langle a|\hat{\mu}|c\rangle \langle b|\hat{\mu}|d\rangle|^2}{4\pi\varepsilon_0 R^3 \Delta_F}$$

Notice that the electric dipole moment, and therefore the dipolar interaction, is invariant by transition to the rotating frame.

2.3.2 Blockade effect and gate setup

Consider two atoms, which can be in three eigenstates $\{|0\rangle, |1\rangle, |2\rangle\}$ of $\hat{H}_0$ (we refer to the explanations in the previous section to justify this reduction), so from the previous section we replace $|c\rangle = |d\rangle = |2\rangle$ while $|a\rangle, |b\rangle$ refer to ghost orbitals chosen to engineer the Van der Waals regime. Let $\hat{I} = |0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2|$ be the qutrit identity map and $B = C_6/R^6$. Considering our earlier explanations, the two-qutrits hamiltonian with zero individual detunings $\Delta_1 = \Delta_2 = 0$ is:

$$\hat{H}_{\text{FA}} + B|22\rangle\langle 22| = (\Omega_1|1\rangle\langle 2| + \Omega_1^*|2\rangle\langle 1|) \otimes \hat{I} + \hat{I} \otimes (\Omega_2|1\rangle\langle 2| + \Omega_2^*|2\rangle\langle 1|) + B|22\rangle\langle 22|$$

Consider a bi-atomic wavefunction $|\psi\rangle$ evolving under that hamiltonian, with a zero initial coordinate in $|22\rangle$: $\langle 22|\psi(0)\rangle = 0$, and let $\hat{G} = \hat{H}_{\text{FA}}/\Omega_{\max}$ define a dimensionless operator. Then, integrating by parts:

$$\begin{aligned}
e^{\imath B\tau}\langle 22|\psi(\tau)\rangle &= \int_0^\tau \frac{d}{dt}(e^{\imath Bt}\langle 22|\dot{\psi}(t)\rangle)dt \\
&= \int_0^\tau e^{\imath Bt}(\langle 22|\dot{\psi}(t)\rangle + \imath B\langle 22|\psi(t)\rangle)dt \\
&= -\imath\int_0^\tau e^{\imath Bt}\langle 22|\hat{H}_{\text{FA}}|\psi(t)\rangle\, dt \\
&= \frac{\imath}{-B\imath}\left((e^{\imath B\tau}\langle 22|\hat{H}_{\text{FA}}|\psi(\tau)\rangle - \langle 22|\hat{H}_{\text{FA}}|\psi(0)\rangle) - \int_0^\tau e^{-\imath Bt}\langle 22|\hat{H}_{\text{FA}}|\dot{\psi}(t)\rangle\, dt\right) \\
&= \frac{-1}{B}\left(\Omega_{\max}(e^{\imath B\tau}\langle 22|\hat{G}|\psi(\tau)\rangle - \langle 22|\hat{G}|\psi(0)\rangle) - \Omega_{\max}^2\int_0^\tau e^{\imath Bt}\langle 22|\hat{G}^2|\psi(t)\rangle\, dt\right) \\
&= \frac{\Omega_{\max}}{B}(\langle 22|\hat{H}_{\text{FA}}|\psi(0)\rangle - e^{\imath B\tau}\langle 22|\hat{H}_{\text{FA}}|\psi(\tau)\rangle) + O\left(\left(\frac{\Omega_{\max}}{B}\right)^2\right)
\end{aligned}$$

We will make the following assumptions:

- The two atoms are extremely close, the interatomic distance vanishes: $(R \to 0)$.
- $|2\rangle$ is a *Rydberg state* [23]: its principal quantum number is in $n > 60$, inducing strong dipole-dipole interactions since $C_6 \propto n^{11}$, a property enhanced by a careful choice of ghost orbitals $|a\rangle, |c\rangle$ from below. As Rydberg states enjoy higher lifetimes than lower fellow excited states [23], the gate operation is reasonably treated by the hamiltonian above while the risk of spontaneous emission from there is low enough (i.e happening on a much slower time scale than the gate operation) to be treated in the Lindblad-Markov method in open quantum systems.

Then under these conditions $B \gg \Omega_{\max}$ and the **blockade effect** occurs: $\langle 22|\psi\rangle \simeq 0$, which introduces a strong nonlinearity as pairs of atoms can be (un-)excited conditionally on one's state. Let us consider the succession of Rabi cycles outlined in [5]:

Period of $t\Omega_{\max}$	$\Omega_1/\Omega_{\max}$	$\Omega_2/\Omega_{\max}$	$\|00\rangle$	$\|10\rangle$	$\|01\rangle$	$\|11\rangle$
$[0,1]$	1	0	$\|00\rangle$	$\imath\|20\rangle$	$\|01\rangle$	$\imath\|21\rangle$
$[1,2]$	0	1	$\|00\rangle$	$\imath\|20\rangle$	$\imath\|02\rangle$	$\imath\|\mathbf{21}\rangle$
$[2,3]$	0	1	$\|00\rangle$	$\imath\|20\rangle$	$-\|01\rangle$	$\imath\|\mathbf{21}\rangle$
$[3,4]$	1	0	$\|00\rangle$	$-\|10\rangle$	$-\|01\rangle$	$-\|11\rangle$

The bold expressions signal the occurence of the Rydberg blockade effect: $|22\rangle$ is shifted out of resonance, so far detuned from all the other states that it can be considered roughly inacesssible to them: when the first atom is in a Rydberg state any further excitation of the second atom is suppressed. Up to further individual phase flips or to exchanging $|0\rangle$ and $|1\rangle$, this effectively implements a controlled-Z (CZ) gate or controlled-phase flip, i.e the diagonal operation $|b_0 b_1\rangle \mapsto (-1)^{b_0 b_1} |b_0 b_1\rangle$, $(b \in \{0,1\}^2)$. Along with single-qubit gates, these constitute a universal set of logical gates [9], that can thus implement any quantum circuit.

2.4 Optimization challenges

The theoretical insights aforementioned prove that Rydberg-interacting neutral atoms constitute genuine quantum hardware material, that can be handled through univeral single- and two-qubit gates. However, those convenient calculations derive from many approximations of the underlying physical theory.

In reference to the original Rydberg-blockade scheme, chapter 4 is dedicated to optimizing a version of the square pulses present therein that is robust against time-dependent noise. While this scheme is the simplest physical implementation of an entangling gate, such operations can also be treated with parallel laser drives evading

the need for individual atom adressing, and/or with a Rabi frequency as a smooth function of time $\Omega(t)$ as we'll refer in later chapters. We refer to [19] for a review of other possible gate setups.

Two-qubit gates based on the Rydberg blockade effect rely on exciting atoms to Rydberg states, which are subject to photon spontaneous emission and black body radiation, making them potentially decay to lower-energy states before the end of the gate operation, introducing errors in the process. This can be fought physically by placing the atoms in cryogenic environments making void around them [29], and by modulating the time-dependent Rabi frequency with pulses minimizing the time spent in Rydberg states as in [4]. Another leading source of errors is the finiteness of the Rydberg blockade potential. At first glance, one could argue that this can be managed by simply arbitrarily enhancing the interaction potential, either by increasing the principal quantum number n of Rydberg states, or decreasing the interatomic distance R, or decreasing the maximum Rabi frequency $\Omega_{\max}$. However:

- As n grows, the average detuning with fellow Rydberg states becomes smaller and smaller (since the energy splittings get tighter and tighter in the order of n^{-1}), so the risk of unwanted leakage to other Rydberg states increases and moves the dynamics out of the convenient two-level picture.

- R is fundamentally lower-bounded by the width of optical traps, characterized by the power of optical tweezers, so it cannot arbitrarily vanish. In [25] this distance is around 3µm

- As seen in the gate setup below, its duration is proportional to $1/\Omega_{\max}$, meaning that decreasing it results in a longer gate, which is then more exposed to Rydberg state decay.

Instead, the effect of Rydberg blockade finiteness is fought by either DRAG-type pulses analytically suppressing the coordinates of unwanted states [11] or numerical pulses drawn from optimal control tools that minimize the finite blockade contribution in perturbation theory as in [4]. Similar tools were also used to limit the effects of (constant-time) control fluctuations, thermal atomic motion and leakage to other atomic orbitals outside of the computational basis and targeted Rydberg states [24].

To summarize, the insights aforementioned lead to a clear identification of physical and numerical noise sources:

- Photonic spontaneous emission

- Dipolar approximation

- Two-level approximation (coupling to other states, dipolar contribution of other states)
- Rotating wave approximation
- Motion of the atoms (which can be treated classically or quantumly!) [19]
- Finiteness of Rydberg interaction
- Imperfection on the control apparatus
- Semi-classical approximation: omitting the quantum nature of the electromagnetic field)
- Non-relativistic approximation: omitting the effects of special relativity)

Therefore, it opens the possibility of a systematic modulation of the control-hardware coupling to fight those noises, which would be subject to optimization procedures, and motivate this thesis.

2.5 Conclusion

Neutral atoms constitute a reliable and promising quantum hardware technology, due to its flexibility, stability and recent technical progress, as well as the relative simplicity of its mathematical descriptions. The long established field of quantum optics allows to derive elegant one- and two-qubit gates on this platform, combined with the remarkable properties of Rydberg orbitals. However, these nice derivations sit on mathematical approximations that may not hold in realistic settings. To correct for those, it is important to robust develop gate pulses limiting the effect of those noise sources.

Bibliography

[1] Nielsen, M.A. and Chuang, I.L., 2010. Quantum computation and quantum information. Cambridge university press. https://profmcruz.wordpress.com/wp-content/uploads/2017/08/quantum-computation-and-quantum-information-nielsen-chuang.pdf (Cited on pages 5, 16 and 25.)

[2] Wootters, W.K., 2001. Entanglement of formation and concurrence. Quantum Inf. Comput., 1(1), pp.27-44. (Cited on pages 19 and 91.)

[3] Manzano, D., 2020. A short introduction to the Lindblad master equation. Aip advances, 10(2). (Cited on page 18.)

[4] Gonzalez-Ballestero, C., 2024. Tutorial: projector approach to master equations for open quantum systems. Quantum, 8, p.1454. https://doi.org/10.22331/q-2024-08-29-1454 (Cited on page 18.)

[5] Breuer, H.P. and Petruccione, F., 2002. The theory of open quantum systems. OUP Oxford. (Cited on page 18.)

[6] Preisser, G., Wellnitz, D., Botzung, T. and Schachenmayer, J., 2023. Comparing bipartite entropy growth in open-system matrix-product simulation methods. Physical Review A, 108(1), p.012616. (Cited on pages 17, 79 and 91.)

[7] Eisert, J., Cramer, M. and Plenio, M.B., 2010. Colloquium: Area laws for the entanglement entropy. Reviews of modern physics, 82(1), pp.277-306. https://doi.org/10.1103/RevModPhys.82.277, https://arxiv.org/abs/0808.3773. (Cited on pages 19 and 75.)

[8] Daniel A. Steck, Quantum and Atom Optics, available online at https://atomoptics.uoregon.edu/~dsteck/teaching/quantum-optics/quantum-optics-notes.pdf (revision 0.16.2, 15 November 2024). (Cited on pages 20, 21, 22 and 23.)

[9] Boßmann, L., Grummt, R. & Kolb, M. On the dipole approximation with error estimates. Lett Math Phys 108, 185–193 (2018). https://doi.org/10.1007/s11005-017-0999-y (Cited on page 22.)

[10] Vitanov, N.V., Rangelov, A.A., Shore, B.W. and Bergmann, K., 2017. Stimulated Raman adiabatic passage in physics, chemistry, and beyond. Reviews of Modern Physics, 89(1), p.015006. (Cited on page 23.)

[11] Theis, L.S., Motzoi, F., Wilhelm, F.K. and Saffman, M., 2016. High-fidelity Rydberg-blockade entangling gate using shaped, analytic pulses. Physical Review A, 94(3), p.032306. (Cited on pages 22 and 26.)

[12] Theis, L.S., Motzoi, F., Machnes, S. and Wilhelm, F.K., 2018. Counteracting systems of diabaticities using drag controls: The status after 10 years (a). Europhysics Letters, 123(6), p.60001. https://iopscience.iop.org/article/10.1209/0295-5075/123/60001 for a review of DRAG, (Not cited.)

[13] Motzoi, F. and Wilhelm, F.K., 2013. Improving frequency selection of driven pulses using derivative-based transition suppression. Physical Review A—Atomic, Molecular, and Optical Physics, 88(6), p.062318. https://journals.aps.org/pra/abstract/10.1103/PhysRevA.88.062318 (Not cited.)

[14] Motzoi, F., Gambetta, J.M., Rebentrost, P. and Wilhelm, F.K., 2009. Simple pulses for elimination of leakage in weakly nonlinear qubits. Physical review letters, 103(11), p.110501. https://doi.org/10.1103/PhysRevLett.103.110501 (Not cited.)

[15] Jaksch, D., Cirac, J.I., Zoller, P., Rolston, S.L., Côté, R. and Lukin, M.D., 2000. Fast quantum gates for neutral atoms. Physical Review Letters, 85(10), p.2208. (Cited on pages 6 and 25.)

[16] Gerry, C.C. and Knight, P.L., 2023. Introductory quantum optics. Cambridge university press. https://ostad.hormozgan.ac.ir/ostad/UploadedFiles/387098/387098-6463002517254326.pdf (Cited on pages 22 and 23.)

[17] Noh, K., Jiang, L. and Fefferman, B., 2020. Efficient classical simulation of noisy random quantum circuits in one dimension. Quantum, 4, p.318. (Cited on pages 75 and 79.)

[18] Phillips, W.D., 1998. Nobel Lecture: Laser cooling and trapping of neutral atoms. Reviews of Modern Physics, 70(3), p.721. (Cited on page 6.)

[19] Gordon, J.P. and Ashkin, A., 1980. Motion of atoms in a radiation trap. Physical Review A, 21(5), p.1606. https://doi.org/10.1103/PhysRevA.21.1606 (Cited on page 27.)

[20] Jandura, S. and Pupillo, G., 2022. Time-optimal two-and three-qubit gates for Rydberg atoms. Quantum, 6, p.712. (Cited on pages xxiv, 26, 53, 54, 61, 79, 81 and 114.)

[21] Branden, D.B., Juhasz, T., Mahlokozera, T., Vesa, C., Wilson, R.O., Zheng, M., Kortyna, A. and Tate, D.A., 2009. Radiative lifetime measurements of rubidium

Rydberg states. Journal of Physics B: atomic, molecular and optical physics, 43(1), p.015002. (Cited on page 7.)

[22] Hölzl, C., Götzelmann, A., Pultinevicius, E., Wirth, M., & Meinert, F. (2024). Long-lived circular Rydberg qubits of alkaline-earth atoms in optical tweezers. Physical Review X, 14(2), 021024. (Cited on page 7.)

[23] Beterov, I. I., Ryabtsev, I. I., Tretyakov, D. B., & Entin, V. M. (2009). Quasiclassical calculations of blackbody-radiation-induced depopulation rates and effective lifetimes of Rydberg n S, n P, and n D alkali-metal atoms with $n \leq 80$. Physical Review A—Atomic, Molecular, and Optical Physics, 79(5), 052504. (Cited on page 7.)

[24] Jandura, Sven, Jeff D. Thompson, and Guido Pupillo. "Optimizing Rydberg gates for logical-qubit performance." PRX Quantum 4.2 (2023): 020336. https://doi.org/10.1103/PRXQuantum.4.020336 (Cited on page 26.)

[25] Evered, S.J., Bluvstein, D., Kalinowski, M. et al. High-fidelity parallel entangling gates on a neutral-atom quantum computer. Nature 622, 268–272 (2023). https://doi.org/10.1038/s41586-023-06481-y (Cited on pages 7 and 26.)

[26] Wang, Y., Wang, Y., Chen, YA. et al. Efficient fault-tolerant implementations of non-Clifford gates with reconfigurable atom arrays. npj Quantum Inf 10, 136 (2024). https://doi.org/10.1038/s41534-024-00945-3 (Cited on page 7.)

[27] Pecorari, L., Jandura, S., Brennen, G.K. et al. High-rate quantum LDPC codes for long-range-connected neutral atom registers. Nat Commun 16, 1111 (2025). https://doi.org/10.1038/s41467-025-56255-5 (Cited on page 8.)

[28] Bluvstein, D., Evered, S.J., Geim, A.A. et al. Logical quantum processor based on reconfigurable atom arrays. Nature 626, 58–65 (2024). https://doi.org/10.1038/s41586-023-06927-3 (Cited on page 8.)

[29] Henriet, L., Beguin, L., Signoles, A., Lahaye, T., Browaeys, A., Reymond, G.O. and Jurczak, C., 2020. Quantum computing with neutral atoms. Quantum, 4, p.327. https://doi.org/10.22331/q-2020-09-21-327 (Cited on pages 6, 7 and 26.)

Chapter 3

Optimization concepts

Sommaire

This chapter introduces the optimization concepts necessary to understand the next chapter. Section 3.1 gives a non exhaustive introduction to artificial evolution and describes the mechanisms of two of its main subclasses : genetic algorithms and evolution strategies. Section 3.2 introduces multi-objective optimization, the art of computing the best trade-offs between contradictory trade-offs in engineering, through the concept of Pareto-optimality. Then, an overview of NSGA-III, a genetic algorithm for Pareto optimization featured in the next chapter, is given, and the chapter ends with a discussion on the topological compatibility between Pareto optimality and usual differentiable optimization.

3.1 Artificial evolution

3.1.1 Generalities

Artificial evolution (AE) is a class of numerical optimization algorithms inspired by the biological theories of evolution and natural selection. Let E be any set, $f : E \mapsto \mathbb{R}^+$ be bounded. In the AE lexikon, when seeking to solve $\min_{x \in E} f(x)$ (up to sign changes and addition of constants, without loss of generality this is representative of realistic problems), f is called the *fitness* function, a typical argument variable x is called an *individual* representing a potential solution and one algorithmic step is called a *generation*. AE works by updating at each generation a *population* $\mathcal{P} \subset E$ of candidate solutions (thus AE algorithms are called *population-based*), of fixed size $\lambda \in \mathbb{N}$, through the *reproductive cycle* made of:

- *Selection* : sampling of population members favoring low cost individuals in a more or less biased manner.
- *Recombination*: synthesis of new individuals ('offspring') out of selected ones, applied with probability $p_c \in [0, 1]$
- *Mutation* : slight modification of individuals on each coordinate at varying intensities, applied with probability $p_m \in [0, 1]$

In AE a collective behavior is sought: the solution **emerges** progressively from the population through the interactions between individuals (recombination) selected for high fitnesses (selection) whereas conventional approaches evolve individual solutions through the f fitness function properties.

Besides any consideration about f, the performance of an AE algorithm is essentially a matter of parameter tuning in order to get a good balance between exploration and exploitation thanks to genetic operators, and keeping the population and value variances in an acceptable range. This mirrors the generic dilemma in

numerical optimization between exploration (enriching the search space not to miss out on potential optimal outside of the current) and exploitation (digging into a potential optimal line currently searched). Indeed, selection and recombination are essentially exploitation mechanisms as they focus on high fitness indivuduals and their offspring which, as approaching the maximum, should have a high fitness too. On the other hand, mutation is essentially an exploration mechanism that enforces diversity in the population to avoid confinement in local maxima and rugged landscapes; and this can be said to a lesser extent about the selection mechanisms that give survival chances to low-fitness individuals so they may open the way to interesting future search directions. The fact that reproductive cycle operators are applied randomly with some prescribed probabilities makes AE algorithms fundamentally **stochastic**, preventing them from privileging search directions without reason and reinforces the explorative and exploitative characters of the reproductive cycle operators. Those features are enhanced as one favors large population sizes where AE algorithms benefit from parallelization (indeed theoretical guarantees of AE work in the limit of an infinite population). However, choosing the wrong population size and selection pressure can lead to poor convergence behavior: if population is too small, convergence may be premature, if it's too high convergence may be too slow. It is the opposite for selection pressure. These two parameters must also be mitigated with computing time/power.

In the most abstract formulation, AE does not leverage any special structure of f beyond its values, hence in highly structured cases, it can perform poorly compared to conventional algorithms. Indeed the components of AE can be thought as templates that can be filled with relevant informations about f. For example, when f is differentiable an AE algorithm can track the neighbourhood of a global maximum while a gradient ascent, ingrained into a modified fitness, locates/niches into the closest local minimum [8]. This approach corrects for the mutually exclusive caveats of both methods and is especially useful when f is highly non-convex (so its "sense of curvature" varies a lot), and thus has a lot of local maxima, wherein gradient methods tend to be trapped. The idea of combining a metaheuristic with differentiable optimization has already been studied in the quantum control context, though using simulated annealing [1].

AE is included in the larger class of nature-inspired *metaheuristics*, alongside other global optimization algorithms such as simulated annealing (metallurgy-inspired), Ant Colony, Artifical Bee Colony (entomology-inspired), Particle Swarm Optimisation (ethology-inspired), against which AE is frequently benchmarked since they are applied in unstructured settings too.
In this section, we give a non-exhaustive short review of two subclasses of AE: genetic algorithms and evolution strategies. Another popular AE algorithm is differential evo-

lution. Multi-objective extensions of various artificial evolution algorithms are outlined in the subsequent section.

3.1.2 Genetic algorithms

Even though Evolution Strategies (ES) (cf. section 3.1.3) predate genetic algorithms, it is Genetic Algorithms (GA) that are best known within the class of artificial evolution algorithms. First proposed by John Holland in 1975 [2], GAs push the biological analogy down to the genetic level. In the GA lexikon, the set of variables to optimize is called a *genome* where each gene is a binary digit value (bit).

$\mathcal{P}$ is the population of fixed size $\lambda \in \mathbb{N}$, $p_c \in [0,1[$ is the recombination probability, $p_m \in [0,1[$ is the mutation probability. GAs are naturally suited for discrete and combinatorial problems, as their original formulation represents solutions as binary strings $\boldsymbol{x} \in \{0,1\}^n$ ($n \in \mathbb{N}$), from where problems in finite sets can be converted without loss. This creates many problems if the optimization problem representation needs integers (or worse, real values) as the modification of a single bit by the mutation operator can result in wild changes in the resulting integer or real value. Continuous problems can be treated by either truncating their variables' binary decomposition (up to the desired accuracy), or by constructing adapted reproductive cycle operators as we mention below.

Genetic selection consists in randomly selecting genomes in $\mathcal{P}$ to recombine, with probability biased towards high fitness and without excluding low-fitness genomes to avoid confinement in local optima. Though more restrictives methods exist, let us describe here four significant selection schemes [9]:

- *Roulette wheel* : Select x in $\mathcal{P}$ with probability $1 - f(x)/(\sum_{z\in\mathcal{P}} f(z))$ [1].
- *Stochastic acceptance*: Select x in $\mathcal{P}$ uniformly, select it for reproduction with another individual selected with probability $1 - (f(x)/(\max_{z\in\mathcal{P}} f(z)))$.
- *Ranking-based*: Sort $\mathcal{P}$ by fitness values $f(z_1) \leq \cdots \leq f(z_i) \leq \cdots \leq f(z_\lambda)$. Select z_i with probability $p(i)$. Here $[\![1,\lambda]\!] \ni i \mapsto p(i)$ is a precomputed probability distribution as a normalized decreasing function of population indices, e.g:
  - linear pattern $p(i) = (s - 2(s-1)(i-1)/(\lambda-1))/\lambda$ with selection pressure parameter $s \in [1,2]$
  - exponential pattern $p(i) = \frac{1-w}{w^{-\lambda}-1} w^{-i}$ with ratio $w \in [0,1[$.

[1] To select from a discrete probability distribution $(p_i)_{i\in\mathbb{N}}$, let $U \sim \mathcal{U}([0,1])$, $I := \min(i \in \mathbb{N} | \sum_{j=1}^{i} p_i > U)$ so that $\mathbb{P}(I = i) = p_i$.

- *Tournament* comes in 2 varieties: a standard k-ary tournament and a stochastic tournament:
  - *Standard k-ary tournament* (with parameter $k \in [\![1, \lambda]\!]$): uniformly sample k elements of $\mathcal{P}$ and select the best of them.
  - *Stochastic tournament*: (with parameters $k = 2, p_t \in [0, 1]$): and select the best with probability p.

  Tournament is the preferred selection method because it is cheap, it can select individuals with an easy to define selection pressure that is independent of the fitness landcape and contrarily to roulette wheel, it can be used for minimizing problems (such as minimizing noise).

In the GA lexikon, recombination is termed 'crossover'. In the discrete binary case, we restrict our review here to the single point crossover, the exchange of their respective heads/tails at a given site (generally uniformly sampled), $j(\sim \mathcal{U}([\![1, n]\!]))$:

$$C : \begin{pmatrix} (x_1, \cdots, x_j, x_{j+1}, \cdots, x_n) \\ (y_1, \cdots, y_j, y_{j+1}, \cdots, y_n) \end{pmatrix} \mapsto \begin{pmatrix} (x_1, \cdots, x_j, y_{j+1}, \cdots, y_n) \\ (y_1, \cdots, y_j, x_{j+1}, \cdots, x_n) \end{pmatrix}$$

Though more complex crossover schemes exist [3]. In the continuous real case: simulated binary crossover (SBX) [12], represented/summarized by the matrix:

$$C : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \frac{1}{2} \begin{pmatrix} 1+\beta & 1-\beta \\ 1-\beta & 1+\beta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Where β is a real parameter that in [12] is sampled after the probability distribution with real parameter $\eta \neq 1$ and density:

$$p(\beta) = \frac{1+\eta}{2} \begin{cases} \beta^{\eta} & \text{if } \beta \in]0, 1] \\ \beta^{-(\eta+2)} & \text{if } \beta \geq [1, +\infty[\end{cases}$$

Notice that its cumulative distribution function (CDF) is

$$F(x) = \int_0^x p(\beta) d\beta = \begin{cases} x^{1+\eta}/2 & \text{if } x \in [0, 1] \\ 1 - x^{-(1+\eta)}/2 & \text{if } x \in [1, +\infty] \end{cases}$$

The inverse CDF is:

$$F^{-1}(y) = \begin{cases} (2y)^{1/(1+\eta)} & \text{if } y \in [0, 1/2] \\ (1-2y)^{-1/(1+\eta)} & \text{if } x \in [1/2, 1] \end{cases}$$

So by the inverse sampling theorem, $\beta = F^{-1}(U)$ with $U \sim \mathcal{U}([0, 1])$.

In the discrete binary setting, the mutation consists of randomly flipping each bit independently with probability $p_m \in [0,1]$:

$$\forall j \in [\![1,n]\!],\ x_j \leftarrow x_j \oplus \mathbb{1}_{u_j \leq p_m} \quad (u_j \sim \mathcal{U}([0,1])\ \text{i.i.d.})$$

Where $\oplus$ be the modulo-2 sum. More complex patterns exist in the continuous variables setting, we only outline here the popular polynomial mutation applied in NSGA-III and explained in [10, 11], in box constraints $x \in [x^l, x^u]$, a mutation parameter $\eta \neq -1$. Apply on each coordinate (with independent $r \sim \mathcal{U}([0,1])$)

$$x \leftarrow x + (x^u - x^l) \times \begin{cases} \left(2r + (1-2r)\left(\frac{x^u - x}{x^u - x^l}\right)^{1+\eta}\right)^{1/(1+\eta)} - 1 & \text{if } r < 0.5 \\ 1 - \left(2(1-r) + 2(r-0.5)\left(\frac{x - x^l}{x^u - x^l}\right)^{1+\eta}\right)^{1/(1+\eta)} & \text{if } r \geq 0.5 \end{cases}$$

A GA is said to have converged when either all the individuals of the population become similar. This results is the stopping of the progression of the improvement of the average/median/maximal fitness across several generations, after which the evolution is stopped after there is no change in fitness value for too many generations. To summarize, Algorithm 1 below lists the pseudo-code of a typical GA.

Input: Population size $\lambda \in 2\mathbb{N}$, Mutation probability $p_m \in [0,1[$, Crossover probability $p_c \in [0,1[$
Output: An approximation y to $\max_{x \in E} f(x)$
Generate initial population $\mathcal{P}$ by uniformly sampling λ elements in E;
while *Termination criteria not met* **do**
  Compute population fitness $(f(x))_{x \in \mathcal{P}}$;
  Update selection scheme from fitness values;
  $\mathcal{P}' \leftarrow \varnothing$;
  for $k = 1$ *to* $\lambda/2$ **do**
    Select independently two distinct genomes (x, x') from $\mathcal{P}$;
    Perform crossover $(x, x') \leftarrow C(x, x')$ with probability p_c (at a uniformly random site in the discrete case);
    Mutate each coordinate of (x, x') independently with probability p_m;
    $\mathcal{P}' \leftarrow \mathcal{P}' \cup \{x, x'\}$
  end
  ; $\mathcal{P} \leftarrow \mathcal{P}'$;
end
$y \leftarrow \max_{x \in \mathcal{P}} f(x)$;

Algorithm 1: Generation loop of a typical GA

In the combinatorial setting $E = \{0,1\}^n$, GAs have been shown to converge asymptotically to the fitness global optimum. In [5] is proved an upper bound on the minimal population size such that the Markov chain generated by the GA passes

through the fitness optimum almost surely. Interestingly, this bound is related to the minimal variation between the fitness values, whereas in quantum computing the performance of adiabatic state preparation is foremost bounded by the inverse squared spectral gap of the target hamiltonian (see chapter 2). In [6] are outlined tight upper bounds about the minimal number of generations necessary to visit the global optimum with desired probability.

In [3], a series of benchmarks are cited that show empirically evidence for nice performances of GA in combinatorial and networking problems. In this vein, GA have been proposed in information theory to formulate new error-correcting codes [13, 14] with comparable performance to conventional ones. Beyond numerical quantities, GAs also optimize the design of abstract structures, such as symbolic functions through genetic programming [4], or neural network architectures through NEAT (Neuronal Evolution of Augmented Topologies) [7].

3.1.3 Covariance Matrix Adaptation Evolution Strategies (CMA-ES)

Evolution Strategies (ES) are a subclass of AE algorithms where historically, the population only evolves thanks to mutation and selection. ES are typically suited for black-box continuous-variables problems, and originated in the 1960s through the pioneering work of Ingo Rechenberg and Hans-Paul Schwefel. We refer to [15, 16] for extensive reviews of the topic.

In this subsection, we only briefly outline a major variant thereof, the *Covariance Matrix Adaptation Evolution Strategy* (CMA-ES), that is featured in chapter 4, and refer to [17] for a thorough review of the topic. The core idea of CMA-ES here is to evolve a normally-distributed population $\mathcal{P} \subset \mathbb{R}^N$ of $\lambda \in \mathbb{N}$ candidate solutions, characterized by a *covariance matrix* $\boldsymbol{C}$ and an average vector $\boldsymbol{m}$, updated accross generations through adaptative step variables.

The purpoted effect is that as generations pass, the usual AE operators (selection, recombination, mutation) push the population towards lower fitness values with a reduced tendency of getting stuck in sub-optimal valleys (local optima). Here below is the CMA-ES algorithm in natural language:

Data: Population size $\lambda \in \mathbb{N}$, Recombination weights $(w_i)_{i=[\![1,\lambda]\!]}, (c_m, c_\sigma, d_\sigma, c_c, c_1, c_\mu) \in [0,1]^5$,
Input: Initial average vector $\boldsymbol{m} \in E$
Output: An approximation to $\min_{\boldsymbol{x}\in E} f(\boldsymbol{x})$
$\boldsymbol{p}_\sigma \leftarrow \boldsymbol{0}$;
$\boldsymbol{p}_c \leftarrow \boldsymbol{0}$;
$\boldsymbol{C} \leftarrow \boldsymbol{I}$;
$g \leftarrow 0$;
while *Termination criteria not met* **do**
  $g \leftarrow g + 1$;
  for $i = 1$ *to* λ **do**
    $\boldsymbol{x}_i \leftarrow \mathcal{N}(\boldsymbol{m}, \sigma\boldsymbol{C})$;
  end
  f-sort the population so that $f(\boldsymbol{x}_1) \le \cdots \le f(\boldsymbol{x}_i) \le \cdots \le f(\boldsymbol{x}_\lambda)$;
  for $i = 1$ *to* λ **do**
    $\boldsymbol{y}_i \leftarrow (\boldsymbol{x}_i - \boldsymbol{m})/\sigma$;
  end
  $\langle \boldsymbol{y} \rangle \leftarrow \sum_{i=1}^{\mu} w_i \boldsymbol{y}_i$;
  $\boldsymbol{m} \leftarrow \boldsymbol{m} + c_m \sigma \langle \boldsymbol{y} \rangle$;
  $\boldsymbol{p}_\sigma \leftarrow (1 - c_\sigma)\boldsymbol{p}_\sigma + \sqrt{c_\sigma(2 - c_\sigma)\mu_{\text{eff}}}\boldsymbol{C}^{-1/2} \langle \boldsymbol{y} \rangle_w$;
  $\sigma \leftarrow \sigma \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\|\boldsymbol{p}_\sigma\|}{a_N} - 1\right)\right)$;
  $h_\sigma = \begin{cases} 1 & \text{if } \|\boldsymbol{p}_\sigma\| < (1.4 + \frac{2}{N+1}) a_N (1 - (1 - c_\sigma)^{2(g+1)})^{1/2} \\ 0 & \text{else} \end{cases}$;
  $\boldsymbol{p}_c \leftarrow (1 - c_c)\boldsymbol{p}_c + h_\sigma \sqrt{c_\sigma(2 - c_\sigma)\mu_{\text{eff}}} \langle \boldsymbol{y} \rangle_w$;
  for $i = 1$ *to* λ **do**
    $w_i^\circ \leftarrow w_i \left(\mathbb{1}_{w_i \ge 0} + \mathbb{1}_{w_i < 0} \frac{n}{\|\boldsymbol{C}^{-1/2}\boldsymbol{y}_i\|^2}\right)$;
  end
  $\boldsymbol{C} \leftarrow (1 + ((1 - h_\sigma)c_c(2 - c_c) - 1)c_1 - c_\mu \sum_{i=1}^{\lambda} w_i)\boldsymbol{C} + c_1 \boldsymbol{p}_c \boldsymbol{p}_c^T + c_\mu \sum_{i=1}^{\lambda} w_i^\circ \boldsymbol{y}_i \boldsymbol{y}_i^T$
end

Algorithm 2: Summary of CMA-ES

Let's explain a few points of this algorithm. a_N refers to the expectation value of the norm of a vector of indepedent standard-normally distributed coordinates:

$$a_N = \mathbb{E}(\|\mathcal{N}(\mathbf{0}, \boldsymbol{I})\|) = \sqrt{2}\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}$$

Where $\Gamma : x \mapsto \int_0^{+\infty} t^x e^{-t} dt$ is the Euler Gamma function. Next, the input 'Data' parameters of the algorithms can be freely defined by the user, but the authors have also recommended default values for them:

- $\lambda \in \mathbb{N}$ is the population size, the original authors of CMA-ES recommend to always have $\lambda \geq 5$, and in particular they propose to work with $\lambda = 4 + \lfloor 3 \ln N \rfloor$.
- $\mu \in [\![1, \lambda]\!]$ is the number of 'best' individuals selected to pass on their information to the next generation, recommended as $\mu = \lambda/2$ (selection of the best half of the population). It is also the number of positive recombination weights.
- Pre-weights recommended in logarithmic scale: $w'_i = \ln((\lambda+1)/2) - \ln(i)$
- Effective positive and negative variances

$$\mu_{\text{eff}} = \frac{(\sum_{i=1}^{\mu} w'_i)^2}{\sum_{i=1}^{\mu} w'^2_i} \quad \bar{\mu}_{\text{eff}} = \frac{(\sum_{i=\mu+1}^{\lambda} w'_i)^2}{\sum_{i=\mu+1}^{\lambda} w'^2_i}$$

- Various coefficients:

$$\alpha_{\mu}^{-} = 1 + \frac{c_1}{c_\mu} \quad \alpha_{\mu_{\text{eff}}}^{-} = 1 + \frac{2\bar{\mu}_{\text{eff}}}{\bar{\mu}_{\text{eff}} + 2} \quad \alpha_{\text{pos def}}^{-} = \frac{1 - c_1 - c_\mu}{n c_\mu}$$

- $(\mu_i)_{i \in [\![1,\lambda]\!]}$ is a collection of real recombination weights such that $\sum_{i=1}^{\mu} w_i = 1$, $w_1 \geq w_2 \geq \cdots \geq w_\mu > 0$. The authors recommend to put:

$$w_i = w'_i \times \begin{cases} \frac{1}{\sum_{j=1}^{\mu} w'_j} & \text{if } w'_i > 0 \\ \frac{-1}{\sum_{j=\mu+1}^{\lambda} w'_j} \min(\alpha_{\mu}^{-}, \alpha_{\mu_{\text{eff}}}^{-}, \alpha_{\text{pos def}}^{-}) & \text{else} \end{cases}$$

- Step size controls:

$$c_m = 1 \quad c_\sigma = \frac{\mu_{\text{eff}} + 2}{n + \mu_{\text{eff}} + 5} \quad d_\sigma = 1 + 2\max\left(0, \sqrt{\frac{\mu_{\text{eff}} - 1}{n+1}}\right) + c_\sigma$$

- CMA:

$$c_c = \frac{4 + \mu_{\text{eff}}/n}{n + 4 + 2\mu_{\text{eff}}/n} \quad = 2 \quad c_1 = \frac{\alpha_{\text{cov}}}{(n+1.3)^2 + \mu_{\text{eff}}}$$
$$c_\mu = \min\left(1 - c_1, \frac{\alpha_{\text{cov}}(1/4 + \mu_{\text{eff}} + 1/\mu_{\text{eff}} - 2)}{(n+2)^2 + \alpha_{\text{cov}}\mu_{\text{eff}}/2}\right)$$

In [16] are cited an important number of examples hinting at the convergence of general ES to the global optimum of at least some nicely behaved spherical functions, though the algorithm success remain mostly empirical in nature. Importantly, CMA-ES is provably related to the natural gradient method [16] of updating distributions in statistical learning in an optimal way. One could expect this connection to explain the empirical performances of CMA-ES. The CMA-ES admits a multi-objective extension [20].

3.2 Multi-objective optimization

3.2.1 Generalities

In science and engineering, it is common to meet problems that require the simultaneously optimization of several contradictory criteria (i.e in the form $\min_{x\in E}(f_1(x), f_2(x))$). Those are called *multi-objectives optimisation problems* (MOP) and show up e.g in quantum computing through the design of logic gates as e.g the following dilemmas:

- A higher maximal Rabi frequency leads to faster gates (reducing the risk of Rydberg state decay) but for a fixed Rydberg interaction strength, it reduces the blockade effect (and therefore the two-qubit gate fidelity).
- Leakage transitions outside of the computational subspace can be adiabatically eliminated by specific pulse shapes (see the previous chapter 2), but this induces longer gates that are more exposed to Rydberg state decay. This also applies to other sources of noise, such as control fluctuations or atomic thermal motion, since by definition, the shortest gate correcting them will be longer than the shortest gate not correcting them.

Abstract MOP may seem unsolvable by definition, but are easier to figure out once reformulated as **computing the best possible trade-offs** between the criteria. The quality of a trade-off can be naturally formulated as the propensity of not being improvable on one criterion without worsening the other ones (i.e avoid moves $x_1 \to x_2$ such that $f_1(x_1) > f_1(x_2)$ and $f_2(x_1) < f_2(x_2)$). This leads to a multivariate comparison relevant in mathematical optimisation, the Pareto ordering $\succ$ [23]:

$$\forall y, y' \in \mathbb{R}^m,\ y \succ y' \iff \begin{cases} \forall i \in [\![1, m]\!],\ y_i \geq y'_i \\ \exists i \in [\![1, m]\!],\ y_i > y'_i \end{cases}$$

Under the natural quality definition aforementioned, a trade-off is called "weakly dominant" if no other element dominates it in this ordering, i.e if it is a weak maximal element thereof. By definition, arbitary vectors cannot be directly Pareto-compared (so

the order is not total and strong optima may not exist), however their trade-off quality can be weakly compared by counting the elements Pareto-dominating them and comparing the numbers (which can always be done when working in a finite set $Y \subset \mathbb{R}^m$ as in the algorithms we will see later). When that 'dominance count' is zero, a trade-off is Pareto-optimal (the individual is called a strong dominant) and belongs to the hereby defined *Pareto front*:

$$P(Y) = \{y \in Y \mid \{y' \in Y \mid y' \succ y\} = \varnothing\}$$

The goal of multi-objectives optimizers is to find out individuals that are as close as possible to the theoretical Pareto front. As an interesting mise-en-abyme, we would like to point out that the choice of hyperparameters of a genetic algorithm is itself many-objectives MOPs.

3.2.2 NSGA-III: A Reference-Point-Based metaheuristic approach to Pareto Multi-Objective Optimization

Multi-objectives extensions have been proposed to simulated annealing [24, 23], differential evolution [25] and CMA-ES [20]. In this subsection, we briefly describe the third Non-dominated Sorting Genetic Algorithm (NSGA-III) [18, 19], as it is featured in chapter 4, and is now an established state of the art method to solve MOPs.

Let E be any set, $\boldsymbol{f} = (f_j)_{j\in[\![1,M]\!]} : E \to \mathbb{R}^M$ be the objectives. Suppose to seek a solution to the MOP $\min_{x\in E}(f_j(x))_{j\in[\![1,M]\!]}$. As its name suggests, the first technical feature of NSGA-III is to sort the population by Pareto optimality degree, as shown in the `Non-dominated-sort` function (Algorithm 4) outlined below. By systematically selecting and recombining Pareto optimal genomes in the population, and diversifying the said population through mutations, the algorithm seeks to approximate the actual set of Pareto-optimal genomes in the whole space. The second technical feature of NSGA-III, is the introduction of **reference points** mechanisms to avoid concentration phenomena (a risk increasing with M) and enforce diversity. The number H of reference points is fixed by the user and the authors recommend to choose it in the form $H = \binom{M+p-1}{p}$ where p is a user-defined parameter controlling their density. The reference points are constructed and managed by the `Normalize` (Algorithm 5), `Associate` (Algorithm 6) and `Niching` (Algorithm 7) functions outlined in the next pages. NSGA-III can be used in either discrete or continuous variables setup, in which case the authors recommend to use the simulated binary crossover and polynomial mutation operators, that were

outlined in the previous section.

Input: Either H reference points Z^s or aspiration points Z^a, parent population P_t
Output: Offspring population P_{t+1}
$S_t \leftarrow \varnothing$;
$i \leftarrow 1$;
$Q_t \leftarrow$(Recombination+Mutation)(P_t);
$R_t \leftarrow P_t \cup Q_t$;
$(F_1, \cdots, F_p) \leftarrow$ `Non-dominated-sort`(R_t);
while $|S_t| < \lambda$ **do**
  $S_t \leftarrow S_t \cup F_i$;
  $i \leftarrow i + 1$;
end
$l \leftarrow i - 1$ (Index of last front included);
if $|S_t| = \lambda$ **then**
  $P_{t+1} \leftarrow S_t$;
  break;
end
else
  $P_{t+1} \leftarrow \cup_{j=1}^{l-1} F_j$;
  $K \leftarrow N - |P_{t+1}|$;
  $(\mathbf{f}^n, Z^r) \leftarrow$ `Normalize`(S_t, Z^s, Z^a);
  for $s \in S_t$ **do**
    $(\pi(s), d(s)) \leftarrow$ `Associate`(S_t, Z^r);
  end
  for $j \in Z^r$ **do**
    $\rho_j \leftarrow \sum_{s \in S_t \setminus F_l} \mathbb{1}_{\pi(s)=j}$;
  end
  $P_{t+1} \leftarrow$ `Niching`$(K, \rho, \pi, d, Z^r, F_l, P_{t+1})$;
end
return P_{t+1};

Algorithm 3: NSGA-III generation t

Input: $\mathcal{P}$
Output: Pareto fronts$(F_1, \cdots, F_p)$
$F_0 \leftarrow \varnothing$;
while $\mathcal{P} \setminus F_i \neq \varnothing$ **do**
  $F_{i+1} \leftarrow \{x \in \mathcal{P} \mid \{z \in \mathcal{P} \mid f(z) \succ f(x)\} = \varnothing\}$;
  $i \leftarrow i + 1$;
end

Algorithm 4: `Non-dominated-sort`

Input: S_t, either H reference points Z^s or aspiration points Z^a
Output: $\mathbf{f}^n, Z^r$,
for $j \in [\![1, M]\!]$ **do**
  $z_j^{\min} \leftarrow \min_{s \in S_t} f_j(s)$;
  $\mathbf{z}^{j,\max} = \arg\min_{s \in S_t} \max_{i \in [\![1,M]\!]} f_j(s)$;
end
Construct the unique hyperplane $\mathcal{H}$ containing the points $(\mathbf{z}^{j,\max})_{j \in [\![1,M]\!]}$;
for $j \in [\![1, M]\!]$ **do**
  $a_j \leftarrow$ Intercept of i-th objective axis and $\mathcal{H}$;
end
for $j \in [\![1, M]\!]$ **do**
  $(\mathbf{f}^n)_j = \frac{f_j(\mathbf{z}^{j,\max}) - z_j^{\min}}{a_j - z_j^{\min}}$;
end
if Z^a *is given* **then**
  $Z^r \leftarrow$ Set of points in Z^a projected to the normalized $\mathcal{H}$;
end
else
  $Z^r \leftarrow Z^s$;
end

Algorithm 5: `Normalize`

Input: Z^r, S_t
Output: $\{(\pi(s), d(s)) \mid s \in S_t\}$
for $s \in S_t$ **do**
  for $w \in Z^r$ **do**
    $d^\perp(s, w) \leftarrow s - w^T s / \|w\|$
  end
  $\pi(s) \leftarrow \arg\min_{w \in Z^r} d^\perp(s, w)$;
  $d(s) \leftarrow d^\perp(s, \pi(s))$;
end

Algorithm 6: `Associate`

Input: $K, \rho, \{(\pi(s), d(s) \mid s \in S_t)\}, Z^r, F_l$
Output: P_{t+1}
$k \leftarrow 1$;
while $k \leq K$ **do**
  $J_{\min} = \arg\min_{j \in Z^r} \rho_j$;
  $c \leftarrow \text{random}(J_{\min})$;
  $I_c \leftarrow \{s \in F_l \mid \pi(s) = c\}$;
  if $I_c \neq \varnothing$ **then**
    if $\rho_c = 0$ **then**
      $G \leftarrow \arg\min_{s \in I_c} d(s)$;
    end
    else
      $G \leftarrow \{\text{random}(I_c)\}$;
    end
    $P_{t+1} \leftarrow P_{t+1} \cup G$;
    $\rho_c \leftarrow \rho_c + 1$;
    $F_l \leftarrow F_l \setminus \{G\}$;
    $k \leftarrow k + 1$;
  end
  else
    $Z^r \leftarrow Z^r \setminus \{c\}$;
  end
end

Algorithm 7: `Niching`

3.2.3 Pareto optimization in differentiable settings

A significant feature of the Pareto order is its topological compatibility with generic necessary conditions in differentiable optimization [21], and further compatibility with

sufficient conditions in convex cases too:

$$x^* \in \arg\min_{x \in E} (g_i(x))_{i \in [\![1,m]\!]} \implies \exists \{\lambda_i\}_{i=1}^m \subset \mathbb{R}^+, \sum_{i=1}^m \lambda_i = 1, \sum_{i=1}^m \lambda_i \nabla g_i(x) = 0$$

It means that a differentiable semi-necessary Pareto optimality conditions is to obey the conventional optimiality necessary condition for a convex combination of the criteria. So to put it short, Pareto optimization remarkably preserves the structure of differentiable optimization problems, greatly simplifying the handling of contradictory criteria. The gradients and differentiation here can be taken in the Banach sense, as this capacity is passed on to the calculus of variations (the field of optimizing integral quantities depend on functions and their differentials) and optimal control theory [21, 22].

3.3 Conclusion

Pareto Optimization is a powerful class of concepts to tackle engineering problems in multiple criteria, which frequently occur in designing quantum logic gates. Its weak formulation allows to adapt confirmed artificial evolution algorithms to deal with those problems in poorly structured context. On the other hand, its relation to differentiable optimization, notably in infinite-dimensional spaces, further confirms the solidity of tools from calculus of variations and optimal control theory under more structured hypotheses.

Bibliography

[1] Ram, M.H., Krithika, V.R., Batra, P. and Mahesh, T.S., 2022. Robust quantum control using hybrid pulse engineering. Physical Review A, 105(4), p.042437. (Cited on page 33.)

[2] Holland, J.H., 1975. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. MIT press. Reprinted in 1992. (Cited on page 34.)

[3] Katoch, S., Chauhan, S.S. and Kumar, V., 2021. A review on genetic algorithm: past, present, and future. Multimedia tools and applications, 80, pp.8091-8126. (Cited on pages 35 and 37.)

[4] Bodenhofer, U., 2003. Genetic algorithms: theory and applications. Lecture notes, Fuzzy Logic Laboratorium Linz-Hagenberg, Winter, 2004. https://www.cs.upc.edu/~belanche/Docencia/cne/material-complementario/Genetic_Bodenhofer.pdf (Cited on page 37.)

[5] Raphaël Cerf. Asymptotic convergence of genetic algorithms. Advances in Applied Probability, 1998, 30 (2), pp.521-550. ⟨10.1239/aap/1035228082⟩. ⟨hal-03702788⟩ (Cited on page 36.)

[6] Greenhalgh, D. and Marshall, S., 2000. Convergence criteria for genetic algorithms. SIAM Journal on Computing, 30(1), pp.269-282. https://doi.org/10.1137/S009753979732565X (Cited on page 37.)

[7] Stanley, K.O. and Miikkulainen, R., 2002. Evolving neural networks through augmenting topologies. Evolutionary computation, 10(2), pp.99-127. (Cited on page 37.)

[8] D'Angelo, G., & Palmieri, F. (2021). GGA: A modified genetic algorithm with gradient-based local search for solving constrained optimization problems. Information Sciences, 547, 136-162. (Cited on page 33.)

[9] Blickle, T. and Thiele, L., 1995. A comparison of selection schemes used in genetic algorithms. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=48a2821e3fe010ed47e4b1521912336601a1f710 (Cited on page 34.)

[10] Carles-Bou, J.L., Galán, S.F. Self-adaptive polynomial mutation in NSGA-II. Soft Comput 27, 17711–17727 (2023). https://doi.org/10.1007/s00500-023-09049-0 (Cited on page 36.)

[11] Hamdan, M., 2011. A dynamic polynomial mutation for evolutionary multi-objective optimization algorithms. International Journal on Artificial Intelligence Tools, 20(01), pp.209-219. (Cited on page 36.)

[12] Deb, K., Sindhya, K. and Okabe, T., 2007, July. Self-adaptive simulated binary crossover for real-parameter optimization. In Proceedings of the 9th annual conference on genetic and evolutionary computation (pp. 1187-1194). https://doi.org/10.1145/1276958.1277190 (Cited on page 35.)

[13] Comellas, F., & Roca, R. (2013, June). Using genetic algorithms to design constant weight codes. In Proceedings of the International Workshop on Applications of Neural Networks to Telecommunications (pp. 119-124). Psychology Press. https://web.mat.upc.edu/francesc.comellas/old-files/codes/codes_f.pdf (Cited on page 37.)

[14] Alba, E., & Chicano, J. F. (2004, March). Solving the error correcting code problem with parallel hybrid heuristics. In Proceedings of the 2004 ACM symposium on Applied computing (pp. 985-989). https://doi.org/10.1145/967900.968101 (Cited on page 37.)

[15] Beyer, HG., Schwefel, HP. Evolution strategies – A comprehensive introduction. Natural Computing 1, 3–52 (2002). https://doi.org/10.1023/A:1015059928466. Available at https://saksagan.ceng.metu.edu.tr/courses/ceng713/documents/Evolution%20strategies%20a%20comprehensive%20introduction.pdf (Cited on page 37.)

[16] Nikolaus Hansen, Dirk V. Arnold, Anne Auger. Evolution Strategies. Janusz Kacprzyk; Witold Pedrycz. Handbook of Computational Intelligence, 871-898, Springer, 2015, 978-3-622-43504-5. Available at http://www.cmap.polytechnique.fr/~nikolaus.hansen/es-overview-2015.pdf. (Cited on pages 37 and 40.)

[17] Hansen, Nikolaus. "The CMA evolution strategy: A tutorial." arXiv preprint arXiv:1604.00772 (2016). https://arxiv.org/abs/1604.00772 (Cited on page 37.)

[18] Deb, K. and Jain, H., 2013. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints. IEEE transactions on evolutionary computation, 18(4), pp.577-601. https://doi.org/10.1109/TEVC.2013.2281535. Accessible on the author's page as https://www.egr.msu.edu/~kdeb/papers/k2012009.pdf. (Cited on page 41.)

[19] Jain, H. and Deb, K., 2013. An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part II: Handling constraints and extending to an adaptive approach. IEEE Transactions on

evolutionary computation, 18(4), pp.602-622. Accessible on the author's page as https://www.egr.msu.edu/~kdeb/papers/k2012010.pdf. (Cited on page 41.)

[20] Igel, C., Hansen, N. and Roth, S., 2005. The multi-objective variable metric evolution strategy part I. Internal Report IR-INI 2005-04, Ruhr-Universitat Bochum. (Cited on pages 40 and 41.)

[21] Lovison, A. and Cardin, F., 2022. A Pareto–Pontryagin Maximum Principle for Optimal Control. Symmetry, 14(6), p.1169. http://doi.org/10.3390/sym14061169 (Cited on pages 44, 45 and 109.)

[22] Zhu, Q.J., 2000. Hamiltonian necessary conditions for a multiobjective optimal control problem with endpoint constraints. SIAM Journal on Control and Optimization, 39(1), pp.97-112. (Cited on pages 45 and 109.)

[23] Collette, Y. and Siarry, P., 2004. Multiobjective optimization: principles and case studies. Springer Science & Business Media. (Cited on pages 40 and 41.)

[24] Amine, K., 2019. Multiobjective simulated annealing: Principles and algorithm variants. Advances in Operations Research, 2019(1), p.8134674. (Cited on page 41.)

[25] Kukkonen, S. and Lampinen, J., 2005, September. GDE3: The third evolution step of generalized differential evolution. In 2005 IEEE congress on evolutionary computation (Vol. 1, pp. 443-450). IEEE. (Cited on page 41.)

Chapter 4

Evolutionary reduction of the laser noise impact on quantum gates

Sommaire

This work was published in the proceedings of the conference 'Complex Computational Ecosystems' as:

Ley, T., Leonteva, A.O., Schachenmayer, J., Collet, P. Evolutionary Reduction of the Laser Noise Impact on Quantum Gates. Complex Computational Ecosystems 2023. Lecture Notes in Computer Science, vol 13927. Springer, Cham. https://doi.org/10.1007/978-3-031-44355-8_5

4.1 Abstract

As the size of quantum hardware progressively increases, the conjectured computational advantages of quantum technologies tend to be threatened by noise, which randomly corrupts the design of quantum logical gates. Several methods already exist to reduce the impacts of noise on that matter. However, a reliable and user-friendly one to reduce the noise impact has not been presented yet. Addressing this issue, this paper proposes a relevant method based on evolutionary optimisation and modulation of the gate design. This method consists of two parts : a model of quantum gate design with time-dependent noise terms, parameterised by a vector of laser phases, and an evolutionary optimisation platform aimed at satisfying a trade-off between the gate fidelity and a pulse duration-related metric of the time consuming simulation model. This feature is the main novelty of this work. Another advantage is the ability to treat any noise spectrum, regardless of its characteristics (e.g., variance, frequency range, etc). A thorough validation of the method is presented, which is based on empirical averaging of random gate trajectories. It is shown that evolutionary based method is successfully applied for noise mitigation. It is expected that the proposed method will help designing more and more noise-resisting quantum gates.

4.2 Introduction

Quantum computing uses the laws of quantum physics to perform logical operations inaccessible to classical means: e.g., superposition and entanglement. In theory, an ideal quantum computer would solve some mathematical problems with exponentially fewer computing resources than any classical counterpart [1]. It applies to practical fields such as quantum chemistry [2], materials science [3], and cryptography [4].

However, quantum systems are highly prone to noise, coming from interactions with the environment, or from imperfections in the control apparatus (e.g., laser, electronic devices). Due to the analog nature of those quantum systems, the noise critically affects the performance of quantum computers. Moreover, the noise also dramatically reduces the quantum computational power, and would prevent quantum computers from reaching any advantage on practical problems with regard to the classical counterparts. This can be seen by the following result: assuming a constant error rate and without any effective correction strategy, a large variety of random quantum circuits (i.e programs) are efficiently simulated by classical computers [5], [6], [7], which would be far more difficult without noise. The ability to overcome noise is the central question raised by quantum computing skeptics [17]. Indeed, quantum circuits (programs) solving relevant problems, require millions of elementary quantum gates (logical operators). Thus, even with small noise rates, they need efficient strategies to suppress errors. The latter exist in theory, but they require average gate infidelity much lower than what current noise

levels impose. Reducing the impact of noise on the gate design is a highly non-trivial challenge, for which several solutions have been proposed: e.g., [14], [15], [16]. Generally, these methods model the gate infidelity (inaccuracy) with equations that require long simulations, and do not cope well with a high standard deviation or high frequency fluctuations in the noise.

Addressing these issues, in this work, we focus on adapting a flat laser pulse suffering amplitude and dephasing noises, in order to implement a Rydberg blockade two-qubit gate, which is specific to the cold-atoms hardware. The theoretical information about that hardware is largely presented in the literature: e.g., [20]. In order to reduce the noise and simultaneously, to reduce the time spent in the Rydberg state, we formulate a quantum gate design as a continuous non-constrained bi-objective problem based on a time-expensive computer simulation. Taking into account the number of objectives (2) and the large number of decision variables (51), we set 100 candidate solutions in the population, where a computation time of around 1 minutes per each solution. In order to reduce the computational complexity, we use the CPU-parallel implementation of the multi-objective algorithm Non-dominated Sorting Genetic Algorithm III (NSGA-III) provided by the EASEA (*EAsy Specification of Evolutionary Algorithms*) platform described in [19].

To our best knowledge, this work is the first attempt to present a quantum gate design task as a continuous large-scale multi-objective problem, solved by a parallel optimisation algorithm. Our method is not limited by any characteristic of the noise, and allows us to select a pulse sequence (laser phases) and duration to implement the target gate. Moreover, EASEA provides different single and multi-objective optimisation algorithms, which can be easily applied for the problem of laser noise reducing in the further researches in order to investigate an impact of different optimisation techniques on the performance in terms of the accuracy and speed-up.

This paper is organized as follows. Section 2 briefly outlines the related works. Section 3 describes the problem definition. Section 4 presents the proposed method based on the evolutionary optimisation. The experiments and results are shown in Section 5. Section 6 concludes this article.

4.3 Related works

In this section, we briefly present the related works addressing the problem of laser noise impact reduction for quantum gate design presented in the literature. They can be broadly divided in four categories:

Noise mitigation It consists of repeating some gates throughout the circuit, in order to reduce the influence of noise on the computation result [8]. This technique generally assumes that the gate noise is uncorrelated in time. However, correlated

noise has been shown to bypass the current noise mitigation strategies, as experimentally proved in [9]. Thus, non-trivial correlations in time must be accounted for, when analysing noise impact on quantum computation.

Robust optimal control As outlined in [10], it converts noise impact minimisation, into an optimal control problem in an extended space with appropriate boundary conditions, and derives semi-analytical solutions to that problem. However, it only treats time-independent noise in a single direction (e.g amplitude or dephasing). Adapting that technique to time-dependent noise would increase the problem dimension exponentially in the number of harmonics, when considering all the cross products between operators. A slight variation of this approach is "inverse geometric optimisation", as introduced in [11]. The idea is to convert noise impact reduction into a Euler-Lagrange problem, leading to a second order non-linear differential equation. This approach allows to directly track the trajectory of the quantum state in a semi-analytical formulation. However, this method cannot be currently be adapted to either dephasing and/or time-dependent noise, due do its mathematical formulation [11].

Filter function formalism As outlined in [14] [15] [16], the idea is to formulate the average gate infidelity as a function of both the laser controls and the noise power spectral density. Then, it is optimised with differentiable methods such as gradient descent. However, that method fails when the noise intensity is of order 10% of the laser pulse intensity, such that the implicit first-order truncation does not ensure a low enough infidelity.

Estimating the noise impact with Monte Carlo summation of random gate trajectories and optimising the laser pulse with a separate algorithm In [13], the laser control is decomposed in Fourier series, and the Fourier coefficients are optimized through the Nelder-Mead algorithm. In their work, dephasing noise is considered relative, like amplitude noise. This stands in contradiction to the convention in the literature. Our present work follows a similar path, but characterises the laser control with flat pulses, whose phases are optimised by a multi-objective algorithm.

4.4 Problem definition

4.4.1 Context

Optically-trapped cold atoms constitute a promising hardware architecture for long term quantum computing due to long qubit life time, maturity of laser technology and flexibility of two-qubit interactions [20]. The idea of this technology is to encode bits

as two specific orbitals, called "0" and "1", populated by the valence electron of some chosen special atomic species [20].

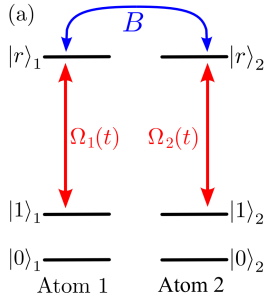


Figure 4.1: Atomic transitions and Rydberg interaction. Illustration adapted from [4].

Upon interaction with a nearby electromagnetic field, electrons switch between orbitals. This is usually done with laser tuned in resonance with the atomic transition frequency between the said orbitals. In the literature, a π-pulse refers to a complete transfer between two orbitals, with a multiplication by the imaginary unit of the relevant orbital coordinate in the quantum state vector. Rydberg states refer to orbitals far away from the nucleus, exhibiting high electric dipole moments. When the resulting electric potential is much larger than the power of the electromagnetic field, the Rydberg blockade prevents two electrons to be simultaneously in a Rydberg state. This happens when atoms are close to each other. The protocol outlined in [18] allows us to design a universal two-qubit gate, using only the Rydberg blockade and four π-pulses between the "1" and a Rydberg orbital. Thus, as we focus on laser noise, we only have to optimise the π-pulse. Figure 4.1 shows a simplified view of the atomic transition described above, where r is a Rydberg state, B is the Rydberg blockade strength and Ω_j is a function characterising the laser pulse.

4.4.2 Mathematical model

Due to the mathematical description of quantum mechanics, the electron's quantum state is represented as a complex unit vector, whose dimension is the number of considered orbitals. In our case, there are 2 : the 1 and r orbitals, as 0 is excluded from the transitions. Since the evolution of the vector is linear and norm-preserving, it is represented as a 2×2 complex unitary matrix $U(t)$, with $U(0)$ being the identity matrix. The interaction between the atom and an electromagnetic field, inducing the orbital transition, is mathematically represented as the following differential equation:

$$\dot{U}(t) = \imath \hat{H}(t) U(t) \tag{4.1}$$

where $\hat{H}(t)$ is the hamiltonian matrix, which corresponds to the energy configuration of the system. The transition between two atomic orbitals is characterised by the complex

Rabi frequency Ω, as shown in the equation below:

$$\hat{H}(t) = \begin{pmatrix} \varepsilon_d(t) & (1+\varepsilon_a(t))\Omega^*(t) \\ (1+\varepsilon_a(t))\Omega(t) & -\varepsilon_d(t) \end{pmatrix} \tag{4.2}$$

where ε_d is the dephasing noise, ε_a is the amplitude noise. In the absence of noise, the pulse will be the fastest, if the Rabi frequency has a constant modulus, as explained in [4]. Since we want the gates to be as fast as possible, in order to avoid other types of noises, we accordingly parameterise the Rabi frequency :

$$\Omega = 2\pi f_{\max} e^{\imath\varphi} \tag{4.3}$$

where $f_{\max}$ is the maximal frequency of the laser pulses and φ is the phases of the laser pulses. In the noiseless case, the phase is constant $\varphi = 0$. Here, we set it as a piece-wise constant function, with a fixed number of slices (N): $\varphi(t) = \varphi_i$ for $t \in [\frac{iT}{N}, \frac{(i+1)T}{N}]$. The solution of the differential Equation 4.1 is a random matrix written as $U(\varphi, t)$. For the sake of clarity, the list of all the mathematical symbols used in this work is summarised in Table 4.1.

This model can be transformed into an optimisation problem. From an optimization point of view, the real-world quantum gate design problem, presented above, is a typical expensive continuous bi-objective optimization problem without constraints. A bi-objective problem aims at optimizing two values of interest (objectives), which are antagonist. The answer is a set of trade-offs solutions (a vector with the input parameters) between the defined objectives. In section 3.3, we present its inputs and objectives.

4.4.3 Values of interest and input parameters

In this study, we consider 2 values of interest (objective functions) and 51 input parameters.

4.4.3.1 Input parameters:

The parameters and their ranges of variation are summed up in Table 4.2, where the following notation is used:

- $\boldsymbol{\varphi}$ is a vector of N elements, which refers to the values of the laser phase in the orbital transitions. In the current model $N = 50$ as a reasonable number required for both processes: simulation and optimisation.
- T is the duration of the model simulation, which has an impact on the time spent by the electron in the Rydberg state (G);
- $f_{\max}$ is the maximal frequency the laser emits to. In our problem definition, we normalise all units with $f_{\max} = 1$, which means that the model can be applied to any value of that frequency. In recent experiments, typically $f_{\max} = 1$ GHz.

Symbol	Significance (unit)
$\|0\rangle$	Ground orbital (–)
$\|1\rangle$	Excited orbital (–)
$\|r\rangle$	Rydberg orbital (–)
$\imath$	Complex imaginary unit (–)
t	Running time ($1/f_{\max}$)
T	Pulse duration ($1/f_{\max}$)
U	Pulse matrix (–)
$\dot{U}$	Derivative of pulse matrix (–)
Ω	Rabi frequency (rad $f_{\max}$)
Ω^*	Complex conjugate thereof (rad $f_{\max}$)
$\hat{H}$	Hamiltonian matrix (rad $f_{\max}$)
U_0	Target pulse matrix (–)
$f_{\max}$	Maximal Rabi frequency (1)
ν_0	Maximal noise frequency ($f_{\max}$)
$\boldsymbol{\varphi}$	Vector of laser phases (rad)
N	Number of time slices/laser phases
F	Pulse infidelity (-)
G	Time in Rydberg state ($1/f_{\max}$)
ε_a	Amplitude noise (rad $f_{\max}$)
ε_d	Detuning noise (rad $f_{\max}$)

Table 4.1: List of symbols used in the problem definition

4.4.3.2 Objective functions

We seek to make $U(\boldsymbol{\varphi}, T)$ (see Equation 1) as close as possible to $U_0 \coloneqq \begin{pmatrix} 0 & \imath \\ \imath & 0 \end{pmatrix}$, because this matrix represents the π-pulse that we are look for, since we base our work on the protocol of [18], which features only π-pulses. This motivates us to define the gate infidelity as follows :

$$F(\boldsymbol{\varphi}, T) \coloneqq \mathbb{E}(1 - \left|\mathrm{Tr}(U(\boldsymbol{\varphi}, T)U_0^\dagger)\right|^2 /4) \tag{4.4}$$

Table 4.2: List of all inputs

Input	Min	Max	Unit
Phases: $\boldsymbol{\varphi}[0-50]$	0	2π	rad
Time: T	1	5	$1/f_{\max}$

A important source of errors in quantum computation, is spontaneous photon emission. In this random process, super excited states like the Rydberg ones, tend to spontaneously revert to a lower-energy state, releasing a photon. This cannot be prevented, so the only way to minimise the risk of occurrence is to minimise the time spent by the electron in the Rydberg orbital. Assuming the electron starts in the 1 orbital, this quantity is given by :

$$G(\boldsymbol{\varphi}, T) := \mathbb{E}\left(\int_0^T |\langle r|U(\boldsymbol{\varphi}, t)|1\rangle|^2 dt\right) \tag{4.5}$$

Let's briefly show that these two quality criteria are antagonistic. Simulating flat pulses leads to the appearance of noise terms of the form $\varepsilon(\frac{jT}{N})$ in the computations. Since the number N of slices is fixed, these terms tend to vanish as T grows to infinity. Thus, the impact of noise can be mitigated by extending the pulse duration, and choosing a trivial $\boldsymbol{\varphi}$. In other words, increasing T tends to reduce F, as well as simplifying the choice of $\boldsymbol{\varphi}$, but it would potentially imply a larger value of G and thus a higher risk of Rydberg state decay. On the other hand, $0 \leq G \leq T$ (because the integrand is lower than 1), so an optimal G would rather be found by decreasing T, but in turn this complexifies the search for optimal phases. Consequently, to find the trade-off between these objectives is required.

4.5 Proposed Method

The goal of this study is to reduce the gate infidelity (i.e the impact of laser noise) by finding the optimal laser phases $\boldsymbol{\varphi}$, without increasing the time spent in the Rydberg state, in order to minimise the risk of Rydberg state decay. Consequently, it will improve the global gate quality. To achieve this goal, we propose a CPU-parallel multi-objective optimisation-based method, since the quantum gate design can be formulated as a continuous bi-objective computationally expensive problem (as was shown above). Indeed, a multi-objective optimization process allows us to find: (i) all possible non dominated solutions – combinations of $\boldsymbol{\varphi}$ and T, which is called the Pareto set; (ii) its image in the objective space, which presents the Pareto front. Thus, we consider to define a set of optimal parameters of the model (i.e., the phases vector of the laser pulses and the duration of the model simulation), by minimizing the objective functions (i.e., overall infidelity and time spent in the Rydberg state), calculated by simulation. The proposed method consists of two following parts:

- A Quantum gate design presented in Section 3.2, which consists of simulating the laser pulse hamiltonian, on noisy trajectories; and computing the objective functions with Monte Carlo average. For the sake of completeness, the noise generation details are provided in online supplementary materials: `https://git.unistra.fr/tsley/thesis`.

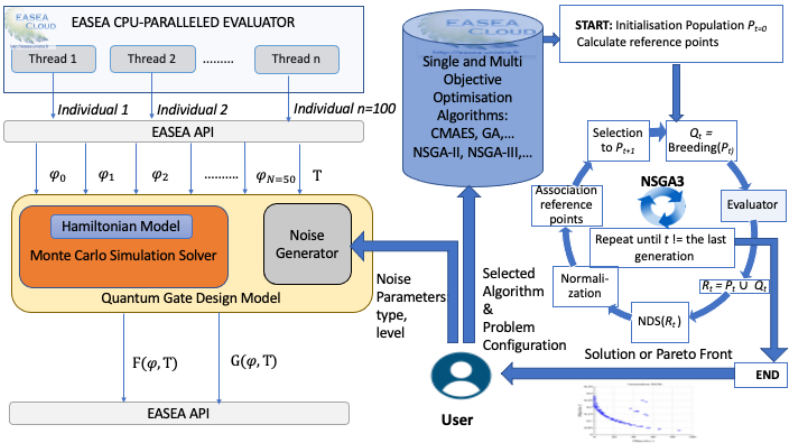


Figure 4.2: Simplified scheme of the proposed method.

- The EASEA platform, which is an open-source transpiler that supports single and multi-objective optimization thanks to its own templates base of different state-of-the-art algorithms [19]. The main advantage of EASEA is that it allows scientists to apply evolutionary algorithms to solve their real-world problems through a user-friendly interface with CPU and/or GPU parallelisation. More precisely, this compiler automatically couples one of the algorithm template files and a problem to be optimised, and provides the C++ based code of the selected algorithm with the integrated problem in it and the compiling file. Consequently, it suits perfectly as a software support for our study. Notably, the quantum gate design model can be modified without changing other parties of the code, due to the very flexible structure of EASEA.

By this concept, we had to select an optimisation algorithm template (**.tpl* file), define the problem in terms of the pseudo-code (**.ez* file) and define the noise parameters, such as the spectrum (white, pink, etc.) and its level w.r.t. maximal intensity of the laser pulse. The simplified scheme of the considered method is schematically depicted in Figure 4.2. It details the programming relation between: (i) the EASEA CPU-parallel evaluator with our quantum gate design model presented on the left and (ii) the selected multi-objective algorithm: Non-dominated Sorting Genetic Algorithm (NSGA-III), shown on the right in Figure 4.2. NSGA-III originally presented in [22], is based on the non-dominated sorting selection mechanism and uses a predefined set of reference points to ensure diversity in the solutions [22]. NSGA-III was selected in this work between other templates, because of the following reasons: (i) it is efficient on several multi- and many-objective problems [24] (in future, the current problem can be extended to a many-objective version); (ii) it does not require any additional hyper-parameters; (iii) it introduces a computationally fast approach, by which the reference

point set is adaptively updated on the fly based on the association status of each reference point over a number of generations [22], [24]; (iv) it can handle constraints without introducing any new parameter, which can be useful in our next research, where some constraints will be taken into account; (v) it can easily be scaled for solving single-objective problems [24]. The following steps explain how this method performs:

1. The selected optimization algorithm, e.g., NSGA-III, starts by randomly generating the initial set of solutions, where each solution consists of the parameters presented in Table 4.2;

2. Its main cycle is repeatedly executed until the last generation (see the right part of Figure 4.2);

3. NSGA-III creates the new child population Q_t from the parent population P_t by the classical breeding operators used in NSGA-III [22]: Simulated Binary crossover (SBX) crossover and Polynomial mutation operator.

4. The *EASEA Evaluator* executes 100 threads to compute F and G in parallel (see the left part of Figure 4.2). Each executed thread works out the model code with its own combination of input parameters and generated noise.

5. Each thread returns the output values (the overall infidelity (F) and the time spent in the Rydberg state (G)) back to the NSGA-III.

6. NSGA-III performs according to its description presented in [22] and goes to the next generation.

Also, important to emphasise that NSGA-III of EASEA is designed for CPU parallel use (several hundred computing threads), which allows us to reduce the runtime.

4.6 Experimental analysis

4.6.1 Experimental design

All values of the setting parameters (both for the simulation model and for the optimisation algorithm) presented below, are the same during all provided experiments. For all experiments we use the problem definition presented in Section 3.3. The decision variables are defined as the presented input parameters. For single-objective study, we use only F as a single optimisation function and applied the well-known Covariance Matrix Adaptation Evolution Strategy (CMA-ES) algorithm [23] from the EASEA plateform. For multi-objective study, as it was presented above, the NSGA-III algorithm is applied.

4.6.1.1 Model Settings

In this paper, we chose to work with $1/f$ noise, also known as pink or flicker noise, since it is widely present in electronic devices, and thus in the quantum hardware that they control. So, we generate 10^6 samples thereof, each time as a sum of sines and cosines with 25 uniformly log-spaced harmonics, in the interval $[f_{\max}, 100 f_{\max}]$, thereby approximating a high-frequency noise. Remind that $f_{\max}$ is the characteristic unit of the problem, and is set to 1. Remind that the optimisation target is a π-pulse to implement a Rydberg blockade two-qubit CZ gate. We run three tests with respective noise levels of 10%, 20% and 30% of the maximal laser pulse intensity.

4.6.1.2 Optimisation Settings

For all experiments, the number of decision variables is 51, according to the problem definition. After many tests, the best parameter settings of CMA-ES bas been found to be: number of generations = 300, population size = 100 and number of threads = 100. The parameter settings of NSGA-III are presented in Table 4.3, where p_c

Table 4.3: Parameter settings of NSGA-III

Parameter	p_c	η_c	p_m	η_m	M	P	H	N	G	Th
Value	1.0	30	1/51	20	2	99	100	100	200	100

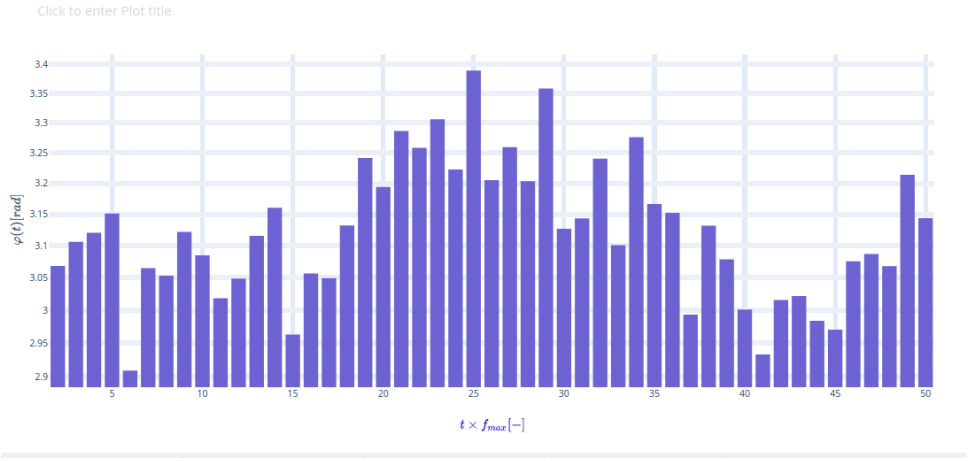
is Simulated Binary (SBX) crossover probability, η_c is crossover distribution index, p_m is polynomial mutation probability, η_m is polynomial distribution index, M is the number of objectives, P is a given integer value, which refers to the number of divisions considered along of each objective axis, H is the number of reference points, N is the population size, G is the number of generation and Th is the number of threads. The selection of P, H and N was made according to the rules defined in [22]: (i) if a problem has a smaller number of objectives, a larger number of P is required (100 is recommended value in bi-objective case); (ii) H is calculated by the approach of Das and Dennis [25]; (iii) the population size is set as the smallest multiple of four higher than the number of reference points (in our case N=H).

4.6.2 Result analysis

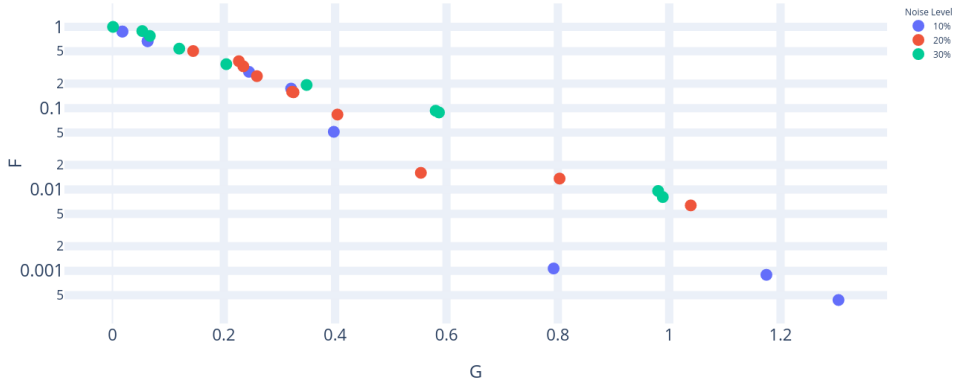
The obtained results are shown in Figure 4.3 and described below.

4.6.2.1 Single-objective

This study started with a single-objective optimisation of F, with a fixed T = 1 and a noise intensity of 10%. After 136 generations of 100 candidates, the algorithm CMA-ES found a solution F = 0.005052. This solution in terms of the coordinates of $\boldsymbol{\varphi}$ is



a) Best solution



b) Pareto Front

Figure 4.3: Obtained results.

presented in log scale in Figure 4.3 (a). This result is better than the gate infidelity $F = 0.0055$ obtained with the naive choice of setting all the 50 phase to be zero: $\boldsymbol{\varphi} = \mathbf{0}$. This result confirmed that artificial evolution can bring some improvements in robust gate designs, and motivated us to explore multi-criteria optimisation for simultaneous antagonistic problems.

4.6.2.2 Multi-objective

The Pareto fronts obtained for three selected noise levels %10 (blue), %20 (red) and %30 (green) of $2\pi f_{\max}$ are presented in Figure 4.3 (b), where the F axis is presented in log scale. As seen from Figure 4.3 (b), the lowest minimum of F was obtained in the case of 10% noise level (presented in blue).

For the rightmost element of the Pareto front, the pulses we find result in a compromise between an infidelity slightly higher than the naive case, and a shorter time spent in the Rydberg state, signalling a preference in reducing Rydberg decay risk. For elements of the leftmost Pareto front, the situation is reversed. This shows that NSGA-III drives the solutions to satisfactory compromises, which should be investigated w.r.t. possible values of the Rydberg decay rate. For all cases, the shape of the Pareto Front is slightly discontinuous, however the solutions are not irregularly distributed in the objective space. It shows that for all selected noise levels, NSGA-III is able to find solutions in the different regions of the objective space, providing the possible variations of trade-off between the overall infidelity and the time spent in the Rydberg state. We assume that these results can be improved by increasing the population size and the number of generations.

4.7 Conclusion

The multi-objective based method we propose aims at introducing a reliable gate design, which allows us to reduce both the laser noise impact and the risk of Rydberg state decay. As a result, this method provides a variety of trade-offs between these two objectives, within a reasonable run-time, thanks to CPU-parallelization. It can be easily applied to different noise and laser pulses models. The problem studied here focuses on decomposing a flat pulse affected with pink noise on both amplitude and phase with different noise intensity levels. With the trust of the multi-objective optimization method, thanks to EASEA platform, it is now possible to look for interesting gate design features for a wide variety of noise with a reduced run-time. Moreover, as the EASEA platform provides different single and multi-objective optimisation algorithms, which can be easily coupled with the problem of laser noise reduction, we prospect to investigate the impact of different optimisation techniques on the performance in terms of the accuracy and speed up in further works. Also, further research with different noise types would be considered. However, our work inherited the drawbacks of the protocol outlined in [18], in the sense that its setup would require an individual laser for each atom, which is difficult to attain. Therefore, beyond simply noise-corrected π-pulses, future works will involve a global pulse model with higher degrees of freedom, similarly to [4]. This would correspond to a single laser for all the atoms, able to design multi-qubit Rydberg blockade gates. It is expected that this method can help to make a qualitative step toward developing robust quantum gates.

4.8 Supplementary material

4.8.1 Symbols table

Symbol	Significance	Dimension (Unit)
$\|0\rangle$	Ground orbital	Dimensionless
$\|1\rangle$	Excited orbital	Dimensionless
$\|r\rangle$	Rydberg orbital	Dimensionless
$\imath$	Complex imaginary unit	Dimensionless
t	Running time	Time $(1/f_{\max})$
T	Pulse duration	Time $(1/f_{\max})$
U	Pulse matrix	Dimensionless
$\dot{U}$	Derivative of pulse matrix	Dimensionless
Ω	Rabi frequency	Angle frequency (rad $\cdot$ Hz)
Ω^*	Complex conjugate of the latter	Angle frequency (rad $\cdot$ Hz)
$\hat{H}$	Hamiltonian matrix	Dimensionless
U_0	Target pulse matrix	Dimensionless
ν_0	Maximal value of the Rabi frequency	Inverse time (Hz)
$f_{\max}$	Maximal frequency reached by the noise	Inverse time (Hz)
φ	Laser phase	Angle (rad)
F	Pulse infidelity	Dimensionless
G	Time in Rydberg state	Time $(1/f_{\max})$
ε_a	Amplitude noise	Angle frequency (rad Hz)
ε_d	Detuning noise	Angle frequency (rad Hz)

Table 4.4: List of symbols

4.8.2 Noise generation

Pink noise is also referred to as "$1/f$ noise", since its power spectral distribution is roughly the inverse function, though with a minimal and maximal frequency cutoffs, since the inverse function is not integrable. Being also named "flicker" noise, it is widely present in electronics, including the ones that control quantum devices. Existing methods to compute pink noise with trigonometric series present a lot of limitations : for example they require a large number of harmonics to ensure convergence to the desired spectrum, and thus are not presented explicitly in the literature. The following method accurately computes pink Noise with a limited number of terms. At the best of our knowledge, it has not been presented elsewhere.

$$C(h) = \int_0^{\log(\nu_{\max})} \cos(2\pi e^x h)dx = \log(\nu_0) \int_0^1 \cos(2\pi \nu_0^s h)ds$$

Indeed, this integral can be approximated up to first order as :

$$\frac{C(h)}{C(0)} = C_M(h) + O\left(\frac{1}{M}\right)$$

Where

$$C_M(h) := \frac{1}{M} \sum_{n=0}^{M-1} \cos\left(2\pi \nu_0^{\frac{n}{M}} h\right)$$

Thus, if (a_n, b_n) are two i.i.d sequences of standard normal random variables, then the expression

$$X(t) := \frac{1}{M} \sum_{n=0}^{M-1} \cos\left(2\pi \nu_{\max}^{n/M} t\right) a_n + \sin\left(2\pi \nu_{\max}^{n/M} t\right) b_n$$

defines a Gaussian process such that $\mathbb{E}(X(t+h)X(t)) = C_M(h)$, thereby providing a good approximation to pink noise.

4.8.3 Optimal coordinates of φ

The phase vector found by the single-objective optimisation mentioned in subsection 5.2 is :

$$\varphi = ($$
$$3.1117210566, 3.0681531508, 3.1062684923, 3.1204162648, 3.1511899718,$$
$$2.9093600733, 3.0648268235, 3.0531992808, 3.1219826294, 3.0848634591,$$
$$3.0181634846, 3.0485478461, 3.1157769830, 3.1606018942, 2.9632053356,$$
$$3.0567039551, 3.0493017295, 3.1321501012, 3.2415897692, 3.1940974328,$$
$$3.2861696576, 3.2583179501, 3.3060153192, 3.2227544442, 3.3887061970,$$
$$3.2053472380, 3.2594614180, 3.2034174002, 3.3582338303, 3.1268169264,$$
$$3.1433784957, 3.2405135080, 3.1008649385, 3.2758823795, 3.1663607416,$$
$$3.1524414836, 2.9931205862, 3.1318277759, 3.0787100614, 3.0011544184,$$
$$2.9331935381, 3.0155597121, 3.0218345718, 2.9837460312, 2.9705769473,$$
$$3.0755616247, 3.0868778008, 3.0677822108, 3.2140962589, 3.1436841099$$
$$)$$

Bibliography

[1] Aaronson, S. and Arkhipov, A., 2011, June. The computational complexity of linear optics. In Proceedings of the forty-third annual ACM symposium on Theory of computing (pp. 333-342). (Cited on page 50.)

[2] Cao, Y., Romero, J., Olson, J.P., Degroote, M., Johnson, P.D., Kieferová, M., Kivlichan, I.D., Menke, T., Peropadre, B., Sawaya, N.P. and Sim, S., 2019. Quantum chemistry in the age of quantum computing. Chemical reviews, 119(19), pp.10856-10915. (Cited on page 50.)

[3] Kairys, P., King, A.D., Ozfidan, I., Boothby, K., Raymond, J., Banerjee, A. and Humble, T.S., 2020. Simulating the Shastry-Sutherland Ising model using quantum annealing. Prx Quantum, 1(2), p.020320. (Cited on page 50.)

[4] Shor, P.W., 1999. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM review, 41(2), pp.303-332. (Cited on page 50.)

[5] Zhou, Y., Stoudenmire, E.M. and Waintal, X., 2020. What limits the simulation of quantum computers?. Physical Review X, 10(4), p.041038. (Cited on page 50.)

[6] Noh, K., Jiang, L. and Fefferman, B., 2020. Efficient classical simulation of noisy random quantum circuits in one dimension. Quantum, 4, p.318. (Cited on page 50.)

[7] Napp, J.C., La Placa, R.L., Dalzell, A.M., Brandao, F.G. and Harrow, A.W., 2022. Efficient classical simulation of random shallow 2D quantum circuits. Physical Review X, 12(2), p.021021. (Cited on page 50.)

[8] Urbanek, M., Nachman, B., Pascuzzi, V.R., He, A., Bauer, C.W. and de Jong, W.A., 2021. Mitigating depolarizing noise on quantum computers with noise-estimation circuits. Physical Review Letters, 127(27), p.270502. (Cited on page 51.)

[9] Schultz, K., LaRose, R., Mari, A., Quiroz, G., Shammah, N., Clader, B.D. and Zeng, W.J., 2022. Impact of time-correlated noise on zero-noise extrapolation. Physical Review A, 106(5), p.052406. (Cited on page 52.)

[10] Van Damme, L., Ansel, Q., Glaser, S.J. and Sugny, D., 2017. Robust optimal control of two-level quantum systems. Physical Review A, 95(6), p.063403. (Cited on page 52.)

[11] Dridi, G., Liu, K. and Guérin, S., 2020. Optimal robust quantum control by inverse geometric optimization. Physical Review Letters, 125(25), p.250403. (Cited on page 52.)

[12] Dawson, C.M. and Nielsen, M.A., 2005. The Solovay-kitaev algorithm. arXiv preprint quant-ph/0505030. (Not cited.)

[13] Huang, C.H. and Goan, H.S., 2017. Robust quantum gates for stochastic time-varying noise. Physical Review A, 95(6), p.062325. (Cited on page 52.)

[14] Le, I.N.M., Teske, J.D., Hangleiter, T., Cerfontaine, P. and Bluhm, H., 2022. Analytic Filter-Function Derivatives for Quantum Optimal Control. Physical Review Applied, 17(2), p.024006. (Cited on pages 51 and 52.)

[15] Hangleiter, T., Cerfontaine, P. and Bluhm, H., 2021. Filter-function formalism and software package to compute quantum processes of gate sequences for classical non-Markovian noise. Physical Review Research, 3(4), p.043047. (Cited on pages 51 and 52.)

[16] Kang, M., Wang, Y., Fang, C., Zhang, B., Khosravani, O., Kim, J. and Brown, K.R., 2023. Designing filter functions of frequency-modulated pulses for high-fidelity two-qubit gates in ion chains. Physical Review Applied, 19(1), p.014014. (Cited on pages 51 and 52.)

[17] Vardi, M.Y., 2019. Quantum hype and quantum skepticism. Communications of the ACM, 62(5), pp.7-7.

[18] Jaksch, D., Cirac, J.I. and Zoller, P., 2000. SL Rolston, R. Côté and MD Lukin. Phys. Rev. Lett, 85, p.2208. (Cited on page 50.)

[19] Collet, P., Lutton, E., Schoenauer, M., Louchet, J. (2000, September). Take it EASEA. In PPSN VI. (Cited on pages 53, 55 and 61.)

[20] Morgado, M. and Whitlock, S., 2021. Quantum simulation and computing with Rydberg-interacting qubits. AVS Quantum Science, 3(2), p.023501. (Cited on pages 51 and 57.)

[21] Jandura, S. and Pupillo, G., 2022. Time-optimal two-and three-qubit gates for Rydberg atoms. Quantum, 6, p.712. (Cited on pages 51, 52 and 53.)

[22] Deb, K. and Jain, H. (2013). An evolutionary many-objective optimization algorithm using reference-point-based non dominated sorting approach, part I: solving problems with box constraints. IEEE transactions on evolutionary computation, 18(4), 577-601. (Cited on pages xxiv, 26, 53, 54, 61, 79, 81 and 114.)

[23] Hansen, N. and Auger, A. (2011, July). CMA-ES: evolution strategies and covariance matrix adaptation. In Proceedings of the 13th annual conference companion on Genetic and evolutionary computation (pp. 991-1010). (Cited on pages 57, 58 and 59.)

[24] Seada, H. and Deb, K. (2015, March). U-NSGA-III: a unified evolutionary optimization procedure for single, multiple, and many objectives: proof-of-principle results. In Evolutionary Multi-Criterion Optimization: 8th International Conference, EMO 2015, Guimarães, Portugal, March 29-April 1, 2015. Proceedings, Part II (pp. 34-49). (Cited on page 58.)

[25] Das, I. and Dennis, J. E. (1998). Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. SIAM journal on optimization, 8(3), 631-657. (Cited on pages 57 and 58.)

(Cited on page 59.)

Chapter 5

Classical emulation of quantum systems

Sommaire

As said in Chapter 1, the classical simulation of general quantum many-body systems is fundamentally asymptotically inefficient, and this inefficiency is the very motivation behind the construction of quantum computers. However, reliable quantum computers are still a long way to come and before they are available, efficient classical algorithms are still required to numerically analyse quantum physical systems in general. At a deeper level, classical algorithms to simulate quantum systems are integrated in the race towards the construction of reliable quantum computers, even if they work only for a limited system size or under any other fundamental limit, and so for those reasons:

- As a crucial step in their numerical analysis of potential future quantum computing hardware, classical simulation algorithms help probing their properties and reliability.

- Quantum computers are extremely prone to errors due to noise and decoherence, so their outputs must be verified by classical means. We acknowledge that for general system sizes, there effectively exist classical verification protocols not requiring any quantum computational power on the verifier side [1], but they are probabilistic and built on Post-Quantum Cryptographic conjectures that may not hold forever. As long as the task is tractable (i.e when small or moderately large machines are considered), the simplest verification is instead to classically simulate the quantum computations and compare the results upon measurements.

- A recent trend in QEC has been to construct non-trivial codes and recovery maps through numerical optimization in order to obtain higher error thresholds (the number below which quantum errors are corrected faster than they occur, allowing clean computations), we refer to chapter 8 for more details about this. Therefore, for small code sizes, in order to verify those non-trivial steps, a simple solution is to classically emulate the circuits involved and compare the results again. Also, precisely estimating the error threshold of QEC codes requires to classically simulate quantum circuits too.

- The development of numerical classical simulation algorithms sheds light on the boundary between classical and quantum computational powers, and we suggest that by identifying the critical parameters leading to quantum advantage, it would also help identiying useful quantum algorithms.

This chapter 5 introduces two classical methods to simulate quantum circuits used in chapter 6. Both can be applied to noiseless circuits through pure states and to noisy circuits through mixed states. Section 5.1 introduces the operator basis and the statevector simulation methods. Section 5.2, the main part of this chapter 5, introduces the Matrix Product States formalism and the updates algorithms.

5.1 Naive exact method

Let $d, n \in \mathbb{N}$. The naive yet exact way to emulate a quantum system of wavefunction $|\psi\rangle \in \mathbb{C}^{d^n}$ is to track all the coefficients $\langle \boldsymbol{b}|\psi\rangle$ ($\boldsymbol{b} \in [d]^n$). This reasoning applies to the emulation of open systems too, though the density operator $\hat{\rho}$ is "flattened" as $|\rho\rangle\!\rangle$ in a suitable operator basis, orthogonal in the *Hilbert-Schmidt* inner product so equipped with the braket : $\langle\!\langle A_1|A_2\rangle\!\rangle = \mathrm{Tr}\left(A_1^\dagger A_2\right)$. When $d = 2$, the two choices of basis are:

- The standard operator basis:

$$|0\rangle\!\rangle = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\ |1\rangle\!\rangle = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\ |2\rangle\!\rangle = \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\ |3\rangle\!\rangle = \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$$

  In this basis, the conversion of operations from the standard ket basis is quite straightforward since $|U\rho U^\dagger\rangle\!\rangle = (|U\rangle\!\rangle\langle\!\langle U|)|\rho\rangle\!\rangle$.

- Pauli matrices:

$$|0\rangle\!\rangle = \frac{\sqrt{2}}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\quad |1\rangle\!\rangle = \frac{\sqrt{2}}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\quad |2\rangle\!\rangle = \frac{\imath\sqrt{2}}{2}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\quad |3\rangle\!\rangle = \frac{\sqrt{2}}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

  A significant advantage here is that hermitian operators (so density operators too) have real coordinates in this basis: $\langle\!\langle \rho|k\rangle\!\rangle \in \mathbb{R}$. This representation is further especially useful if the circuit is only made of Pauli operations, and/or if one is interested into computing Pauli observables (terms of the form $\mathrm{Tr}\left(\rho\sigma^b\right)$) as in chapter 6, since these quantities are then simply coordinates of the evolved operator. However, applying operations in this basis is a bit trickier than in the standard one, as linear maps between operators must be converted in matrices in the Pauli basis first (see below).

A linear map ϕ between operators is also "flattened" into a matrix through :

$$\hat{\phi} = \sum_{k,k'=0}^{d^2-1} \langle\!\langle k'|\phi(|k\rangle\!\rangle)\ |k\rangle\!\rangle\langle\!\langle k'|$$

So that for any operator $\hat{A}$, $|\phi(\hat{A})\rangle\!\rangle = \hat{\phi}|A\rangle\!\rangle$. That being said, in the rest of this chapter we use the simple Dirac braket notation $\langle\cdot|\cdot\rangle$, regardless of the variables inside being pure or mixed states. Notice that in an operator basis, pure states are unit vectors (i.e have norm 1) while strictly mixed states have a norm strictly lower than 1. Let $A \in \mathbb{C}^{(d^m,d^m)}$ represent a linear m-sites operation. To apply A at sites $l_1, \cdots, l_m \in [n]$, the coordinates are updated through:

$$\forall \boldsymbol{b} \in [d]^n,\ \langle \boldsymbol{b}|\psi\rangle \leftarrow \sum_{b'_{l_{1:m}} \in [d]^m} \langle b'_{l_{1:m}}|A|b_{l_{1:m}}\rangle\ \langle \cdots b'_{l_1} \cdots b'_{l_m} \cdots |\psi\rangle$$

5.2 Approximate emulation with Matrix Product States

5.2.1 Motivation, definition and construction

As we saw in Chapter 2, the Schmidt decomposition is fundamentally tied to the entanglement building in a composite quantum system, which in turn is considered as the key factor behind quantum computing's performance. Therefore, in order to numerically analyse this entanglement and identify the boundary between classical and quantum computing, it is interesting to construct a representation of quantum states based on iterated Schmidt decompositions, whose complexity would thus scale with the maximum prescribed Schmidt rank accross the site splittings, instead of the general system dimension. Indeed, this representation was first constructed in [3]. Let $d, n \in \mathbb{N}$, $|\psi\rangle \in \mathbb{C}^{d^n}$ represent the state of an n-qudits system, and let $\chi \in \mathbb{N}$ be its *bond dimension*, the maximal Schmidt rank accross the splittings. For $l \in [n]$, let $(|u_{\alpha_l}^{[:l]}\rangle, |u_{\alpha_l}^{[l:]}\rangle)$ respectively be the and left/right Schmidt vectors of ψ at the splitting between $\mathcal{H}_0 = \mathbb{C}^{d^l}$ (span of the first l qudits) and $\mathcal{H}_1 = \mathbb{C}^{d^{n-l}}$ (span of the next $n-l$ ones) so that:

$$|\psi\rangle = \sum_{\alpha} \lambda_\alpha |u_\alpha^{[:l]}\rangle \otimes |u_\alpha^{[l:]}\rangle$$

In all the next expressions, α indices are implicitly summed in $[\![1, \chi]\!]$ and b indices are are implicitly summed in $[\![1, d]\!]$. Let's set fictious splittings at the boundary:

- At $l = 0:$ $\lambda_0^{[0]} = 1$, $u_0^{[:0]} = (1)$, $u_0^{[0:]} = \psi$ and for $\alpha > 0:$ $\lambda_\alpha^{[0]} = 0$, $u_\alpha^{[:0]} = 0$, $u_\alpha^{[0:]} = 0$

- At $l = n-1:$ $\lambda_0^{[n-1]} = 1$, $u_0^{[:n-1]} = \psi$, $u_0^{[n-1:]} = (1)$ and for $\alpha > 0:$ $\lambda_\alpha^{[n-1]} = 0$, $u_\alpha^{[:n-1]} = 0$, $u_\alpha^{[n-1:]} = 0$

For $l \in [n-1]$, $\alpha, \alpha' \in [\chi], b \in [d]$, let $\Gamma^{[l]b}_{\alpha,\alpha'} := \langle b, u^{[l+1:]}_{\alpha'} | u^{[l]}_{\alpha} \rangle / \lambda^{[l+1]}_{\alpha'}$. We can write inductively:

$$
\begin{aligned}
|\psi\rangle &= \sum_{\alpha_1} \lambda^{[1]}_{\alpha_1} |u^{[:1]}_{\alpha_1}\rangle \otimes |u^{[1:]}_{\alpha_1}\rangle \\
&= \sum_{\alpha_1} \lambda^{[1]}_{\alpha_1} \left(\sum_{\alpha_0} \sum_{b=0}^{d-1} \lambda^{[0]}_{\alpha_0} \Gamma^{[0]b}_{\alpha_0,\alpha_1} |b\rangle \right) \otimes |u^{[1:]}_{\alpha_1}\rangle \\
&= \sum_{\alpha_1} \lambda^{[1]}_{\alpha_1} \left(\sum_{\alpha_0} \sum_{b_0=0}^{d-1} \lambda^{[0]}_{\alpha_0} \Gamma^{[0]b_0}_{\alpha_0,\alpha_1} |b_0\rangle \right) \otimes \sum_{\alpha_2} \lambda^{[2]}_{\alpha_2} \left(\sum_{b_1=0}^{d-1} \Gamma^{[1]b_1}_{\alpha_1,\alpha_2} |b_1\rangle \right) |u^{[2:]}_{\alpha_2}\rangle \\
&\vdots \\
&= \sum_{\alpha_l} \lambda^{[l]}_{\alpha_l} \left(\sum_{\boldsymbol{b} \in [d]^l} \prod_{j=0}^{l-1} \sum_{\alpha_j} \lambda^{[j]}_{\alpha_j} \Gamma^{[j]b_j}_{\alpha_j,\alpha_{j+1}} |\boldsymbol{b}\rangle \right) \otimes |u^{[l:]}_{\alpha_l}\rangle \\
&= \sum_{\alpha_l} \lambda^{[l]}_{\alpha_l} \left(\sum_{\boldsymbol{b} \in [d]^l} \prod_{j=0}^{l-1} \sum_{\alpha_j} \lambda^{[j]}_{\alpha_j} \Gamma^{[j]b_j}_{\alpha_j,\alpha_{j+1}} |\boldsymbol{b}\rangle \right) \otimes \left(\sum_{b=0}^{d-1} |b\rangle \otimes \sum_{\alpha_{l+1}} \Gamma^{[l]b_l}_{\alpha_l,\alpha_{l+1}} \lambda^{[l+1]}_{\alpha_{l+1}} |u^{[l+1:]}_{\alpha_{l+1}}\rangle \right) \\
&= \sum_{\alpha_{l+1}=0}^{\chi_{l+1}-1} \lambda^{[l+1]}_{\alpha_{l+1}} \left(\sum_{\boldsymbol{b} \in [d]^{l+1}} \prod_{j=0}^{l} \sum_{\alpha_j} \lambda^{[j]}_{\alpha_j} \Gamma^{[j]b_j}_{\alpha_j,\alpha_{j+1}} |\boldsymbol{b}\rangle \right) \otimes |u^{[l+1:]}_{\alpha_{l+1}}\rangle \\
&\vdots \\
&= \sum_{\boldsymbol{b} \in [d]^n} \prod_{j=0}^{n-1} \sum_{\alpha_j} \lambda^{[j]}_{\alpha_j} \Gamma^{[j]b_j}_{\alpha_j,\alpha_{j+1}} |\boldsymbol{b}\rangle
\end{aligned}
$$

Putting this in the form of tensors yields:

$$
|\psi\rangle = \sum_{\boldsymbol{b} \in [d]^n} \prod_{i=1}^{n} \mathrm{Tr}\Big(\Gamma^{[i]b_i} \, \mathrm{Diag}(\lambda^{[i]}) \Big) |\boldsymbol{b}\rangle
$$

Due to its form, this expression is called a *Matrix Product State* (MPS) when applied to pure states, and *Matrix Product Density Operator* when applied to mixed states. Conversely, as long as the orthogonality conditions are satisfied (see below) the MPS representation readily gives the Schmidt decomposition through:

$$
|u^{[:l]}_{\alpha_l}\rangle = \sum_{\boldsymbol{b} \in [d]^l} \prod_{i=1}^{l} \mathrm{Tr}\Big(\Gamma^{[i]b_i} \, \mathrm{Diag}(\lambda^{[i]}) \Big) |\boldsymbol{b}\rangle \quad |u^{[l:]}_{\alpha_l}\rangle = \sum_{\boldsymbol{b} \in [d]^{n-l}} \prod_{i=l+1}^{n} \mathrm{Tr}\Big(\Gamma^{[i]b_i} \, \mathrm{Diag}(\lambda^{[i]}) \Big) |\boldsymbol{b}\rangle
$$

5.2.2 Local operations on MPS

A single-site operation $U \in \mathbb{C}^{(d,d)}$ is applied on index $l \in [\![1, n-1]\!]$ by modifying there only the superscript indexing of the tensor at this index:

$$
\forall b \in [d], \ \Gamma^{[l]b} \leftarrow \sum_{a=0}^{d-1} \Gamma^{[l]a} \langle a|U|b\rangle
$$

Since this is done in all the subscript tensor indices, the complexity of this operation scales in $\Theta(d\chi^2)$ basic arithmetic operations.

Let $V \in \mathbb{C}^{(d^2,d^2)}$ represent a linear two-sites operation. It is applied on sites $l, l+1$ by modifying $\Gamma^{[l]}, \lambda^{[l]}, \Gamma^{[l+1]}$ through another SVD. Remind that at that site, through the Schmidt decomposition:

$$|\psi\rangle = \sum_\alpha \lambda_\alpha^{[l]} |u_\alpha^{[:l]}\rangle \otimes |u_\alpha^{[l:]}\rangle = \sum_{\alpha_0,\alpha_1} |u_{\alpha_0}^{[:l-1]}\rangle \otimes |w_{\alpha_0,\alpha_1}\rangle \otimes |u_{\alpha_1}^{[l+1:]}\rangle$$
$$\text{Where} \quad |w_{\alpha_0,\alpha_1}\rangle = \lambda_{\alpha_0}^{[l-1]} \lambda_{\alpha_1}^{[l+1]} \sum_{\alpha,b_0,b_1} \lambda_\alpha^{[l]} \Gamma_{\alpha_0,\alpha}^{[l]b_0} \Gamma_{\alpha,\alpha_1}^{[l+1]b_1} |b_0 b_1\rangle$$

Therefore, applying V at sites $l, l+1$:

$$|\psi'\rangle := (I^{\otimes m} \otimes V \otimes I^{\otimes(n-m-2)}) |\psi\rangle = \sum_{\alpha_0,\alpha_1} |u_{\alpha_0}^{[:l-1]}\rangle \otimes V |w_{\alpha_0,\alpha_1}\rangle \otimes |u_{\alpha_1}^{[l+1:]}\rangle$$

Define a matrix M with elements:

$$M_{d\alpha_0+b_0, d\alpha_1+b_1} = \langle b_0 b_1 | V | w_{\alpha_0,\alpha_1} \rangle$$

Let $v^{[:l]}, v^{[l:]}$ be the singular vectors of M, and $\lambda'^{[l]}$ be its singular values, and define:

$$\Gamma'^{[l]b_0}_{\alpha_0,\alpha} = \frac{v^{[:l]}_{\alpha, d\alpha_0+b_0}}{\lambda_{\alpha_0}^{[l-1]}} \quad \Gamma'^{[l+1]b_1}_{\alpha,\alpha_1} = \frac{v^{[l:]*}_{\alpha, d\alpha_1+b_1}}{\lambda_{\alpha_1}^{[l+1]}}$$

It follows that:

$$M_{d\alpha_0+b_0, d\alpha_1+b_1} = \sum_\alpha \lambda_{\alpha_0}^{[l-1]} \lambda'^{[l]}_\alpha \lambda_{\alpha_1}^{[l+1]} \Gamma'^{[l]b_0}_{\alpha_0,\alpha} \Gamma'^{[l+1]b_1}_{\alpha,\alpha_1}$$

Plugging this into the expression of $|\psi'\rangle$:

$$\begin{aligned} |\psi'\rangle &= \sum_{\alpha_0,\alpha_1} |u_{\alpha_0}^{[:l-1]}\rangle \otimes \sum_{b_0,b_1} M_{d\alpha_0+b_0, d\alpha_1+b_1} |b_0 b_1\rangle \otimes |u_{\alpha_1}^{[l+1:]}\rangle \\ &= \sum_{\alpha_0,\alpha_1} |u_{\alpha_0}^{[:l-1]}\rangle \otimes \sum_{b_0,b_1,\alpha} \lambda_{\alpha_0}^{[l-1]} \lambda'^{[l]}_\alpha \lambda_{\alpha_1}^{[l+1]} \Gamma'^{[l]b_0}_{\alpha_0,\alpha} \Gamma'^{[l+1]b_1}_{\alpha,\alpha_1} |b_0 b_1\rangle \otimes |u_{\alpha_{l+1}}^{[l+1:]}\rangle \end{aligned}$$

By uniqueness of the SVD in decreasing order, one verifies that this is the correct tensor update and the gate is effectively applied on the MPS. This operation essentially consists in a SVD of a $(d\chi, d\chi)$-matrix , therefore its complexity scales in $\Theta(d^3\chi^3)$ basic arithmetic operations.

5.2.3 Truncation and orthogonality issue

Non-trivial two-sites operations generally increase entanglement in the system so that the Schmidt rank at the splitting grows after the gate. So when imposing a bond

dimension χ, after a two-qubit gate the Schmidt terms must be cut beyond χ and the first χ Schmidt values are renormalized to satisfy $\sum_{\beta=1}^{\chi} \lambda_\beta^2 = 1$. This yields:

$$\forall \alpha \in [\chi],\ \lambda_\alpha \leftarrow \frac{\lambda_\alpha}{(\sum_{\beta=1}^{\chi} \lambda_\beta^2)^{1/2}}$$

When emulating an open quantum system, some two-sites operations might not be unitary: this happens for example if V is the intrisic noise channel associated to a two-qubit qubit evolved from a faulty hardware, as seen in chapter 6. As it typically turns a pure state into a mixed state, it reduces the Hilbert-Schmidt norm: $\langle\!\langle \rho | V^\dagger V | \rho \rangle\!\rangle < \langle\!\langle \rho | \rho \rangle\!\rangle$. After applying it on $\hat{\rho}$ at $l, l+1$ ($l \in [n-1]$), generally for $j \neq l$

$$\nu := \sum_\alpha \lambda_\alpha^{[l]2} < \sum_\alpha \lambda_\alpha^{[j]2} = 1$$

As the normalization condition is not uniformly respected, the MPDO representation is not loyal to the Schmidt decomposition anymore and the tensors are not orthogonal. This is problematic, for example if one wants to track the evolution of the operator space entanglement entropy at some splitting unaffected by V. The key to restore orthogonality is to apply the previous procedure with $V = I$, the two-sites identity matrix, at all the other splittings $j \neq l$. Then, the purity is updated with $\mathrm{Tr}\big(\hat{\rho}^2\big) \leftarrow \nu\, \mathrm{Tr}\big(\hat{\rho}^2\big)$ and the Schmidt values are updated with $\lambda_\alpha^{[l]2} \leftarrow \lambda_\alpha^{[l]2}/\nu$

5.2.4 Inner products with MPS

Consider two quantum states ψ, ψ' with respective MPS representation $(\lambda, \Gamma), (\lambda', \Gamma')$. Their inner product is:

$$\langle \psi' | \psi \rangle = \mathrm{Tr}\left(\sum_{b_0} \Gamma^{[0]b_0} \lambda^{[0]} \dots \left(\sum_{b_{n-1}} \Gamma^{[n-1]b_{n-1}} \lambda^{[n-1]} \lambda'^{[n-1]} \Gamma'^{[n-1]b_{n-1}\dagger} \right) \dots \lambda'^{[0]} \Gamma'^{[0]b_0\dagger} \right)$$

The computational complexity of their inner products scales in $\Theta(nd\chi)$. If $|\psi'\rangle$, $|\psi\rangle$ share the same tensors except for indices in $[\![l, m]\!]$:

$$\langle \psi' | \psi \rangle = \mathrm{Tr}\left(\sum_{b_l=1}^{d} \Gamma^{[l]b_l} \lambda^{[l]} \dots \sum_{b_m=1}^{d} \Gamma^{[m]b_m} \lambda^{[m]} \lambda'^{[m]} \Gamma'^{[m]b_m\dagger} \dots \lambda'^{[l]} \Gamma'^{[l]b_l\dagger} \right)$$

This is known as *tensor contraction*. In particular, to compute expectation values of products of single-qubit operators:

$$\langle \psi | \prod_{t=l}^{m} U_t | \psi \rangle = \sum_{\boldsymbol{\alpha}_{l-1:m+1}} \lambda_{\alpha_{l-1}}^{[l-1]2} \lambda_{\alpha_{m+1}}^{[m+1]2} \prod_{t=0}^{m} \lambda_{\alpha_t}^{[t]2} \sum_{b_t, b_t'=0,0}^{1,1} \langle b_t | U | b_t' \rangle\, \Gamma_{\alpha_{t-1},\alpha_t}^{[t]b_t} \Gamma_{\alpha_t,\alpha_{t+1}}^{[l]b_t'}$$

5.2.5 Computational interest of MPS

As the classical emulation of significant operations detailed above run polynomially in χ, this form is especially interesting in situations where $\chi = \text{poly}(n)$. A wide variety of interesting composite quantum states statisfy area scaling laws: their maximal entanglement entropy (accross sites splittings) scales polynomially in their boundary surface [2], which itself generally scales polynomially in system size. This implies that their maximal Schmidt rank is reasonably bounded in system size, suiting them for MPS emulation. It has also been observed numerically that the Schmidt entropy in random quantum circuits is heavily damped by the noise, with Schmidt values decaying fast. Therefore, such noisy circuits are reasonably classically emulable with MPDO. Recent studies [3] have demonstrated that noisy quantum circuits in one dimension can be efficiently simulated classically due to noise-induced limitations on entanglement growth. Specifically, the entanglement entropy of the Matrix Product Density Operator (MPDO) representation saturates at a finite value, independent of system size, which prevents the exponential growth of computational resources typically required for noiseless quantum circuits. The maximal entanglement entropy depends only on the noise strength, with larger noise rates leading to stronger entanglement suppression. As a result, beyond a critical noise threshold, classical algorithms based on MPDOs can efficiently approximate the circuit dynamics, highlighting a fundamental distinction between the computational power of noisy and noiseless quantum systems.

5.3 Conclusion

The simulation of open quantum systems requires to identify a correct base to expand density operators therein, and various mathematical and computational arguments points at the Pauli basis for so. While this basis allows to express density operators as real-valued statevectors, this exhaustive approach should be reserved to a low number of interacting systems, since it does not leverage any other particular information about the studied variables.

On the other hand, the approach of *Matrix Product States* has been recently gaining a lot of traction due to its success in approximating ground energies and other observables of categories of relevant quantum systems. As it is based on repeated *Schmidt decompositions*, this method artificially restricts the entanglement entropy (EE) of the system numerically analysed by fixing a maximal *Schmidt rank*, yielding computationally cheap and accurate descriptions of systems known their limited EE, such as noisy quantum circuits and systems affected by area scaling laws. More generally (even in the absence of EE bounds), MPS simulations allow to track the evolution of metrics relevant to quantum systems (e.g observables eigenvalues) with regard to EE, which is

widely seen as the main variable behind quantum advantage. Therefore, MPS simulations help identiying, empirically and numerically, the tasks and subjects which would be more prone to showcase a quantum computational advantage. Finally, while MPS methods with fixed low Schmidt rank stll cannot emulate general noiseless quantum circuit, they are relevant to compare properties of noisy quantum circuits.

Bibliography

[1] Mahadev, U., 2022. Classical verification of quantum computations. SIAM Journal on Computing, 51(4), pp.1172-1229. https://doi.org/10.1137/20M1371828, https://arxiv.org/abs/1804.01082. (Cited on page 69.)

[2] Eisert, J., Cramer, M. and Plenio, M.B., 2010. Colloquium: Area laws for the entanglement entropy. Reviews of modern physics, 82(1), pp.277-306. https://doi.org/10.1103/RevModPhys.82.277, https://arxiv.org/abs/0808.3773. (Cited on pages 19 and 75.)

[3] Vidal, G., 2003. Efficient classical simulation of slightly entangled quantum computations. Physical review letters, 91(14), p.147902. (Cited on page 71.)

[4] Huang, X.R., Ding, Z.X., Hu, C.S., Shen, L.T., Li, W., Wu, H. and Zheng, S.B., 2018. Robust Rydberg gate via Landau-Zener control of Förster resonance. Physical Review A, 98(5), p.052324. (Not cited.)

[5] Jaksch, D., Cirac, J.I., Zoller, P., Rolston, S.L., Côté, R. and Lukin, M.D., 2000. Fast quantum gates for neutral atoms. Physical Review Letters, 85(10), p.2208. (Cited on pages 6 and 25.)

[6] Gerry, C.C. and Knight, P.L., 2023. Introductory quantum optics. Cambridge university press. (Cited on pages 22 and 23.)

[7] Zheng, W., Zhang, Y., Dong, Y., Xu, J., Wang, Z., Wang, X., Li, Y., Lan, D., Zhao, J., Li, S. and Tan, X., 2022. Optimal control of stimulated Raman adiabatic passage in a superconducting qudit. npj Quantum Information, 8(1), p.9. (Not cited.)

[8] Noh, K., Jiang, L. and Fefferman, B., 2020. Efficient classical simulation of noisy random quantum circuits in one dimension. Quantum, 4, p.318. (Cited on pages 75 and 79.)

[9] Nielsen, M.A. and Chuang, I.L., 2010. Quantum computation and quantum information. Cambridge university press. (Cited on pages 5, 16 and 25.)

[10] Wootters, W.K., 2001. Entanglement of formation and concurrence. Quantum Inf. Comput., 1(1), pp.27-44. (Cited on pages 19 and 91.)

Chapter 6

The impact of post-gate noise compensation schemes on standard quantum simulation metrics

Sommaire

6.1 Introduction

In this chapter, we derive the full quantum channel representation of a noisy Rydberg-blockade two-qubit controlled phase gate gate using recently proposed such setup. The noise, either in frequency and intensity, arises as time-dependent stochastic fluctuations of the optical control apparitus, leading some atoms to stay in Rydberg states after the gate. To compensate for this effect of the noise, three optical schemes are considered: either a forced optical coherent transition back to a specific computational basis state, or letting the Rydberg state decay to a statistical mixture of both. These three schemes are compared for the two different noise types and different levels thereof. They are numerically tested through simulations of both ferromagnetic and antiferromagnetic kicked transverse Ising models. Starting from these simulations, we compare the noise compensation scheme through the Pauli-ZZ correlations of nearest-neighbors and next-nearest-neighbors qubits, the circuit fidelity and operator entanglement entropies. We show that for each model, choosing the right post-gate projection scheme significantly improve the metrics in question.

The effect of random noise on quantum circuits has long been studied, notably on variational circuits, Haar-distributed random circuits, and evolutions under Heisenberg-type hamiltonians. It is now an established fact that random noise dramatically reduces the computing power of quantum circuits, making them efficiently simulable with classical means, notably Matrix Product Density Operators (MPDO) or as multiple Quantum Trajectories of pure states in Matrix Product States (QT+MPS), with limited bond dimension. However, those studies tend to focus on purely depolarizing, dephasing or amplitude damping channels [1, 2, 3, 4].

In this work, we propose to complete the picture by looking at the specific error channels coming from control hardware noise, which are somewhat overshadowed by those studies. More precisely, we consider the promising platform of Rydberg-interacting neutral-atomic qubits. In particular, we take the recently-proposed time-minimal phase gate proposed in [4] as the standard protocol in our study. Since this mostly consists of controlled excitation between a fixed qubit ground state $|1\rangle$ and a Rydberg state $|2\rangle$, the effect of the noise is to leave some atoms in $|2\rangle$ after the gate operation. To mitigate this effect of noise and send the atoms back in the computational basis, three choices are left to the controller:

- Force an optical transition back to $|1\rangle$, the '1' scheme
- Force an optical transition back to $|0\rangle$, the '0' scheme
- Let the atom autonomously decay from$|2\rangle$ to the statistical mixture $(|0\rangle\langle 0| + |1\rangle\langle 1|)/2$ (which is supposed to be the default situation).

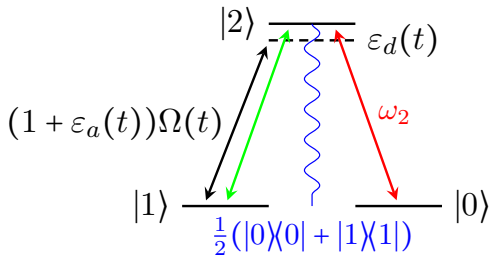


Figure 6.1: Level diagram representation of the post-gate setup, where the black line evokes the original noisy drive. The green line indicates the '1' scheme, the red line indicates the '0' scheme and the blue indicates the '01' scheme.

To simplify the picture, we assume that:

- These three situations are the only possible ways for a excited state to fall back in the computational subspace
- The optical transitions are done in a multi-photon process such that the intermediate orbitals are dipole-forbidden first with $(|0\rangle, |1\rangle)$, then with $|2\rangle$, to keep all the atoms in the computational subspace at the end.
- Those correcting transitions are all noiseless
- Indeed, all other quantum gates besides controlled-phase ones are noiseless

Our fist task is to compute, for each of those scenarii, the two-qubit error channel associated to a noisy local hamiltonian evolution followed by a noise compensation scheme. We consider two different types of noise: frequency and intensity. This computation is eased by a few assumptions on the control noise, that we outline in section 5.

Our second task is to embed those derived channels in the numerical simulation of a relevant hamiltonian, in order to compare the efficiency of the post-gate noise compensation schemes. For this, we choose the kicked transverse Ising model, which is a standard benchmark in probing quantum supremacy claims [5], and in mitigating them through MPS simulations [6]. Upon emulating it, we compute four metrics : circuit fidelity, Pauli-ZZ nearest-neighbour and next-nearest-neighbour correlations, and operator space entanglement entropy (OE). Thus on this generic example, we ask how different noise channels induce a higher fidelity when simulating exact dynamics, how they introduce more or less operator entanglement, how efficient are the MPDO representation of states they induce.

Our work consists in understanding fundamental differences induced by various noise compensation schemes, so we emulate the quantum systems at small scales. Indeed, tangible tendencies already appear at a relatively low number of qubits, so higher dimensions are not crucially needed. For the rest of this work, we remind the definition of Pauli operators:

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^y = \begin{pmatrix} 0 & -\imath \\ \imath & 0 \end{pmatrix} \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

6.2 Noisy evolution

6.2.1 Setup

Consider an atom with ground orbitals $|0\rangle, |1\rangle$ serving as computational states, and a *Rydberg* orbital $|2\rangle$ mediating inter-atomic interactions. To engineer the quantum gates, orbitals $1, 2$ in each atoms are coupled by an external light field and the atoms exert dipole-dipole interactions on each other. Let's recall some results about symmetric Rydberg-blockade two-qubit gates.

Let $\xi \in \mathbb{R}^2$, $K_1 := |01\rangle\langle 01| + |10\rangle\langle 10|, K_2 := |11\rangle\langle 11|, W := |00\rangle\langle 00| + e^{\imath\xi_1} K_1 + e^{\imath\xi_2} K_2$ is called a `CZ`-type gate when $\xi_2 + 2\xi_1 = \pi$. The choice of those angles influences the minimum duration of the gate, leading to potentially fewer decay errors but also requiring to append W with two parallel single qubit rotations $e^{\imath\xi_0|1\rangle\langle 1|}$ to get an actual controlled-phase flip. Notice that the choice of these angles has no influence on the error channels, therefore without loss of generality we just consider them to be $(0, \pi)$.

6.2.2 Noiseless gate

It has been proved in [4], that in the Rydberg blockade regime, W arises from the joint simultaneous evolution of two atoms under the block-diagonal hamiltonian

$$\hat{H}^{(0)} := \sum_{b=1}^{2} P_b \hat{H}_b^{(0)} P_b^\dagger$$

Where the projectors P_b and single qubit hamiltonians H_b are:

$$\begin{aligned} P_1 &:= (|01\rangle + |10\rangle)\langle 0| + (|02\rangle + |20\rangle)\langle 1| \\ P_2 &:= |11\rangle\langle 0| + |w\rangle\langle 1| \quad |w\rangle := \frac{1}{\sqrt{2}}(|12\rangle + |21\rangle) \\ \hat{H}_b^{(0)}(t) &:= \sqrt{b}(\Omega(t)\,|0\rangle\langle 1| + \Omega(t)^*\,|1\rangle\langle 0|) \end{aligned}$$

Here Ω is the complex Rabi frequency, whose modulus is considered constant. The noiseless propagator inherits the block-diagonality from the hamiltonian:

$$U^{(0)}(\tau) := \mathcal{T} e^{\imath \int_0^\tau \hat{H}^{(0)}(t)dt} = \sum_{b=0}^{1} P_b U_b^{(0)}(\tau) P_b^\dagger$$
$$U_b^{(0)}(\tau) := \mathcal{T} e^{\imath \int_0^\tau \hat{H}_b^{(0)}(t)dt}$$

6.2.3 Noisy evolution

The "leakage" to Rydberg oribtals is assumed to stem solely from stochastic fluctuations in the control fields over the atoms. Thus, let's add some noise terms, so the atoms actually evolve under the infinitesimal hamiltonian:

$$\hat{H}^{(0)}(t)dt + \varepsilon \hat{H}^{(1)}(t)dZ_t$$

Where $(Z_t)_t \geq 0$ is a stochastic process specified later, whose differential averages vanishes: $\mathbb{E}(dZ_t) = 0$. Like the control, the perturbation is block-diagonal:

$$\hat{H}^{(1)} = \sum_{b=1}^{2} P_b \hat{H}_b^{(1)} P_b^\dagger \quad \hat{H}_b^{(1)} \in \mathbb{C}^{(2,2)}$$

Two types of noise perturbations are considered, so :

$$\forall b \in \{1,2\},\ \hat{H}_b^{(1)} = \begin{cases} \hat{H}_b^{(0)} & \text{Intensity noise} \\ |1\rangle\langle 1| & \text{Frequency noise} \end{cases}$$

The perturbed propagator of the biatomic system is a matrix stochastic process with:

$$U(\tau, \varepsilon) := \mathcal{T} e^{\imath \int_0^\tau \hat{H}^{(0)}(t)dt + \varepsilon \hat{H}^{(1)}(t)dZ_t}$$

It inherits the block-diagonality of the hamiltonians:

$$U(\tau,\varepsilon) = \sum_{b=1}^{2} P_b U_b(\tau,\varepsilon) P_b^\dagger - P_b P_b^\dagger + I \quad U_b(\tau,\varepsilon) = \mathcal{T} e^{\imath \int_0^\tau \hat{H}_b^{(0)}(t)dt + \varepsilon \hat{H}_b^{(1)}(t)dZ_t}$$

Let's isolate the perturbation term in a separate unitary operator:

$$\tilde{U}_b(\tau,\varepsilon) := U_b(\tau,0)^\dagger U_b(\tau,\varepsilon)$$

So that, conversely:

$$U_b(\tau,\varepsilon) = U_b(\tau,0)\tilde{U}_b(\tau,\varepsilon)$$

Using the interaction picture representation (or from perturbation theory), the first term is another propagator

$$\tilde{U}_b(\tau,\varepsilon) := \mathcal{T} e^{-\imath\varepsilon \int_0^\tau Q_b(t)dZ_t}$$

Where the interaction picture hamiltonian is:

$$Q_b(t) \coloneqq U_b(0,t)^\dagger \hat{H}_b^{(1)}(t) U_b(0,t)$$

Likewise, the perturbation in the biatomic propagator can be isolated in a block-diagonal term :

$$\tilde{U}(\tau,\varepsilon) \coloneqq U(\tau,0)^\dagger U(\tau,\varepsilon) = \sum_{b=0}^{1} P_b \tilde{U}_b P_b^\dagger - P_b P_b^\dagger + I$$

So that

$$U(\tau,\varepsilon) = U(\tau,0)\tilde{U}(\tau,\varepsilon)$$

From now on, let's assume a perfect drive in the noiseless case, i.e $\Pi^{\otimes 2} U(0,\tau)\Pi^{\otimes 2} = W$.

6.3 Quantum channel representation

6.3.1 Projection channels

While Rydberg states tend to live longer than lower fellow excited states, they are still prone to *photon spontaneous emission*, and can be optically brought back to the ground states. Let $\Pi \coloneqq |0\rangle\langle 0| + |1\rangle\langle 1|$ be the projector onto the computational basis. Let $A \in \{\{0\},\{\{1\},\{0,1\}\}$. The "qutrit to qubit" channel of state de-excitation from $|2\rangle$ to elements of A, is:

$$f_A : \hat{\rho} \mapsto \Pi^{\otimes 2}\hat{\rho}\Pi^{\otimes 2} + g_A(\rho) \quad g_A \coloneqq \frac{1}{|A|}\sum_{a\in A} g_{\{a\}}$$

$$g_{\{a\}} : \hat{\rho} \mapsto \sum_{b=0}^{1} |ba\rangle\langle b2|\,\hat{\rho}\,|b2\rangle\langle ba| + |ab\rangle\langle 2b|\,\hat{\rho}\,|2b\rangle\langle ab|$$

Notice that this definition deliberately omits $|22\rangle$ terms because we assume the validity of the *Rydberg blockade* regime, in which those terms have zero coordinates.

The case $A = \{0,1\}$ corresponds to a spontaneous photon emission from $|2\rangle$ to a uniform statistical mixture of $|0\rangle, |1\rangle$. Then, this induces two types of errors in the computational basis, σ^x and σ^z. The singleton case $A = \{a\}$, occurs for example if after the faulty gate, a laser couples $|2\rangle$ to a short-lived orbital quickly decaying only to $|a\rangle$. This two-photon process is also necessary to avoid any excitation from the ground state to the Rydberg manifold again.In this case, there are only σ^z errors if $a = 1$ and only σ^x errors if $a = 0$.

A	$\{0\}$	$\{1\}$	$\{0,1\}$
σ^x	1	0	1
σ^z	0	1	1

6.3.2 Noisy gates

In the rest of this work, let's assume that computational errors arise solely from electrons left in the Rydberg state after the faulty gate, de-excited back into the computational basis after the gate either through spontaneous photon emission or a forced coherent optical scheme. The average noisy gate channel under study is:

$$\Phi : \rho \mapsto f_A(\mathbb{E}(\tilde{U}\rho\tilde{U}^\dagger)) = \mathbb{E}(f_A(\tilde{U}\rho\tilde{U}^\dagger))$$

The expectation and projection commute with each other.

6.3.3 Kraus operators

Let's start with the part still in the computational subspace after the faulty gate. Let ρ be a deterministic density operator, A be any matrix.

$$\tilde{U}\rho\tilde{U}^\dagger = (\tilde{U}-A)\rho(\tilde{U}-A)^\dagger + A\rho(\tilde{U}-A)^\dagger + (\tilde{U}-A)\rho A^\dagger + A\rho A^\dagger$$

Choosing $A \coloneqq \mathbb{E}(\tilde{U})$ yields $\mathbb{E}(A\rho(\tilde{U}-A)^\dagger + (\tilde{U}-A)\rho A^\dagger) = 0$ so that:

$$\mathbb{E}(\tilde{U}\rho\tilde{U}^\dagger) = \mathbb{E}((\tilde{U}-A)\rho(\tilde{U}-A)^\dagger) + A\rho A^\dagger$$

Therefore setting $K_0 \coloneqq \Pi^{\otimes 2}\mathbb{E}(\tilde{U})\Pi^{\otimes 2}$, $B \coloneqq \Pi^{\otimes 2}\tilde{U}\Pi^{\otimes 2} - \mathbb{E}(\Pi^{\otimes 2}\tilde{U}\Pi^{\otimes 2})$ yields

$$\mathbb{E}(\Pi^{\otimes 2}\tilde{U}\rho\tilde{U}^\dagger\Pi^{\otimes 2}) = K_0\rho K_0^\dagger + \mathbb{E}(B\rho B^\dagger)$$

Setting $z_b \coloneqq \langle 0|\tilde{U}_b|0\rangle$, $s_b \coloneqq z_b - \mathbb{E}(z_b)$ for $b \in \{1,2\}$, this results in:

$$K_0 = |00\rangle\langle 00| + \sum_{c=1}^{2}\mathbb{E}(z_c)K_c \quad B = \sum_{c=1}^{2} s_c K_c$$

Developing the terms:

$$B\rho B^\dagger = \sum_{c=1}^{2}|s_c|^2 K_c\rho K_c^\dagger + s_c s_{3-c}^* K_c\rho K_{3-c}^\dagger$$

Let $K_{3+c} \coloneqq (K_1 + \imath^{2(c\%2)+\lceil c/2\rceil}K_2)/2$ for $c \in [\![0,3]\!]$. Developing terms on the right:

$$K_1\rho K_2 = K_3\rho K_3^\dagger - K_4\rho K_4^\dagger + \imath(K_5\rho K_5^\dagger - K_6\rho K_6^\dagger)$$
$$K_2\rho K_1 = K_3\rho K_3^\dagger - K_4\rho K_4^\dagger - \imath(K_5\rho K_5^\dagger - K_6\rho K_6^\dagger)$$
$$zK_1\rho K_2 + z^*K_2\rho K_1 = \Re(z)(K_3\rho K_3^\dagger - K_4\rho K_4^\dagger) + \Im(z)(K_5\rho K_5^\dagger - K_3\rho K_3^\dagger)$$

Therefore, to expand the expectation value in B:

$$B\rho B^\dagger = \sum_{c=1}^{2}|s_c|^2 K_c\rho K_c^\dagger + \Re(s_1 s_2^*)(K_3\rho K_3^\dagger - K_4\rho K_4^\dagger) + \Im(s_1 s_2^*)(K_5\rho K_5^\dagger - K_6\rho K_6^\dagger)$$

Let's turn to the Rydberg state decay part. Let $m_c := 1 - |z_c|^2$ for $c \in \{1, 2\}$. Then, for $b \in \{0, 1\}$ (when the denominator is zero, the numerator is zero too):

$$\frac{\langle b2|\tilde{U}\rho\tilde{U}^\dagger|b2\rangle}{\langle b1|\rho|b1\rangle} = \frac{\langle 2b|\tilde{U}\rho\tilde{U}^\dagger|2b\rangle}{\langle 1b|\rho|1b\rangle} = 2^{-b} m_{b+1}$$

For $a \in \{0, 1\}$, the action of the decay channel is thus:

$$g_{\{a\}}(\tilde{U}\rho\tilde{U}^\dagger) = \sum_{b=0}^{1} 2^{-b} m_b \left(|ba\rangle\langle b1|\,\rho\,|b1\rangle\langle ba| + |ab\rangle\langle 1b|\,\rho\,|1b\rangle\langle ab|\right)$$

The eventual Kraus decomposition of the noise channel is:

$$\Phi(\rho) = \sum_{i=0}^{13} \lambda_i K_i \hat{\rho} K_i^\dagger$$

Where the Kraus operators and coefficients are listed and detailed in the table below:

$\rightarrow$	A	$\{0\}$	$\{1\}$	$\{0,1\}$
i	K_i	λ_i		
1	$\|01\rangle\langle 01\| + \|10\rangle\langle 10\|$	$\mathbb{E}(\|s_1\|^2)$		
2	$\|11\rangle\langle 11\|$	$\mathbb{E}(\|s_2\|^2)$		
3	$(K_1 + K_2)/2$	$\Re(\mathbb{E}(s_1 s_2^*))$		
4	$(K_1 - K_2)/2$	$-\Re(\mathbb{E}(s_1 s_2^*))$		
5	$(K_1 + \imath K_2)/2$	$-\Im(\mathbb{E}(s_1 s_2^*))$		
6	$(K_1 - \imath K_2)/2$	$\Im(\mathbb{E}(s_1 s_2^*))$		
7	$\|00\rangle\langle 01\|$	$\mathbb{E}(m_1)$	0	$\frac{\mathbb{E}(m_1)}{2}$
8	$\|00\rangle\langle 10\|$			
9	$\|10\rangle\langle 11\|$	$\frac{\mathbb{E}(m_2)}{2}$		$\frac{\mathbb{E}(m_2)}{4}$
10	$\|01\rangle\langle 11\|$			
11	$\|01\rangle\langle 01\|$	0	$\mathbb{E}(m_1)$	$\frac{\mathbb{E}(m_1)}{2}$
12	$\|10\rangle\langle 10\|$			
13	$\|11\rangle\langle 11\|$		$\mathbb{E}(m_2)$	$\frac{\mathbb{E}(m_2)}{2}$

6.4 Approximate channel representation

Let $a, b \in \{0, 1\}$. In the form presented above, the averages presented in the coefficients may be hard to compute analytically, because they involve the time-ordered noisy propagator. Therefore, we approximate the latter with perturbation theory. Let's start by truncating its Dyson expansion at second order:

$$\tilde{U}_b(\tau, \varepsilon) = I - \imath\varepsilon R_b^1(\tau) - \varepsilon^2 R_b^2(\tau) + O(\varepsilon^3)$$

Where the operators are:

$$R_b^1(\tau) \coloneqq \int_0^\tau Q_b(t)dZ_t \quad R_b^2(\tau) \coloneqq \int_0^\tau Q_b(t)R_b^1(t)dZ_t$$

Notice that the second order is not only desirable for accuracy reasons, but it is also necessary because the first order terms vanish upon expectation. Letting $r_b^{(k)} \coloneqq \langle 0|R_b^{(k)}|0\rangle$, coefficients of the noisy propagator admit the following Taylor development:

$$\begin{aligned} z_b &= 1 - \imath r_b^1\varepsilon - r_b^2\varepsilon^2 + O(\varepsilon^3) \\ z_a z_b^* &= 1 + \imath(r_b^1 - r_a^1)\varepsilon + (r_a^1 r_b^1 - r_a^2 - r_b^2)\varepsilon^2 + O(\varepsilon^3) \end{aligned}$$

Since $\mathbb{E}(dZ_t) = 0$, then $\mathbb{E}(R_b^1(\tau)) = 0$ and:

$$\begin{aligned} \mathbb{E}(z_b) &= 1 - \mathbb{E}(r_b^2)\varepsilon^2 + O(\varepsilon^3) \\ \mathbb{E}(z_a)\mathbb{E}(z_b)^* &= 1 - \mathbb{E}(r_a^2 + r_b^2)\varepsilon^2 + O(\varepsilon^3) \\ \mathbb{E}(z_a z_b^*) &= 1 + \mathbb{E}(r_a^1 r_b^1 - r_a^2 - r_b^2)\varepsilon^2 + O(\varepsilon^3) \\ \mathbb{E}(s_a s_b^*) &= \mathbb{E}((z_a - \mathbb{E}(z_a))(z_b - \mathbb{E}(z_b))^*) \\ &= \mathbb{E}(z_a)\mathbb{E}(z_b)^* - \mathbb{E}(z_a z_b^*) \\ &= \mathbb{E}(r_a^1 r_b^1)\varepsilon^2 + O(\varepsilon^3) \end{aligned}$$

Therefore, all but the first Kraus operators scale in ε^2:

$$\Phi(\rho) = \tilde{K}_0\rho\tilde{K}_0 + \varepsilon^2 \sum_{i=1}^{13} \tilde{\lambda}_i K_i \rho K_i^\dagger$$

In order to ensure the trace-preserving character of Φ, the first term is modified as:

$$\tilde{K}_0 = |00\rangle\langle 00| + \sum_{b=1}^{2} \left(1 - 2\varepsilon^2\mathbb{E}(r_b^2)\right)^{1/2} K_b$$

$\forall b \in \{1, 2\}$ let $r_b^{12} \coloneqq (r_b^1)^2$, $c_b \coloneqq 2r_b^2 - r_b^{12}$. The Kraus operators and approximate coefficients are listed and detailed in the table below:

→	A	$\{0\}$	$\{1\}$	$\{0,1\}$
i	K_i	$\tilde{\lambda}_i$		
1	$\|01\rangle\langle 01\| + \|10\rangle\langle 10\|$	$\mathbb{E}(r_1^{12})$		
2	$\|11\rangle\langle 11\|$	$\mathbb{E}(r_2^{12})$		
3	$(K_1 + K_2)/2$	$\mathbb{E}(r_1^1 r_2^1)$		
4	$(K_1 - K_2)/2$	$-\mathbb{E}(r_1^1 r_2^1)$		
5	$(K_1 + \imath K_2)/2$	0		
6	$(K_1 - \imath K_2)/2$			
7	$\|00\rangle\langle 01\|$	$\mathbb{E}(c_1)$	0	$\frac{\mathbb{E}(c_1)}{2}$
8	$\|00\rangle\langle 10\|$			
9	$\|10\rangle\langle 11\|$	$\frac{\mathbb{E}(c_2)}{2}$		$\frac{\mathbb{E}(c_2)}{4}$
10	$\|01\rangle\langle 11\|$			
11	$\|01\rangle\langle 01\|$	0	$\mathbb{E}(c_1)$	$\frac{\mathbb{E}(c_1)}{2}$
12	$\|10\rangle\langle 10\|$			
13	$\|11\rangle\langle 11\|$		$\mathbb{E}(c_2)$	$\frac{\mathbb{E}(c_2)}{2}$

Notice that in this second-order approximation, all the operators and coefficients are real-valued, making Φ hermitian: $\Phi = \Phi^\dagger$. This comes from the first-order terms in $\tilde{U}$ averaging into zero, while second order terms are conjugation-invariant anyway.

6.5 Noise specifications

Let's give analytic formulas for R_b^1, R_b^2. Let's assume that the stochastic process Z is correlated, gaussian, stationary and markovian, which are frequent assumptions in quantum noise theory. Up to a rescaling absorbed by ε anyway, this makes Z a linear combination of independent *Ornstein-Uhlenbeck* processes [7] with varying correlation times and volatilities. However, for the sake of simplicity we consider only one term with unit correlation time, so that $Z_0 \sim \mathcal{N}(0, 1/2)$ and Z satisfies the stochastic differential equation:

$$dZ_t = -Z_t dt + dW_t$$

Where W is a Wiener process. However, to simplify the calculations, we approximate this setting by considering instead $Z_0 = 0$. One gets the analytical expressions:

$$Z_t = \int_0^t e^{s-t} dW_s \quad dZ_t = dW_t - \int_0^t e^{s-t} dW_s dt$$

Let $a, b \in \{0,1\}$. Plugging this into the previous expressions of the residuals:

$$R_b^1(\tau) = \int_0^\tau M_b^+(\tau,t) dW_t \quad \text{where} \quad M_b^\pm(\tau,t) := Q_b(t) - \int_t^\tau e^{\pm t-s} Q_b(s) ds$$

So in the higher-left coordinate:

$$r_b^1 = \int_0^\tau \langle 0|M_b^+(\tau,t)|0\rangle \; dW_t$$

Then, using the *Ito isometry*:

$$\mathbb{E}(r_a^1 r_b^1) = \int_0^\tau \langle 0|M_a^+(\tau,t)|0\rangle \langle 0|M_b^+(\tau,t)|0\rangle \, dt$$

Plugging the analytical expressions of R_b^1 and dZ_t in R_b^2:

$$\begin{aligned} R_b^2(\tau) &= \int_0^\tau Q_b(t) R_b^1(t) dZ_t = \int_0^\tau Q_b(t) \int_0^t M_b^+(t,s) dW_s \left(dW_t - \int_0^t e^{u-t} dW_u dt \right) \\ &= \int_0^\tau \int_0^\tau Q_b(t) M_b^+(t,s) \mathbb{1}_{s\leq t} dW_s dW_t - \int_0^\tau Q_b(t) \left(\int_0^t \int_0^t e^{u-t} M_b^+(t,s) dW_s dW_u \right) dt \end{aligned}$$

Taking the expectation value and using the Ito isometry:

$$\mathbb{E}(R_b^2(\tau)) = \int_0^\tau Q_b(t) M_b^+(t,t) dt - \int_0^\tau Q_b(t) \left(\int_0^t e^{s-t} M_b^+(t,s) ds \right) dt$$

Rearranging the integrals:

$$\begin{aligned} \int_0^t e^s M_b^+(t,s) ds &= \int_0^t e^s \left(Q_b(s) - \int_s^t e^{s-u} Q_b(u) du \right) ds \\ &= \int_0^t e^s Q_b(s) ds - \int_0^t \int_s^t e^{2s-u} Q_b(u) du ds \\ &= \int_0^t e^s Q_b(s) ds - \int_0^t \int_0^u e^{2s} ds e^{-u} Q_b(u) du \\ &= \int_0^t e^s Q_b(s) ds - \int_0^t \frac{e^{2u}-1}{2} e^{-u} Q_b(u) du \\ &= \int_0^t e^s Q_b(s) ds - \int_0^t \frac{e^u - e^{-u}}{2} Q_b(u) du \\ &= \int_0^t \frac{e^s + e^{-s}}{2} Q_b(s) ds \\ \int_0^\tau Q_b(t) \int_0^t e^{s-t} M_b^+(t,s) ds dt &= \int_0^\tau Q_b(t) \int_0^t \frac{e^{s-t} + e^{-s-t}}{2} Q_b(s) ds dt \\ &= \int_0^\tau Q_b(s) \int_s^\tau \frac{e^{s-t} + e^{-s-t}}{2} Q_b(t) dt ds \\ &= \int_0^\tau Q_b(t) \int_t^\tau \frac{e^{t-s} + e^{-t-s}}{2} Q_b(s) ds dt \end{aligned}$$

Reminding that $M_b^\pm(t,t) = Q_b(t)$, and accounting for the previous simplification, we recognize familiar terms:

$$\begin{aligned} \mathbb{E}(R_b^2(\tau)) &= \int_0^\tau Q_b(t) \left(Q_b(t) - \int_t^\tau \frac{e^{t-s} + e^{-t-s}}{2} Q_b(s) ds \right) dt \\ &= \int_0^\tau Q_b(t) \left(\frac{M_b^+ + M_b^-}{2} \right) (\tau, t) dt \end{aligned}$$

Hence in the upper-left coordinate:

$$\mathbb{E}(r_b^2) = \langle 0| \int_0^\tau Q_b(t) \left(\frac{M_b^+ + M_b^-}{2} \right) (t,s) dt \, |0\rangle$$

6.6 Simulation details

6.6.1 Hamiltonians to simulate, and their circuits

Consider the transverse-field nearest neighbour Ising model on n qubits:

$$\hat{H} = \sum_{j=1}^{n} \sigma_j^x \pm \sum_{j=1}^{n-1} \sigma_j^z \sigma_{j+1}^z$$

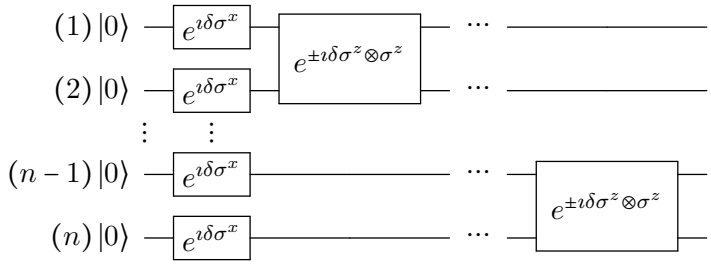
Where $(+,-)$ respectively indicate the *ferromagnetic* and *antiferromagnetic* cases. We seek to emulate the discretized time evolution of both those hamiltonians from 0 to t_f in $m \in \mathbb{N}^*$ angular time steps $\delta := \pi t_f / m$, so to apply the unitary operator $e^{\imath \pi t_f \hat{H}} = (e^{\imath \delta \hat{H}})^m$. Owing to Trotter's approximation, we replace $e^{\imath \delta \hat{H}}$ with $U_x U_z$, where:

$$U_x := \exp\left(\imath\delta \sum_{j=1}^{n} \sigma_j^x\right) = \bigotimes_{j=1}^{n} e^{\imath\delta\sigma^x}$$

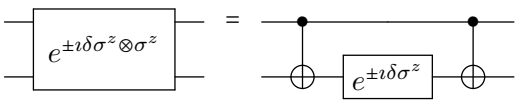
$$U_z := \exp\left(\pm\imath\delta \sum_{j=1}^{n-1} \sigma_j^z \sigma_{j+1}^z\right) = \prod_{j=1}^{n-1} I_{2^{j-1}} \otimes e^{\pm\imath\delta\sigma^z\otimes\sigma^z} \otimes I_{2^{n-j-1}}$$

Since we emulate the kicked transverse Ising model, the Trotter error is not relevant anyway, and we just apply $(U_x U_z)^m$. To run the circuits in the noisy scenarii, the derived noise channels must be inserted around the $e^{\pm\imath\delta\sigma^z\otimes\sigma^z}$ operations (and only them, as we assumed other gates to be noiselss). Let's work that out.

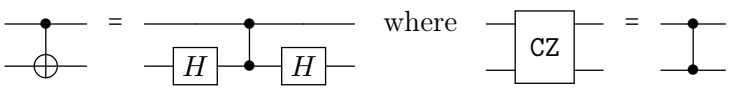
Showing only the two first and two last lines, one layer of the noiseless circuit of $U_x U_z$ is schematically:



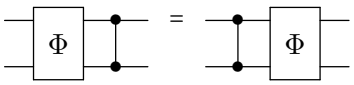
A Pauli-ZZ rotation can be decomposed in terms of controlled-bit flips and Pauli-Z rotation as:



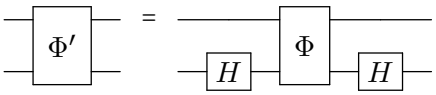
The controlled- bit flip and phased flip are related through:



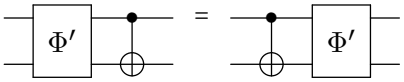
By definition of the error channel derived earlier, W must be appended by Φ, either from the left or the right since both of them are hermitian. Therefore, in second order approximation the average channel of noisy controlled phase flip is:



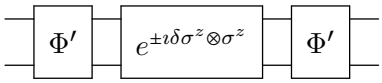
In the σ^x-basis of the second qubit, the noise channel is:



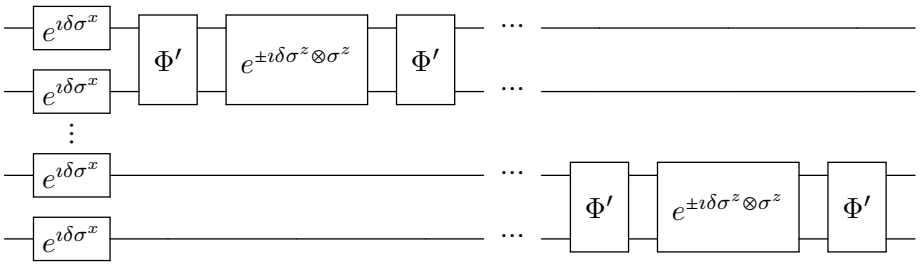
The noisy controlled bit flip is thus:



Therefore, the noisy circuit of $e^{\pm\imath\delta\sigma^z\otimes\sigma^z}$ is:



And one noisy circuit layer is:



6.6.2 Quantities to compute

Let's start with a short note on Operator space entanglement entropy (OE). The *Schmidt decomposition* theorem states that if $\rho \in \mathcal{H}_0 \otimes \mathcal{H}_1$, there exists an integer $\chi \in \mathbb{N}$ named 'Schmidt rank', decreasing positive real numbers named 'Schmidt values' $(\lambda_\alpha)_{\alpha\in[\chi]}$, and two orthonormal families of density matrices named 'Schmidt vectors' $\rho_\alpha^b \in \mathcal{H}_b$ $(b \in \{0,1\},\ \mathrm{Tr}\left(\rho_\alpha^b \rho_\beta^b\right) = \mathbb{1}_{\alpha=\beta})$ such that:

$$\frac{\rho}{\mathrm{Tr}(\rho^2)} = \sum_{\alpha=1}^{\chi} \lambda_\alpha \rho_\alpha^0 \otimes \rho_\alpha^1$$

The *Operator Space Entanglement Entropy* of ρ between $\mathcal{H}_0$ and $\mathcal{H}_1$ is the Shannon entropy of those Schmidt values:

$$S(\rho) = -\sum_{\alpha=1}^{\chi} \lambda_\alpha^2 \log_2(\lambda_\alpha^2)$$

In the pure state case, S is twice the Von Neumann entropy, that quantifies the entanglement in the bipartition of $\mathcal{S}_{01}$ between $\mathcal{S}_0$ and $\mathcal{S}_1$. In the mixed state case, the interpretation does not hold anymore and OE is simply an upper bound on the actual entanglement (it does not quantify it per se). For reference, a genuine measure of entanglement of mixed states is the *entanglement of formation* described in [8] and whose formula is:

$$\mathcal{E}(\hat{\rho}) := \frac{1}{2} \inf \left(\sum_{r=1}^{d^n} p_r S(|\psi_r\rangle\langle\psi_r|) \;\middle|\; p_r \geq 0,\ \sum_{r=1}^{d^n} p_r = 1\ \hat{\rho} = \sum_{r=1}^{d^n} p_r\, |\psi_r\rangle\langle\psi_r| \right)$$

The MPDO decomposition $\rho = \sum_{\boldsymbol{b}\in[d]^n} \mathrm{Tr}\big(\prod_{i=1}^n \mathrm{Diag}(\lambda^{[i]})\Gamma^{[i]b_i}\big) \bigotimes_{i=1}^n \sigma_{b_i}$, with bond dimension χ, readily gives the Schmidt decomposition at each splitting between the i first and $n-i$ last qubits, for $i \in [\![1, n]\!]$. Since $\sum_{\alpha=1}^{\chi} \lambda_\alpha^2 = 1$, $S(\rho) \leq \log_2(\chi)$ and the evolution of $S(\rho)$ asymptotically quantifies the bond dimension required to emulate ρ. Consequently, OE expresses the efficiency of the MPDO representation of an open quantum system, and is therefore related to the hardness of classically approximating them [1, 4]. Therefore, comparing operator entanglements between different schemes leads to identifying which scheme would require the highest bond dimension to classically emulate the evolution under the noisy hamiltonian.

OE can be compared to *trajectory entanglement* (TE) [2]. As a mixed state is generally derived as the statistical average of pure states, its TE is twice the average of their Von Neumann entropies. Mathematically, if $\rho = \mathbb{E}(|\psi\rangle\langle\psi|)$ for a random distribution of pure states, then the TE of ρ is $S^{\mathrm{TE}}(\rho) = \mathbb{E}(S(|\psi\rangle\langle\psi|))$. Just like OE, TE is not a measure of entanglement. In the simulation of the Heisenberg XXZ model, it has been observed on the long run, that under dephasing and spontaneous emission noises, (so different from colored control noise) $S(\rho) < S^{\mathrm{TE}}(\rho)$ indicating that MPDO are a more efficient way to classically emulate quantum dynamics in some regimes[2]. That being said, let's turn to the list of quantities sought to compute through the circuit emulations.

Let L be the quantum channel corresponding to one layer of the noisy circuit, and for $m \in \mathbb{N}$, $\rho_m := L^m(|0\rangle\langle 0|^{\otimes n})$ be the quantum state after m circuit layers, and $|\psi_m\rangle$ result from the noiseless evolution after m layers. In the following lines, let $i = \lfloor n/2 \rfloor$, signaling that we evaluate localized quantities at the middle of the array to avoid any boundary effects. We compare the quantities $c_1(m), c_2(m), F(m), O(m)$, defined as:

- Nearest-Neighbouring ($b = 1$) and Next-Nearest-Neighbouring ($b = 2$) Pauli-ZZ correlations: $c_b(m) = \mathrm{Tr}\big(\sigma_i^z \sigma_{i+b}^z \rho_m\big) - \mathrm{Tr}(\sigma_i^z \rho_m)\,\mathrm{Tr}\big(\sigma_{i+b}^z \rho_m\big)$
- Circuit fidelity $F(m) = \langle \psi_m | \rho_m | \psi_m \rangle$. The overlap between the noisy and noiseless density matrices, it simply measures how much the noise affects the quality of the emulation.
- Operator space entanglement entropy $S(m) = -\sum_{\alpha=1}^{\chi} \lambda_\alpha^2 \log_2(\lambda_\alpha^2)$ where λ are the Schmidt values of the splitting between the spans of the first i and last $n-i$ qubits.

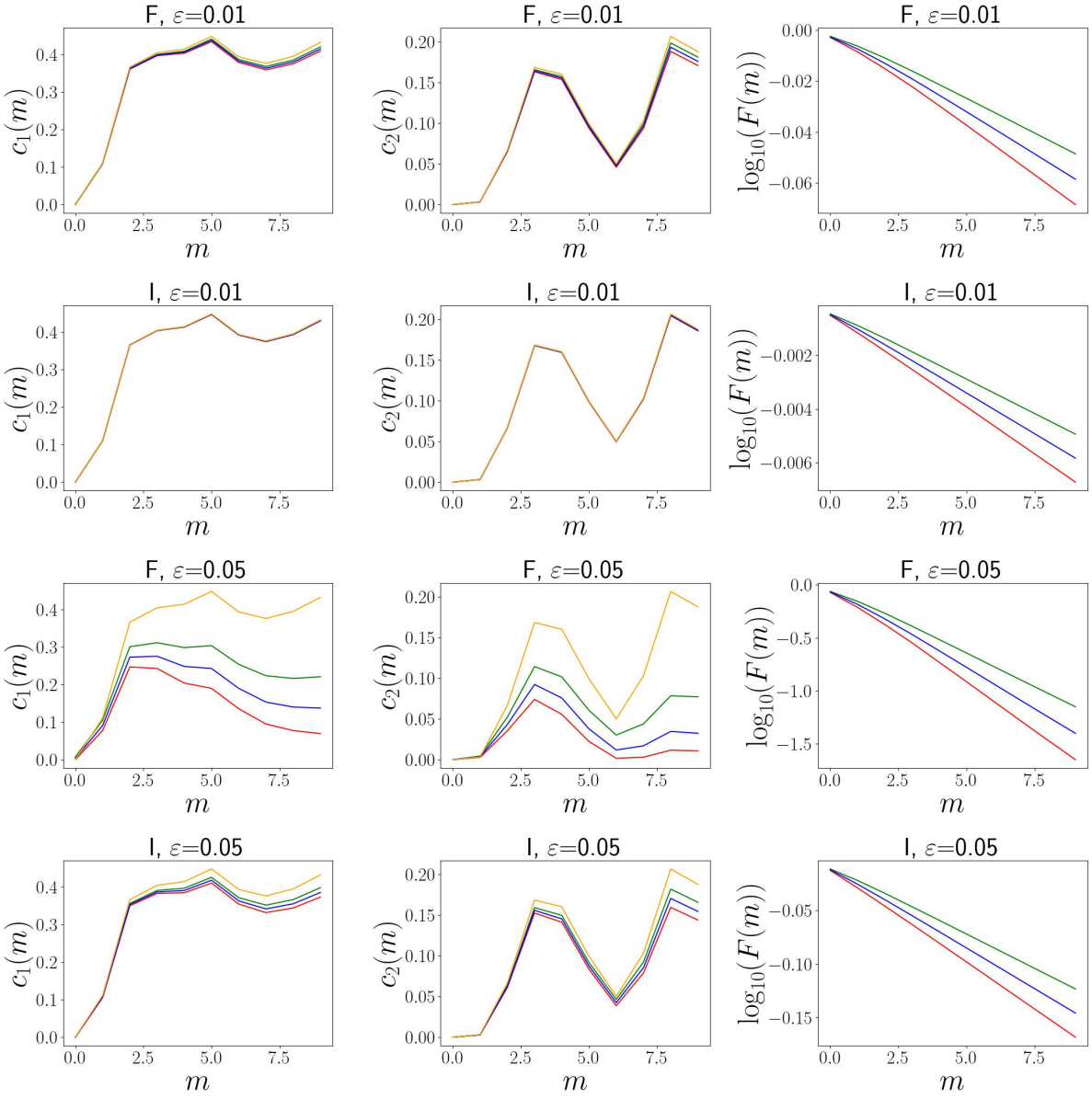


Figure 6.2: Evolution of several simulation metrics under a n = 10-qubits antiferromagnetic Ising hamiltonian. 'F' indicates frequency noise and 'I' indicates intensity noise. The color pattern for the schemes : {(Red, '0'), (Blue, '01'), (Green, '1'), (Orange, 'Noiseless')}

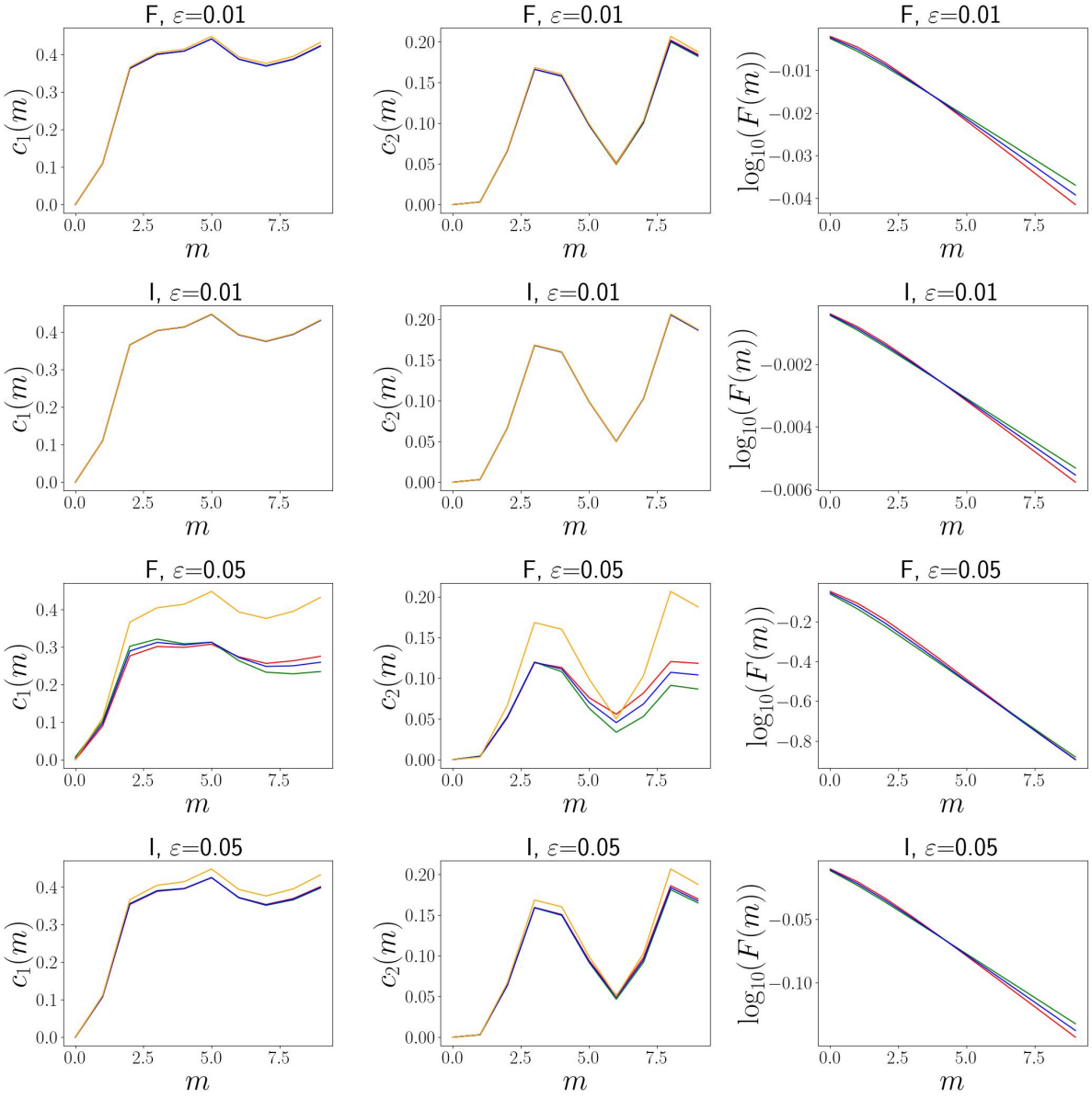


Figure 6.3: Evolution of several simulation metrics under a n = 10-qubits ferromagnetic Ising hamiltonian. 'F' indicates frequency noise and 'I' indicates intensity noise. The color pattern for the schemes : {(Red, '0'), (Blue, '01'), (Green, '1'), (Orange, 'Noiseless')}

6.7 Results

6.7.1 Exact emulations

We first emulate each hamiltonian, antiferromagnetic (+) then ferromagnetic (−), exactly on $n = 10$ qubits by tracking the full density matrix of the qubit register, in depth $m = 10$ and duration $t_f = 1$. Here are our observations:

- **Antiferromagnetic case** (+):
  - **Pauli correlations**: The observations here apply for both Nearest-Neighbor and Next-Nearest-Neighbor correlations. At low noise ($\varepsilon = 0.01$), despite the circuit fidelity already degraded, the noisy Pauli correlations are still in line with the noiseless regime, though they deviate more from there in frequency than in intensity noise. At a higher level of noise ($\varepsilon = 0.05$), the correlations are highly damped and degraded, yet again this is more visible in frequency than in intensity noise. In all cases, the '1' scheme preserves best the correlations, followed by '01' and finally '0'.
  - **Circuit fidelity**: The '1' scheme unequivocally preserves best the circuit fidelity in all the scenario there, always followed by '01' and '0' in decreasing order of performance. We notice that the fidelity drops much faster in frequency than in intensity noise, consistently with the damping in correlations being higher in frequency than in intensity noise.
- **Ferromagnetic case** (−):
  - **Pauli correlations**: The observations here apply for both Nearest-Neighbor and Next-Nearest-Neighbor correlations. Like previously, at low noise ($\varepsilon = 0.01$) the noisy Pauli correlations are still in line with the noiseless regime, though they deviate more from there in frequency than in intensity noise. Again, at a higher level of noise ($\varepsilon = 0.05$), the high damping of correlations are is more visible in frequency than in intensity noise. Generally, the '0' scheme preserves best the Pauli correlations, though the differences are lighter than in the antiferromagnetic case.
  - **Circuit fidelity**: Deviations between compensation schemes appear here too, but in weaker tendenciesthan in antiferromagnetic simulations. We observe that in the beginning the '0' scheme showcases a slightly higher fidelity, then the roles are reversed after about 4 circuit layers, and then the '1' scheme preserves best the circuit fidelity. Notice that this does not align with the observations on Pauli correlations.

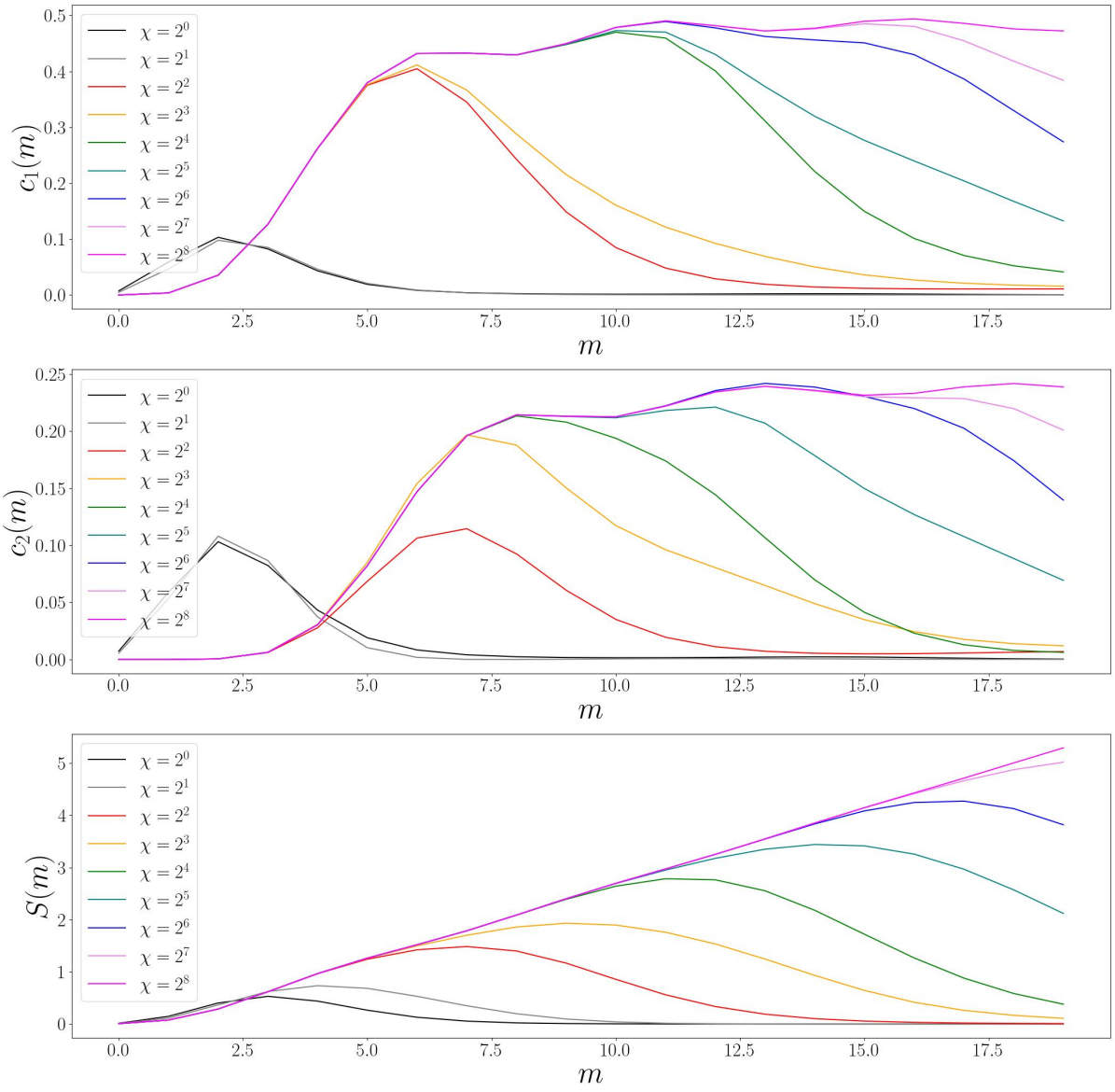


Figure 6.4: Convergence of noiseless standard metrics under a $n = 20$-qubits ferromagnetic Ising hamiltonian accross doubling bond dimensions

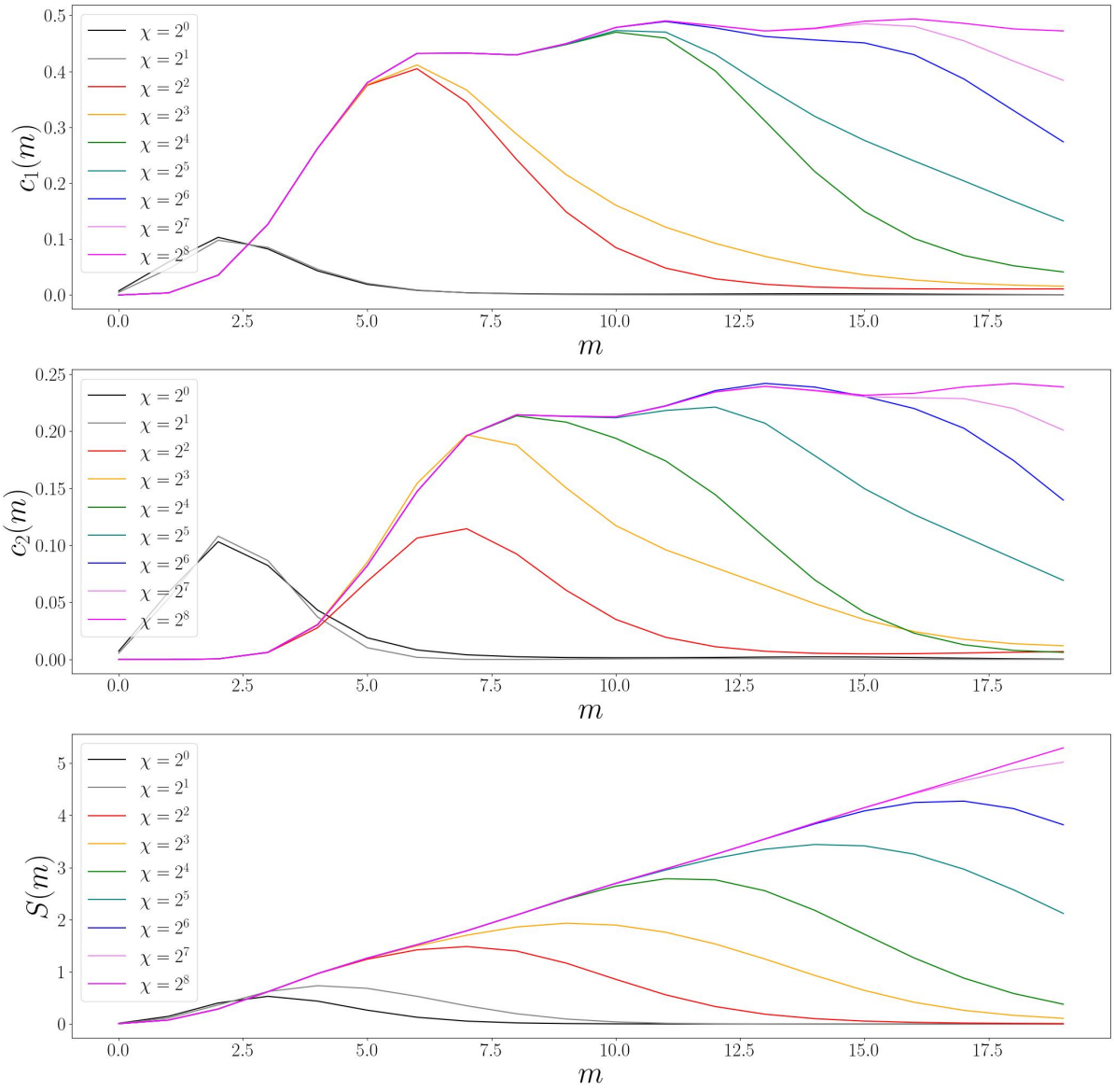


Figure 6.5: Convergence of noiseless standard metrics under a n = 20-qubits antiferromagnetic Ising hamiltonian accross doubling bond dimensions

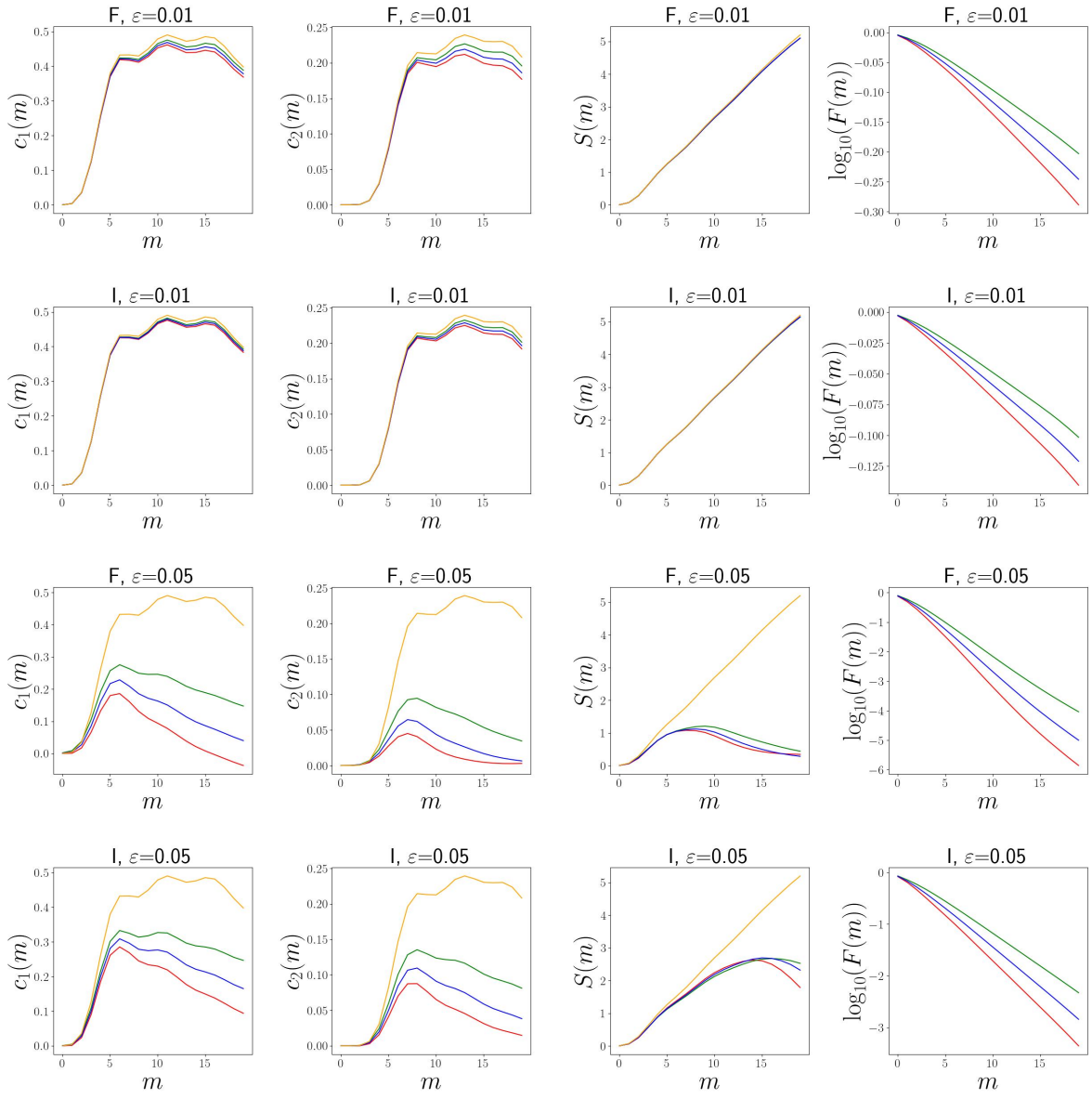


Figure 6.6: Noisy evolution of standard metrics under a n = 20-qubits antiferromagnetic Ising hamiltonian accross varying noise compensation schemes. 'F' indicates frequency noise and 'I' indicates intensity noise. The color pattern for the schemes : {(Red, '0'), (Blue, '01'), (Green, '1'), (Orange, 'Noiseless')}

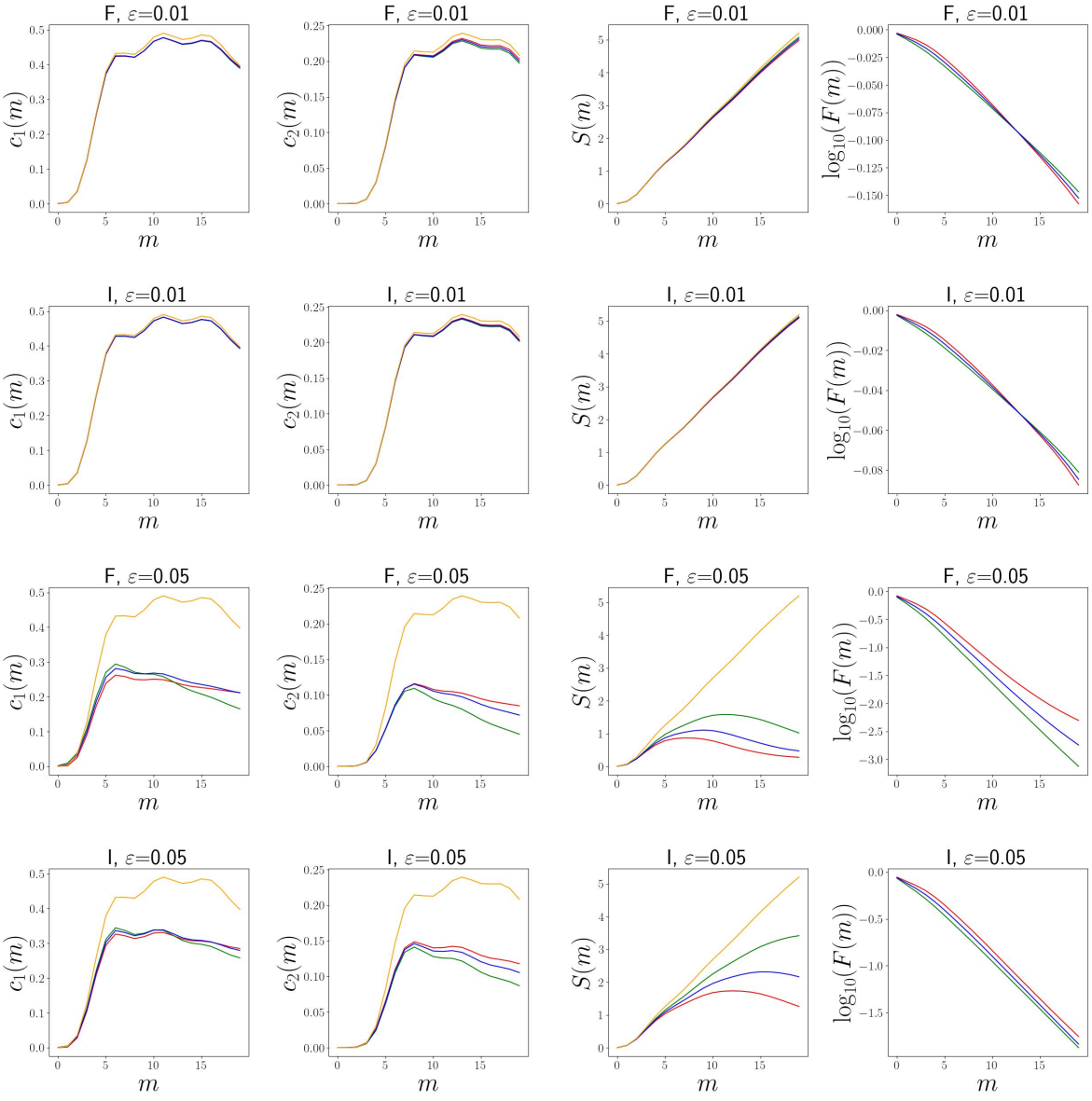


Figure 6.7: Noisy evolution of standard metrics under a n = 20-qubits ferromagnetic Ising hamiltonian accross varying noise compensation schemes. 'F' indicates frequency noise and 'I' indicates intensity noise. The color pattern for the schemes : {(Red, '0'), (Blue, '01'), (Green, '1'), (Orange, 'Noiseless')}

In order to identify the relevant bond dimension for the approximate emulations, we emulate the noiseless antiferromagnetic and ferromagnetic $n = 20$-qubits Ising hamiltonians in depth $m = 20$, duration $t_f = 1$, and with bond dimensions $\chi = 2^k$, $k \in [\![1, 8]\!]$; and collect the relevant metrics. We observe that the quantities start converging at $\chi = 128$, therefore we choose this as the bond dimension of the next emulations. Convergences plots of the relevant metrics are shown in figures 4 and 5.

To verify the simulation results on a higher number of qubits, we emulate the same circuits but using Matrix Product Density Operators instead of full density matrices. We emulate both hamiltonians with $n = 20$ qubits in depth $m = 20$, duration $t_f = 1$, and a bond dimension $\chi = 128$. On all these emulations, we add the comparison of operator space entanglement entropies.

- **Antiferromagnetic case** (+):
  - **Pauli correlations**: The same remarks as in the exact emulation case can be made here. We only notice that tendecies appear as early as $\varepsilon = 0.01$, and that the divergence at $\varepsilon = 0.05$ between the effect of frequency and intensity noise, in each scheme, is lighter than in the antiferromagnetic exact emulation case. The '1' scheme preserves best the Pauli correlations, followed by '01' and '0' in decreasing order of performance.
  - **Circuit fidelity**: Again, The '1' scheme sensibly has the highest circuit fidelity, followed by '01' and '0' in decreasing order of performance; consitently with the results in exact emulation.
  - **Operator Space Entanglement Entropy**: At noise level $\varepsilon = 0.01$, there is almost no difference in OE among the schemes. At $\varepsilon = 0.05$ in both types of noise, the OE of scheme '1' peaks higher and later than the other schemes, followed by '01' and '0' in this order.
- **Ferromagnetic case** (−):
  - **Pauli correlations**: In nearest-neighbour correlations, tendencies hardly appear even at $\varepsilon = 0.01$. At $\varepsilon = 0.05$ though the '0' scheme is very slightly above the others, followed by '01' and '1'; and not much difference can be observed between frequency and intensity noise. In nearest-neighbour correlations, the situation is different: the relative divergences between the schemes are much larger, and the correlations are clearly more damped in frequency than in intensity noise. Again, the '0' scheme showms the correlations closest to the noiseless regime, folloed by '01' and '1'.

- **Circuit fidelity**: At noise level $\varepsilon = 0.01$ (both types), the initial ranking of circuit fidelity until the thirteenth layer is '0'-'01'-'1', and is then reversed after that, consistently with the exact emulation scenario. At noise level $\varepsilon = 0.05$, the '0' scheme gives a better fidelity than the other schemes, which stands against the observations in exact emulation.
- **Operator Space Entanglement Entropy**: Compared to the previous run, the tendencies are much more visible here and the '1' scheme has the highest OE, followed by '01' and '0'.

6.7.2 Analysis

The base observation here is that for any metric, there is always a forced optical projection scheme that is a better choice than the natural decay. In the noiseless scenarii, there is no difference in correlations and operator entanglement between the anti-ferromagnetic and ferromagnetic hamiltonians. But when accounting for noise, each compensation schemes showcases divergent metrics between the two types of hamiltonians. In other words, the preservation of those metrics depends both on the noise compensation scheme and the nature of the hamiltonian under study. Indeed strikingly, despite Kraus coefficients being quite low (remember that they scale in ε^2), the nature of the noise impacts a lot the differences in standard metrics accross the schemes. We also note that in the ferromagnetic case, the noisy scenario diverge more from the noiseless one (i.e correlations are more damped and fidelities collapse faster) than in the antiferromagnetic case, and on the other hand the quantitative differences between the schemes are more marked in the antiferromagnetic emulation than in the ferromagnetic one.

In both hamiltonians, the circuit fidelity collapses after a few layers for all the different noise types, levels and compensation schemes. Let's start with a few general remarks about the evolution of two-body Pauli correlations in both exact and approximate emulations. At a moderatley low level of noise ($\varepsilon = 0.01$), they quite still hold, in both noises, and slightly better so in intensity than in frequency noise; despite the fidelity collapsing. At this level, there is no saliant difference between the compesation schemes. At a higher level of noise ($\varepsilon = 0.05$) they are much more damped and significantly deviate from the noiseless case, and tangible differences between the schemes emerge. Again, this effect is more visible in frequency noise than in intensity noise.

In the exact and approximate antiferromagnetic emulations, in all types and levels of noise considered, the ranking of the schemes in all metrics is '1-01-0': this confirms the intuitive view that the noise is best "compensated" by de-exciting the atom back

to $|1\rangle$ after staying in $|2\rangle$ for too long. However, in the ferromagnetic case, those observations are reversed in Pauli correlations and the order is rather '0-01-1'. The case of fidelity is subtler, as we notice that in the first half of the emulation, the fidelity ranking is '0-01-1' too, and is reversed after. As the reversal happens roughly at half the size of the array, we hypothesize that it is actually a boundary effect: this is confirmed by the high noise case ($\varepsilon = 0.05$) in the approximate evolution, where the reversal does not happen and the boundary effects arrive later due to the noise hampering the dynamics. We deduce from there that in the ferromagnetic case, it is generally better to de-excite the atom back to the $|0\rangle$ instead.

The operator space entanglement entropies start diverging when the level of noise becomes moderately high ($\varepsilon = 0.05$). The OE ranking is '1'-'01'-'0', more precisely the OE bump studied in [4] is always higher and appears later in the '1' scheme, followed by '01' and '0'. This suggests that the '1' scheme has the highest bond dimension requirements in a MPDO emulation (of the evolution under the Ising hamiltonians involving the present digital gate setup), is less efficiently classically simulable and deviates less from the quantum nature of the dynamics. This is first a surprising result because this scheme features only σ^z errors. But worse, in the ferromagnetic emulation a quite paradoxal situation happens: the '0' improves the fidelity but is less 'quantum'! It results that the qualitative symmetry in circuit fidelity and correlations matching between the ferromagnetic and antiferrroamgnetic hamiltonians, is independent of the operator space entanglement entropy and of the MPDO bond dimensions.

6.8 Conclusion

In this chapter, we have seen how a generic characterization of the control noise leads to completely computing its associated two-qubit noise channels, using second-order perturbation theory and standard stochastic calculus tools. From there, the different optical compensation schemes relevant for perturbations of the Rydberg-atomic setup are characterized by varying distributions of Kraus operators (and coefficients) in the derived channels. Plugging these channels in the emulation of an antiferromagnetic and a ferromagnetic Ising hamiltonians reveal interesting insights, notably that the degrees of preservation of each standard emulation metric are not entirely aligned on each other and can be negatively correlated on some figures of interest. Another observation is that the noise compensation schemes are symmetrically distributed on each metric around the statu quo reaction. For each orientation of the transverse Ising hamiltonian, we identify an optical scheme that preserves best the relevant simulation metrics, which is to revert the atoms back to a certain computational basis state. Interestingly, this choice seems independent of the operator entanglement growth differences induced by the choice of post-gate optical projection. An interesting question would be to apply

different optical schemes throughout the emulations of the hamiltonians, which can also be drawn from ensembles beyond the Ising family.

Bibliography

[1] Daraban, Ruben, Fabrizio Salas-Ramírez, and Johannes Schachenmayer. "Non-unitarity maximizing unraveling of open quantum dynamics." SciPost Physics 18, no. 2 (2025): 048. (Cited on pages 79 and 91.)

[2] Preisser, G., Wellnitz, D., Botzung, T. and Schachenmayer, J., 2023. Comparing bipartite entropy growth in open-system matrix-product simulation methods. Physical Review A, 108(1), p.012616. (Cited on pages 17, 79 and 91.)

[3] Noh, K., Jiang, L. and Fefferman, B., 2020. Efficient classical simulation of noisy random quantum circuits in one dimension. Quantum, 4, p.318. (Cited on pages 75 and 79.)

[4] Wellnitz, D., Preisser, G., Alba, V., Dubail, J. and Schachenmayer, J., 2022. Rise and fall, and slow rise again, of operator entanglement under dephasing. Physical Review Letters, 129(17), p.170401. (Cited on pages 79, 91 and 102.)

[5] Kim, Y., Eddins, A., Anand, S. et al. Evidence for the utility of quantum computing before fault tolerance. Nature 618, 500–505 (2023). https://doi.org/10.1038/s41586-023-06096-3 (Cited on page 80.)

[6] Patra, S., Jahromi, S.S., Singh, S. and Orús, R., 2024. Efficient tensor network simulation of IBM's largest quantum processors. Physical Review Research, 6(1), p.013326. (Cited on page 80.)

[7] Doob, J.L., 1942. The Brownian movement and stochastic equations. Annals of Mathematics, 43(2), pp.351-369. (Cited on page 87.)

[8] Wootters, W.K., 2001. Entanglement of formation and concurrence. Quantum Inf. Comput., 1(1), pp.27-44. (Cited on pages 19 and 91.)

[9] Jandura, S. and Pupillo, G., 2022. Time-optimal two-and three-qubit gates for Rydberg atoms. Quantum, 6, p.712. (Cited on pages xxiv, 26, 53, 54, 61, 79, 81 and 114.)

Chapter 7

A dynamic randomization approach to optimize Rydberg-blockade multi-qubit gates

Sommaire

7.1 Introduction

The theory of Optimal Control (OC) has been attracting a growing interest as a theoretical tool to fight many challenges met in the conception of quantum logic gates [1], including faulty controls, spontaneous photon emissions and leakage to energy levels outside of the computational basis, or finiteness of Rydberg blockade in the neutral-atomic context. The core idea is that since the effect of those phenomena is dependent on the external control over the hardware qubit (generally an electromagnetic field), then by tweaking the Rabi oscillation cycle between the desired states, one can maximize the gate fidelity *and* suppress or minimize as much as possible the effect of those error sources. The natural idea is often to discretize the time interval and treat the control as a finite vector of values $\boldsymbol{a}$ in over that time partition and optimize the fidelity and robustness metrics, summarized by a function F, with regard to that vector. Algorithms such as GRAPE provide numerical shortcuts to compute the gradient of the objective function, which is then fed into a gradient descent method converging to at least to a local minimum of that objective [2]. When looking for a global minimum, this approach can be completed by a heuristic algorithm such as simulated annealing, that would locate the neighboorhood of the global minimum. In uncertain and stochastic, more realistic contexts, this objective function can also be minimized using Artificial Evolution (AE) algorithms such as the Covariance-Matrix-Adaptation Evolution Strategy (CMA-ES).

The reference frequency ω of modern lasers tends to grow over time, allowing them to send shorter and shorter light pulses. Indeed, referring to the argument in chapter 2, when ω is sufficiently higher than the fluctuations of the source electromagnetic field then the Rabi cycle dynamics can be formulated in continuous time. This implies that the time discretization aforementioned becomes tighter and tighter, so discretizing the control results in a significant number of variables in the problem, raising its difficulty. This issue is solved by the cornerstone result of OC theory, the Pontryagin Minimality Principle (PMP) that reduces any deterministic OC problem to a boundary value problem (i.e a nonlinear ordinary differential equations with mixed initial and endtime conditions), which itself is numerically reduced down to a root-finding problem whose size is the dimension of the controlled trajectory, drammatically reducing the number of variables. The PMP is further detailed in section 7.2; and it admits a discrete-time formulation too. Once that associated root-finding problem is solved, the optimal control is straightforwardly reconstructed from its solution. However, the boundary value problem induced by the PMP involves highly nonlinear functions, suggesting that the associated numerical root-finding has an ill-behaved gradient and may fall into the computationally worst-case setting, requiring a number of function calls that is exponential in the trajectory dimension.

At a deeper level, the minimum of the objective function (with regard to the control) in the control problem can be considered instead as a function of the starting point and time of the trajectory, which is called the *value function* (VF). It completely characterizes the OC problem and satisfies a nonlinear PDE called the Hamilton-Jacobi-Bellman (HJB) PDE, though in a particular sense cal led the 'viscosity sense'. Similarly to the PMP, computing the VF bypasses the optimization with regard to the control; since the latter is generally implicitly computed as a byproduct of numerical approximations to the VF. Beyond this, computing the VF also bypasses the rootfinding step of the PMP, but at the extent of solving the HJB equation, whose computational complexity is exponential in the trajectory dimension. However, as seen in the next pages, there exist stochastic approximations to the VF computed through Monte Carlo methods.

In this chapter, following the aforementioned trend in the litterature, we analyse the derivation of Rydberg-atomic two-qubit gates pulses robust with regard to the aforementioned perturbations, through the lens of optimal control theory. We propose to solve the resulting OC problem through the framework of value functions, for two different statistical assumptions on the noise process:

- In section 7.2, we show that thanks to a connection between semilinear partial second-order differential equations and forward-backwad stochastic differential equations, optimal control problems benefit from the vanishing viscosity approximation of the value function. Our proposed main contribution is to reduce the complexity of initializing the optimal costate of OC problems, and in particular to make it independent of the dimension. The content of this section is mostly conjectural and requires further mathematical investigation. We propose to apply this to the minimization of exposition to spontaneous photon emission and of the impact of stochastic control noise (either frequency or intensity). Here, the noise is a stationary process with a given power spectral density.

- In section 7.3, we consider the control noise as a continuous-time markovian process, represented by a SDE. Our argument is vindicated by the recently derived SDE for pink noise. We propose to solve the resulting stochastic optimal control problem via the recent 'control randomization method'.

7.2 Dynamic randomization in deterministic robust control: A Monte-Carlo method for costate initialization

Disclaimer: this section constitutes open research directions and the results outlined there are not entirely analytically proven. In all the expressions in this chapter, functions are considered as smooth as the argument requires.

7.2.1 Generalities about optimal control

Let $\cdot$ refer to the canonical inner product. Let $d \in \mathbb{N}$ be the trajectory dimension, $x_0 \in \mathbb{R}^d$ be the trajectory starting point, $t_0 \geq 0$ be its starting time and $\tau > t_0$ its fixed end time, $\mathcal{A}$ be a compact metric space. A typical OC problem is mathematically written:

$$\text{minimize} \qquad \int_{t_0}^{\tau} f(q(t), a(t))\, dt + h(q(\tau)) \tag{7.1}$$

$$\text{subject to} \qquad q(t_0) = x_0,\ \forall t \in [t_0, \tau],\ a(t) \in \mathcal{A},\ \dot{q}(t) = \mu(q(t), a(t)) \tag{7.2}$$

Where:

- $\mu : O \times \mathcal{A} \to \mathbb{R}^d$ constraints the internal dynamics. It is assumed to be uniformly lipshitz-continuous in its first variable.
- $f : O \otimes \mathcal{A} \to \mathbb{R}$ is the instantaneous rate cost, also called a *lagrangian* in relation to classical mechanics. It is assumed to be uniformly lipshitz-continuous in its first variable.
- $h : \mathbb{R}^d \to \overline{\mathbb{R}}$ is the endpoint cost, that can be tailored to signal an endpoint constraint on the trajectory.

The natural method to solve the OC problem is to discretize the time interval into a finite grid $\{t_0, \cdots, t_n := \tau\}$ and the integrals as sums, resulting in a total cost depending on $\boldsymbol{a} = (a(t_i)_{i \in [\![0,n]\!]})$. In quantum control problems where the endpoint cost h is the fidelity with regard to a certain target state, the total cost is efficiently differentiable in $\boldsymbol{a}$ because the derivatives of $x(t_i)$ in $\boldsymbol{a}$ are straightforward to compute from the unitary update rules, fitting well into the linear cost [2].

Alternatively, *Pontryagin's Maximality Principle* (PMP) emphasizes the constrained structure of the continuous OC problem and solves it through a Lagrange multiplier-like dynamical dual variable, called the *costate*. Define the *control hamiltonian*:

$$H(p, x) = \min_{a \in \mathcal{A}} \mu(x, a) \cdot p + f(x, a)$$

Where the minimum above is assumed to be reached at a unique $\hat{a}(x,p)$. The PMP [10] states that if (a,q) is an optimal control/trajectory pair then there exists an optimal *costate* variable p such that:

$$\begin{aligned}
a(t) &= \hat{a}(q(t),p(t)) \\
\dot{q}(t) &= \nabla_p H(p(t),q(t)) = \mu(q(t),a(t)) \\
\dot{p}(t) &= -\nabla_q H(p(t),q(t)) = \partial_x \mu(q(t),a(t)) \cdot p(t) + f(q(t),a(t)) \\
p(\tau) &= \nabla h(q(\tau))
\end{aligned}$$

The last equation is called the *transversality condition* and constitutes the boundary condition of the hamiltonian system q,p, along with the starting point $q(0) = x_0$. It is replaced by $q(\tau) = x_1$ when h signals a constraining trajectory target $x_1 \in \mathbb{R}^d$. The OC problem is solved by solving this boundary value problem, i.e by computing the initial costate $p(0)$ such that the boundary conditions are satisfied. Since the dimension of the trajectory variable is usually significantly lower than the discretization scale, the PMP dramatically reduces the complexity of the OC problem, from determining a whole function to computing the initial costate from where the OC is automatically constructed.

The PMP generally gives only *necessary* optimality conditions, and is not guaranteed to give *sufficient* ones. However, if $\mathcal{A}$ is a convex set and h,l,f are convex and differentiable functions, then the PMP provides *sufficient* conditions too, and the minimizer is global [9]. These conditions are generally satisfied in quantum optimal control. When the trajectory is constrained in a manifold M, the associated costate lies in the tangent space TM of the manifold at the trajectory point [8]. In other words:

$$(\forall t \in [0,\tau],\, q(t) \in M) \implies (\forall t \in [0,\tau],\, p(t) \in TM(q(t)))$$

This further reduces the search landscape of the initial costate. If $M = g^{-1}(\{0\}) = \{x \in \mathbb{R}^d \mid g(x) = 0\}$ for some $g : \mathbb{R}^d \to \mathbb{R}$ then $\forall a \in M$, $TM(a) = \{\nabla g(a)\}^\perp = \{x \in \mathbb{R}^d \mid \langle \nabla g(a), x\rangle = 0\}$. As an outlook, the PMP can be generalized to multi-objective OC problems [21, 22].

7.2.2 Hardness of shooting method

Since $q(0) = x_0$ is fixed, $(q(\tau),p(\tau))$ is entirely determined by $p(0)$. Since the remaining boundary conditions is an element of $\mathbb{R}^d$.

Therefore, this provides a function $F : \mathbb{R}^d \to \mathbb{R}^d$ such that the endpoint condition of the boundary value problem induced by the PMP is $F(p(0)) = 0$. Since $(q(\tau),p(\tau))$ is efficiently differentiated with regard to $p(0)$ in a process called 'adjoint differentiation', so is F and it is suitable for a derivative-based algorithm such as Newton's method

or gradient descent. However, the high nonlinearity of the system makes F 'chaotic', i.e highly sensitive to the original costate and the jacobian ∂F possibly ill-behaved, indicating that an efficient initialization is required for any derivative-based root-finding method to converge.

A good such preconditioner is the *characteristic bisection method* (CBM) [12], which generalizes the one-dimensional dichotomic algorithm to identify the root of F. It works by first constructing a polyhedron of 2^d vertices such that, every element of $\{-1, 1\}^d$ is the sign pattern of the image by F of a unique vertex therein, which ensures the presence of a root thereinside. Then, the polygon is progressively refined by identifying at each step a polyhedron's diagonal whose midpoint respects the global sign patterns, and cutting the polygon at that point. The process continues until the volume of the polyhedron shrinks below an input threshold, at which point the latest cut point is considered a good approximate root. The CBM showcases a linear convergence rate, as it gains one bit of precision at each algorithmic step (this can be seen in one dimension as each cut in half of the interval decides indicates the current bit of the root).

In one dimension, the bisection method offers the best worst case complexity with regard to the interval size [3], so it is reasonably be expected to hold in higher dimensions too. Among the existing root-finding algorithms, the bisection method has the weakest requirements, as it merely requires appropriate sign patterns in the values of a continuous function, and nothing else about its regularity. Indeed, a few years later, the authors of the CBM have generalized the intermediate value theorem to a multivariate setting [13], as a deeper theoretical basis of the CBM. These two facts strongly suggest that the CBM is the 'minimal' multivariate root-finding algorithm, offering the best worst-case complexity and weakest requirements in several dimensions, while being the slowest in the average case.

The CBM requires to evaluate the of signs of images of all the vertices of an hypercube in $\mathbb{R}^d$, and therefore calls F at least 2^d times. Due to our previous remark, this reveals the profound combinatorial nature of root-finding, and heavily suggests that the intrisic worst case query complexity of root-finding is in $\Omega(2^d)$ calls to the function to nullify, making it prone to the *curse of dimensionality*. A `Fortran77` implementation of the CBM is available online [1], with a user guide in [14].

[1] https://people.sc.fsu.edu/~jburkardt///f77_src/toms666/toms666.html

7.2.3 Value function and vanishing viscosity

The OC problem is encapsulated by the *value function*

$$\begin{aligned} v_0(x_0,t_0) = \quad & \text{minimize} \quad && \int_{t_0}^{\tau} f(q(t),a(t))\,dt + h(q(\tau)) \\ & \text{subject to} \quad && q(t_0)=x_0,\ \forall t\in[t_0,\tau],\ a(t)\in\mathcal{A},\ \dot{q}(t)=\mu(q(t),a(t)) \end{aligned}$$

A hallmark analytic result of OC theory is that v_0 is the unique function such that $v_0(\cdot,\tau)=h$ and that is a **viscosity solution** (see the definition in [11]) of the Hamilton-Jacobi-Bellman equation:

$$\partial_t v_0(x,t) + H(\nabla v_0(x,t),x) = 0$$

It is related to the PMP in the sense that, in previous notations, if v_0 is differentiable at $x_0 = q(t_0)$ then $p(t_0) = \nabla v_0(x_0,t_0)$. Therefore, approximating the value function and its gradient (at differentiability points) leads to initializing the PMP costate and effectively solving the OC problem. Let's now turn to such an approximation.

For $\varepsilon > 0$, let v_ε be the unique function such that $v_\varepsilon(\cdot,\tau)=h$ and that is a classical solution to

$$\partial_t v_\varepsilon + H(\nabla v_\varepsilon,\cdot) = \varepsilon\Delta v_\varepsilon$$

This always uniquely exists through parabolic theory. The principle of **vanishing viscosity** states that generically, v_ε converges uniformly to v_0 on any compact subset K as ε vanishes. More precisely, there exists a positive constant C_K (which actually is the Lipshitz constant of μ over K, uniformly in a) such that [11]:

$$\sup_{(x,t)\in K} \|v_0(x,t) - v_\varepsilon(x,t)\| \leq C_K \varepsilon^{1/2}$$

One is thus looking at similar convergence rates for the gradients at points of differentiability. Indeed, in [6, 7], under more detailed but still general conditions on H that we'll consider true there, integral estimates on the convergence rates of gradients are given in the form of:

$$\max_{t\in[0,\tau]} \left(\int_{x\in Q} \|\nabla v_\varepsilon(x,t) - \nabla v_0(x,t)\|^2 dx \right)^{1/2} < C_\beta \varepsilon^\beta$$

Where $C_\beta > 0$ and $\beta \in [1/4, 1]$ are constants **independent of** d). This integral estimate cannot by itself exactly bound the gradient convergence at a fixed point. However, classic measure-theoretic results such as Markov's inequality imply that when $x \in Q$ is randomly sampled, the integrand is bounded at that point with high probability. Therefore, we propose to replace Q from $[-1,1]^d$ to $x_0 + r[-1,1]^d$ (where r is a fixed small radius), i.e to solve the PDE actually in a neighboorhood of x_0, so that this estimate results in an inequality holding with high probability (through Markov's inequality):

$$\|\nabla v_\varepsilon(x_0',t_0) - \nabla v_0(x_0',t_0)\| < C_\beta \varepsilon^\beta$$

7.2.4 Nonlinear Feynman-Kac formula and numerical Monte Carlo method for semi-linear PDE's

A PDE is called *semi-linear* when the terms of highest differentiation order appears only linearly, while lower-order terms are free to appear nonlinearly. Much like its linear version, the nonlinear Feynman-kac formula fundamentally ties second-order semi-linear partial differential equations to forward-backward stochastic differential equations (FBSDE).

Let $b : \mathbb{R}^d \to \mathbb{R}^d, \sigma : \mathbb{R}^d \to \mathbb{R}^{(d,d)}$, $g : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times [0,\tau] \to \mathbb{R}, h : \mathbb{R}^d \to \mathbb{R}$, $x_0 \in \mathbb{R}^d, t_0 \in \mathbb{R}^+, \tau \in [t_0, +\infty]$. The nonlinear Feynman-Kac theorem states that:

$$\begin{cases} X_{t_1} &= x_0 + \int_{t_0}^{t_1} b(X_t,t)dt + \int_{t_0}^{t_1} \sigma(X_t,t)dW_t \\ Y_{t_1} &= h(X_\tau) - \int_{t_1}^{\tau} g(X_t,Y_t,Z_t,t)dt - \int_{t_1}^{\tau} Z_t dW_t \end{cases} \iff$$

$$\begin{cases} \exists w : \mathbb{R}^d \times [t_0,\tau] \to \mathbb{R} \\ w(\cdot,\tau) = h \\ \partial_t w(x,t) + b(x,t)\cdot\nabla w(x,t) + \mathrm{Tr}\big(\sigma(x,t)^T\sigma(x,t)\cdot\nabla^2 w(x,t)\big) + g(x,w(x,t),\sigma(x,t)\nabla w(x,t),t) = 0 \\ Y_{t_1} = w(X_{t_1},t_1) \\ Z_{t_1} = \sigma(X_{t_1},t_1)\nabla w(X_{t_1},t_1) \end{cases}$$

In particular, the latter property implies at the starting times the deterministic equalities

$$Y_{t_0} = w(x_0,t_0),\ Z_{t_0} = \sigma(x_0)\nabla w(x_0,t_0)$$

For a concise outline of FBSDEs in the context of semi-linear PDEs, we refer to section 2 of [16] and references therein. For deeper mathematical details about the relation between the processes Y and Z, we refer to [24].

We do not elaborate here much on the numerical solutions to the FBSDE system, yet we just note that they require to compute conditional expectations to erase the future randomness from the current-time processes. This is in general a quite tricky calculation, carried by Monte Carlo methods through either functional regressions, empirical projections (as in [20]) or Malliavin calculus, as analytically explained in [17].

Numerical solutions (X^π, Y^π, Z^π) to the equation system on the left are typically indexed by a time interval partition π of maximal discretization scale $|\pi|$. Under Lipshitz-continuity assumptions of (g,h) (which are satisfied in the sought applications anyway), the typical performance of numerical schemes scales as:

$$\max_{t\in[\![t_0,\tau]\!]} \mathbb{E}(|Y_t^\pi - Y_t|^2) + \int_{t_0}^{\tau} \mathbb{E}(|Z_t - Z_t^\pi|^2)dt = O(|\pi|)$$

A relative caveat of this inequality is to bound only a global integral squared averages of differences $\int_0^\tau \mathbb{E}(|Z_t^\pi - Z_t|^2)dt$ instead of local pointwise terms $|Z_{t_i}^\pi - Z_{t_i}|$ which is far more interesting in the cases requiring a tight approximation of the adjoint gradient process. As explained in [24], this is likely because conventional schemes treat the numerical adjoint gradient process Z^π rather as an integral of other terms, rather than a process soundly defined at a single time step.

The divergence between aforementioned global and local bounds tends to appear in rather time-discontinuous processes, 'jumping' high enough to concentrate values at sufficiently many time points to break any uniform bound, and thinly so enough to keep the integral bounds low. It is proved in [24] that Z is generally Hölder-continuous, so we expect this continuity property to suppress those unwanted concentrations and control further the divergence between integral and pointwise estimates, so that a scheme minimizing the above quantity shows a similar performance to bound $|Z_{t_0}^\pi - Z_{t_0}|$ with high probability, a property that we'll assume here. We notice that in [24], under stringe conditions on the BSDE generator g and with a more refined numerical scheme based on Malliavin calculus, a much more favorable convergence rate is proved:

$$\max_{t\in[\![t_0,\tau]\!]} \mathbb{E}(|Y_t^\pi - Y_t|^2 + |Z_t^\pi - Z_t|^2) = O(|\pi|)$$

The key ingredient of this bound is to properly time-localize the adjoint gradient process Z by relating it to the value process Y through Malliavin calculus (in general conditions on h, g), although for now this trick gives results only under stringe assumptions on g that are not even satisfied in our setup. But we mention this bound though, because it is a promising research direction in our opinion though.

Due to this reasoning and the previous equalities and scalings, the numerical schemes featured in [16, 24] approximate $(w(x_0,t_0), \sigma(x_0)\nabla w(x_0,t_0))$ within an error of order $O(|\pi|^2)$. Standard analysis of Monte Carlo methods prove that to approximate these quantities within overall error ε, $O((\varepsilon|\pi|)^{-1/2})$ random trajectory samples are required, and roughly that scale of arithmetic operations are needed (since the numer of random samples dominates the computational complexity calculation), and these numbers are independent of the dimension d.

7.2.5 Our proposal and key takeaway

Now, let's join the two observations about the vanishing viscosity approximation of the value function and its gradient, and the stochastic numerical scheme for semi-linear PDE's. The next idea is to work with a constant diffusion matrix and a zero drift term. Let $\alpha > 0$ be a small positive constant. For $x \in \mathbb{R}^d, y \in \mathbb{R}, z \in \mathbb{R}^d, t \in [t_0, \tau]$, let $\sigma(x,t) = \alpha^{1/2} I$, $b(x,t) = \mathbf{0}$, $g(x,y,z,t) = H(x, \alpha^{-1/2} z) = \alpha^{-1/2} H(x,z)$. Then

$w = v_\alpha$ where w is the solution to the PDE in the right column of the NLFK theorem. Consequently, the previously outlined stochastic numerical scheme approximates $(v_\alpha(x_0,t_0), \alpha^{1/2}\nabla v_\alpha(x_0,t_0)$ at either a desired x_0 or at a point close by. Considering the arguments for the vanishing viscosity gradient convergence, a careful choice of the constant α leads to obtain a ε-close approximation to the initial costate $\nabla v_0(x_0,t_0)$ within $O((\varepsilon|\pi|)^{-1/(2\beta)})$ random trajectory samples and roughly that scale of arithmetical operations. Importantly, this complexity bound is independent of d, which contrasts with the root-finding worst case arguably met in optimal control.

7.2.6 Use cases in quantum control

7.2.6.1 Sensitivity to photon spontaneous emission / Minimizing the time spent in the Rydberg state

Consider here the synthesis of a multi-controlled phase gate for neutral atoms, with notations borrowed from [4]. Let $\Pi := \mathrm{Diag}((1,\cdots,1)) \otimes |1\rangle\langle 1|$ be the projector into the Rydberg manifold, and:

$$\sigma^+ := \mathrm{Diag}((1,\cdots,\sqrt{m})) \otimes |1\rangle\langle 0|$$
$$\sigma^- := \mathrm{Diag}((1,\cdots,\sqrt{m})) \otimes |0\rangle\langle 1|$$

The fidelity loss due to spontaneous photon emission from the Rydberg state to the computational subspace is approximately ΓT_R where Γ is the decay rate and T_R is the time spent in the Rydberg states [4], equal to the integral cost term in the equation below. Thus, in the infinite Rydberg blockade regime and assuming a noiseless drive, the pulse $\Omega(t)$ synthesing a multicontrolled phase gate robust against Rydberg state decay is found by solving the optimal control problem:

$$\text{minimize} \int_0^\tau \langle\psi(t)|\Pi|\psi(t)\rangle\, dt + A(1 - \langle\phi|\psi(\tau)\rangle|^2)$$
$$\text{subject to } \imath\,|\dot{\psi}(t)\rangle = (\Omega(t)\sigma^+ + \Omega(t)^*\sigma^-)\,|\psi(t)\rangle \quad |\Omega(t)| \le \Omega_{\max} \quad |\psi(0)\rangle = \mathrm{Diag}(1,\cdots,1)\otimes|0\rangle$$

Where the (very large) penalizing positive constant A enforces the constraint of equality with the target $|\phi\rangle := (e^{\imath\xi_q})_{q\in[\![1,m]\!]} \otimes |0\rangle$. The PMP evolution equation with costate χ is then:

$$\imath\,|\dot{\chi}(t)\rangle = (\Omega(t)\sigma^+ + \Omega(t)^*\sigma^-)\,|\chi_b(t)\rangle + \Pi\,|\psi_b(t)\rangle$$
$$\imath\,|\dot{\psi}(t)\rangle = (\Omega(t)\sigma^+ + \Omega(t)^*\sigma^-)\,|\psi_b(t)\rangle$$
$$\Omega(t) = \imath(\langle\chi(t)|\sigma^+|\psi(t)\rangle - \langle\chi(t)|\sigma^+|\psi(t)\rangle^*)$$

7.2.6.2 Sensitivity to colored control noise

Consider the general problem of driving a quantum system from ψ_{init} to ψ_{targ}, through a hamiltonian $\hat{H}$ deterministically controlled by a time-dependent signal $a(t)$ (lying in

a compact set $\mathcal{A}$, generally equal to the unit complex disk) inducing a small random perturbation $\varepsilon s(t)\hat{G}(a(t))$ on the system. The quality of the evolution is quantified by the average fidelity, yielding the control problem:

$$\begin{aligned}
&\text{minimize } \mathbb{E}\left(1 - |\langle\psi(\tau;\varepsilon)|\psi_{\text{targ}}\rangle|^2\right)\\
&\text{subject to } |\psi(0;\varepsilon)\rangle = |\psi^{\text{init}}\rangle\\
&\qquad a(t) \in \mathcal{A}\\
&\qquad \imath\,|\dot{\psi}(t;\varepsilon)\rangle = (\hat{H}(a(t)) + \varepsilon s(t)\hat{G}(a(t)))\,|\psi(t;\varepsilon)\rangle
\end{aligned}$$

Since ε is an experimental parameter prone to fluctuations in real setups, this problem is rather solved in perturbation theory by Taylor-expanding the trajectory in second order with regard to ε:

$$|\psi(t;\varepsilon)\rangle = |\psi^0(t)\rangle + \varepsilon\,|\psi^1(t)\rangle + \varepsilon^2\,|\psi^2(t)\rangle + O(\varepsilon^3)$$

Consequently:

$$|\langle\psi^0(\tau)|\psi(\tau;\varepsilon)\rangle|^2 = 1 + \varepsilon\Re(\langle\psi^0(\tau)|\psi^1(\tau)\rangle) + \varepsilon^2(\Re(\langle\psi^0(\tau)|\psi^2(\tau)\rangle) + |\langle\psi^0(\tau)|\psi^1(\tau)\rangle|^2) + O(\varepsilon^3)$$

Plugging the Taylor expansion in the noisy dynamics yield:

$$\imath\begin{pmatrix}|\dot{\psi}^0(t)\rangle\\|\dot{\psi}^1(t)\rangle\\|\dot{\psi}^2(t)\rangle\end{pmatrix} = \begin{pmatrix}\hat{H}(a(t)) & 0 & 0\\ s(t)\hat{G}(a(t)) & \hat{H}(a(t)) & 0\\ 0 & s(t)\hat{G}(a(t)) & \hat{H}(a(t))\end{pmatrix}\begin{pmatrix}|\psi^0(t)\rangle\\|\psi^1(t)\rangle\\|\psi^2(t)\rangle\end{pmatrix}$$

Differentiating the inner products:

$$\begin{aligned}
\frac{d}{dt}\langle\psi^0(t)|\psi^1(t)\rangle &= \langle\dot{\psi}^0(t)|\psi^1(t)\rangle + \langle\psi^0(t)|\dot{\psi}^1(t)\rangle\\
&= \imath\left(\langle\psi^0(t)|\hat{H}(a(t))|\psi^1(t)\rangle - \langle\psi^0(t)|\hat{H}(a(t))|\psi^1(t)\rangle - s(t)\,\langle\psi^0(t)|\hat{G}(a(t))|\psi^0(t)\rangle\right)\\
&= -\imath s(t)\,\langle\psi^0(t)|\hat{G}(a(t))|\psi^0(t)\rangle
\end{aligned}$$

Adding that $|\psi^1(0)\rangle = 0$, the hermitianity of $\hat{G}$ implies that $\imath\langle\psi^0(\tau)|\psi^1(\tau)\rangle \in \mathbb{R}$ so $\Re(\langle\psi^0(t)|\psi^1(t)\rangle) = 0$: the first-order ε-dependence of the endtime fidelity vanishes. Let's look at the second order term:

$$\begin{aligned}
\frac{d}{dt}\langle\psi^0(t)|\psi^2(t)\rangle &= \langle\dot{\psi}^0(t)|\psi^2(t)\rangle + \langle\psi^0(t)|\dot{\psi}^2(t)\rangle\\
&= \imath\left(\langle\psi^0(t)|\hat{H}(a(t))|\psi^2(t)\rangle\,dt - \langle\psi^0(t)|\hat{H}(a(t))|\psi^2(t)\rangle\,dt - \langle\psi^0(t)|\hat{G}(a(t))|\psi^1(t)\rangle\,s(t)\right)\\
&= -\imath\,\langle\psi^0(t)|\hat{G}(a(t))|\psi^1(t)\rangle\,s(t) = \imath\,\langle\psi^1(t)|\hat{H}(a(t))|\psi^1(t)\rangle - \langle\dot{\psi}^1(t)|\psi^1(t)\rangle\\
\frac{d}{dt}\Re(\langle\psi^0(t)|\psi^2(t)\rangle) &= -\Re(\langle\dot{\psi}^1(t)|\psi^1(t)\rangle) = -\frac{1}{2}(\langle\dot{\psi}^1(t)|\psi^1(t)\rangle + \langle\psi^1(t)|\dot{\psi}^1(t)\rangle\\
&= -\frac{1}{2}\frac{d}{dt}\langle\psi^1(t)|\psi^1(t)\rangle
\end{aligned}$$

Therefore, since $|\psi^1(0)\rangle = |\psi^2(0)\rangle = 0$:

$$|\langle\psi^0(\tau)|\psi(\tau;\varepsilon)\rangle|^2 = 1 + \varepsilon^2(|\langle\psi^0(\tau)|\psi^1(\tau)\rangle|^2 - \langle\psi^1(\tau)|\psi^1(\tau)\rangle) + O(\varepsilon^3)$$

In a robust control perspective, we seek to ensure that the noiseless part falls into the desired target through a penalized fidelity term weighted by a large positive constant A, while minimizing the first-order trajectory perturbation. So overall the probabilistic OC problem is:

$$\begin{aligned}
&\text{minimize } \mathbb{E}\left(A(1 - |\langle\psi^0(\tau)|\psi_{\text{targ}}\rangle|^2) + |\langle\psi^0(\tau)|\psi^1(\tau)\rangle|^2 - \langle\psi^1(\tau)|\psi^1(\tau)\rangle\right)\\
&\text{subject to } (|\psi^0(0)\rangle, |\psi^1(0)\rangle) = (|\psi^{\text{init}}\rangle, 0)\\
&\qquad a(t) \in \mathcal{A}\\
&\qquad \imath\begin{pmatrix}|\dot{\psi}^0(t)\rangle\\ |\dot{\psi}^1(t)\rangle\end{pmatrix} = \begin{pmatrix}\hat{H}(a(t)) & 0\\ s(t)\hat{G}(a(t)) & \hat{H}(a(t))\end{pmatrix}\begin{pmatrix}|\psi^0(t)\rangle\\ |\psi^1(t)\rangle\end{pmatrix}
\end{aligned}$$

Let's assume that s is a stationary stochastic process (i.e the covariance between two instances of the process depends only on the time lag between them), so consequently it admits a harmonic decomposition [15]:

$$s(t) = \sum_{k=1}^{N} a_{2k-1}\cos(\omega_k t) + a_{2k}\sin(\omega_k t)$$

Where $N \in \mathbb{N}$, $\omega \in \mathbb{R}^N$ and its distribution reflects the noise frequency spectrum, and a is a sequence of $2N$ centered, reduced and uncorelated random variables (i.e $\forall i, j \in [\![1, 2N]\!]$, $\mathbb{E}(a_i) = 0$, $\mathbb{E}(a_i^2) = 1$, $\mathbb{E}(a_i a_j) = 0$ when $i \neq j$). We refer to the supplementary material for more details about this decomposition. To account for the effect of each frequency on the noisy trajectory, let's introduce the solutions to the following equations:

$$\begin{aligned}
|\dot{\psi}^1_{2k-1}(t)\rangle &= \cos(\omega_k t)\hat{G}(a(t))\,|\psi^0(t)\rangle + \hat{H}(a(t))\,|\psi^1_{2k-1}(t)\rangle\\
|\dot{\psi}^1_{2k}(t)\rangle &= \sin(\omega_k t)\hat{G}(a(t))\,|\psi^0(t)\rangle + \hat{H}(a(t))\,|\psi^1_{2k}(t)\rangle
\end{aligned}$$

Along with the initial conditions $|\psi^1_j(0)\rangle = 0$, by linearity of differentiation and the statistical assumptions, one has:

$$\begin{aligned}
|\psi^1\rangle &= \sum_{j=1}^{2N} a_j\,|\psi^1_j\rangle\\
\mathbb{E}(\langle\psi^1(\tau)|\psi^1(\tau)\rangle) &= \sum_{j=1}^{2N}\langle\psi^1_j(\tau)|\psi^1_j(\tau)\rangle\\
\mathbb{E}(|\langle\psi^0(\tau)|\psi^1(\tau)\rangle|^2) &= \sum_{j=1}^{2N}|\langle\psi^0(\tau)|\psi^1_j(\tau)\rangle|^2
\end{aligned}$$

Therefore, the OC problem can be reformulated as to account for robustness against each frequency of the noise spectrum:

$$\begin{aligned}
&\text{minimize } A(1 - |\langle\psi^0(\tau)|\psi_{\text{targ}}\rangle|^2) + \sum_{k=1}^{2N+1} |\langle\psi^0(\tau)|\psi_k^1(\tau)\rangle|^2 - \langle\psi_k^1(\tau)|\psi_k^1(\tau)\rangle \\
&\text{subject to } (|\psi^0(0)\rangle, \cdots, |\psi_i^1(0)\rangle), \cdots) = (|\psi^{\text{init}}\rangle, \cdots, 0, \cdots) \\
&\qquad a(t) \in \mathcal{A} \\
&\qquad \imath|\dot{\psi}^0(t)\rangle = \hat{H}(a(t))|\psi^0(t)\rangle, \ \forall k \in [\![1, N]\!] : \\
&\qquad \imath|\dot{\psi}_{2k-1}^1(t)\rangle = \hat{H}(a(t))|\psi_{2k-1}^1(t)\rangle + \cos(\omega_k t)\hat{G}(a(t))|\psi^0(t)\rangle \\
&\qquad \imath|\dot{\psi}_{2k}^1(t)\rangle = \hat{H}(a(t))|\psi_{2k}^1(t)\rangle + \sin(\omega_k t)\hat{G}(a(t))|\psi^0(t)\rangle
\end{aligned}$$

Accounting for many frequencies ($N \gg 1$) leads to a highly-dimensional problem, vindicating our search for a preconditioner independent of the trajectory's dimensionality.

7.2.7 Discussion

Optimization problems featuring implicit constrained dynamics are usually solved by discretizing the steering variable, and minimizing the overall cost along that vector. In the quantum optical context, the simplicity of these dynamics and costs allow efficient numerical shortcuts to compute the cost's gradient as in the GRAPE algorithm. However, as the time discretization step decreases down to an approximate continuous formulation, the number of variables to optimize over through this approach increases, losing algorithmic efficiency in the process. The elegant mathematical theory of *optimal control* solves this issue through a powerful tool to sensibly reduce the number of variables, the Pontryagin Minimum Principle. It essentially converts the optimal control problem to a boundary value problem, numerically discretized into a root-finding problem. However, the high nonlinearity of the PMP's boundary value problem (even if the original control is linear) likely implies that the shooting methods to solve it will necessarily scale poorly in the trajectory dimensionality.

An alternative path is to leverage the powerful *value function* that characterizes the OC problem, along with its associated nonlinear Hamilton-Jacobi-Bellman equation, and whose gradient is the root sought above. Out of various geometric arguments, one can reasonably conjecture that both the VF and its gradient can be approximated by a related function (and its gradient) satisfying a semi-linear PDE. On the other hand, the nonlinear Feynman-Kac theorem states that solutions of second-order semi-linear PDE can be pointwise expressed with systems of forward-backward stochastic differential equations, which are themselves computed using Monte Carlo methods. Another reasonable conjecture here is that the convergence bounds also applies to the gradient of the sought function, featured in the numerical schemes too. By joining the dots,

one obtains a Monte Carlo method to initialize the costate variable of the PMP, thereby solving the OC problem, that avoids the curse of dimensionality while sitting on reasonable conjectures, vindicated by certain bounds in related problems.

This dimension-independent feature is especially crucial when designing quantum gates that are robust against many independent frequency rays of a control noise source, the decomposition there stemming from the stationarity assumption of the latter.

7.3 Control randomization: application in stochastic problems

7.3.1 Reminders about stochastic control

Let $\mathcal{A}$ be a compact metric space, $\mu : \mathbb{R}^d \times \mathcal{A} \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \times \mathcal{A} \to \mathbb{R}^{d,d}$ be uniformly Lipshitz-continuous and bounded functions. Let W be a Wiener process in $\mathbb{R}^d$, $\tau > 0$ be the fixed finite trajectory endtime. A stochastic optimal control problem consists of computing the minimizing steering a_t^* in the definition of the *value function*:

$$\begin{aligned} v(x_0, t_0) := \quad & \text{minimize} \quad && \mathbb{E}\left(\int_{t_0}^{\tau} f(X_t, a_t) dt + h(X_\tau) \right) \\ & \text{subject to} \quad && X_{t_0} = x_0,\ dX_t = \mu(X_t, a_t)dt + \sigma(X_t, a_t)dW_t,\ a_t \in \mathcal{A} \end{aligned}$$

A hallmark result of stochastic control theory is that v is the unique function such that $v(\cdot, \tau) = h$ and that is a *viscosity solution* (we refer to [11] for the formal definition of the concept) over $\mathbb{R}^d \times [0, \tau]$ to the *Hamilton-Jacobi-Bellman* (HJB) partial differential equation (PDE) [11]:

$$\partial_t v + \min_{a \in \mathcal{A}} \mu(\cdot, a) \cdot \nabla v + \mathrm{Tr}\big(\sigma^T \sigma(\cdot, a) \nabla^2 v\big) f(\cdot, a) = 0$$

Let X^* be an optimal forward controlled trajectory process and $Y_t := v(X_t^*, t)$. The *control randomization method* summarized in [16] and detailed in [18] constructs a backward stochastic differential equation (BSDE) satisfied by Y. Its main ingredient is to reformulate the control as a Poisson jump process I (expressed below) valued in $\mathcal{A}$ and independent of the forward trajectory, and to show that at each time step the transition dY_t derives from a minimization over those auxiliary processes. The presence of control-related terms makes this BSDE tricky to formulate here, therefore in this subsection we only elaborate on its numerical solution. The theoretical underpinnings of the continuous setup are outlined in [18], the bounds on the discrete-time approximation are detailed in [19], the numerical algorithm is completely outlined in [20].

7.3.2 Connection to Poisson measures and numerical scheme

Let's now describe the process I replacing the control. Suppose that $\mathcal{A}$ is supported by a finite measure λ of finite total mass $M := \lambda(\mathcal{A}) < +\infty$. Let S be a sequence of random variables exponentially-i.i.d with parameter M and ζ be a sequence of random variables uniformly-i.i.d under λ/M in $\mathcal{A}$. Then:

$$\forall t \in [t_0, \tau],\ I_t := \zeta_k,\ k := \min\left(n \in \mathbb{N} \mid\ t > t_0 + \sum_{i=0}^{n} S_i\right)$$

Notice that this process does not need to be further discretized in the numerical scheme, being already so. Numerical schemes are typically characterized by a partition $\pi = \{t_0, \cdots, t_n := \tau\}$ of the time interval in $n \in \mathbb{N}^*$ time steps $\Delta t_i = t_{i+1} - t_i$, with modulus $|\pi| = \max_{i \in [\![0, n-1]\!]} |t_{i+1} - t_i|$. The scheme outlined in [20] consists in the following recursion:

$$\begin{array}{ll} X^\pi_{t_0} = x_0 & X^\pi_{t_{i+1}} = X^\pi_{t_i} + \mu(X^\pi_{t_i}, I^\pi_{t_i})\Delta t_i + \sigma(X^\pi_{t_i}, I_{t_i})\Delta W^\pi_{t_i} \\ Y^\pi_{t_n} = h(X_{t_n}) & \mathcal{Y}^\pi_{t_i} = \mathbb{E}(Y^\pi_{t_{i+1}} | X^\pi_{t_i}) + \Delta t_i f(X^\pi_{t_i}, I^\pi_{t_i}) \quad Y_{t_i} = \min\limits_{a \in \mathcal{A}} \mathbb{E}(\mathcal{Y}^\pi_{t_i} \mid X_{t_i}, I_{t_i} = a) \end{array}$$

In particular, the last minimization in a gives the optimal control at $t = t_0$. The conditional expectations featured in the numerical scheme are approximated in [20] through empirical projections, though they can be computed through other analytical methods such as Malliavin calculus [17]. According to [16], under general hypotheses the numerical performance of this scheme is:

$$\max_{i \in [\![0,n]\!]} |Y^\pi_{t_i} - Y_{t_i}| = O(|\pi|^{1/6})$$

7.3.3 Application in quantum control

Consider the general problem of driving a quantum system from an origin $|\psi^{\text{init}}\rangle$ to a target $|\psi_f\rangle$ under a deterministic hamiltonian $\hat{H}$ controlled by a time-varying signal $a_t \in \mathcal{A}$ (lying in a compact set $\mathcal{A}$, generally equal to the unit complex disk), inducing a small stochastic perturbation $\varepsilon dy_t \hat{G}(a_t)$ in this evolution on the system. The quality of the evolution is quantified by the average fidelity, yielding the stochastic control problem:

$$\begin{aligned} &\text{minimize } \mathbb{E}\left(1 - |\langle \psi(\tau;\varepsilon) | \psi_{\text{targ}} \rangle|^2\right) \\ &\text{subject to } |\psi(0;\varepsilon)\rangle = |\psi^{\text{init}}\rangle \\ &\qquad\qquad a_t \in \mathcal{A} \\ &\qquad\qquad \imath |\dot{\psi}(t;\varepsilon)\rangle = (\hat{H}(a_t) + \varepsilon dy_t \hat{G}(a_t)) |\psi(t;\varepsilon)\rangle \end{aligned}$$

As ε is an experimental parameter fluctuating accross setups, this problem is usually solved in perturbation theory, so the trajectory is decomposed as: $|\psi(t;\varepsilon)\rangle = |\psi^0_t\rangle +$

$\varepsilon |\psi_t^1\rangle + \varepsilon^2 |\psi_t^2\rangle + O(\varepsilon^3)$. Consequently:

$$|\langle\psi_\tau^0|\psi(\tau;\varepsilon)\rangle|^2 = 1 + \varepsilon\Re(\langle\psi_\tau^0|\psi_\tau^1\rangle) + \varepsilon^2(\Re(\langle\psi_\tau^0|\psi_\tau^2\rangle) + |\langle\psi_\tau^0|\psi_\tau^1\rangle|^2) + O(\varepsilon^3)$$

Plugging the Taylor expansion in the noisy dynamics yield:

$$\imath d\begin{pmatrix}|\psi_t^0\rangle\\|\psi_t^1\rangle\\|\psi_t^2\rangle\end{pmatrix} = \begin{pmatrix}dt\hat{H}(a_t) & 0 & 0\\ dy_t\hat{G}(a_t) & dt\hat{H}(a_t) & 0\\ 0 & dy_t\hat{G}(a_t) & dt\hat{H}(a_t)\end{pmatrix}\begin{pmatrix}|\psi_t^0\rangle\\|\psi_t^1\rangle\\|\psi_t^2\rangle\end{pmatrix}$$

Since ψ^0 is deterministic, there are no dW_t terms in $d|\psi_t^0\rangle$ and infinitesimally speaking:

$$\langle d\psi_t^0|d\psi_t^1\rangle = \langle d\psi_t^0|d\psi_t^2\rangle = 0$$

Using the Ito product rule:

$$\begin{aligned}d\langle\psi_t^0|\psi_t^1\rangle &= \langle d\psi_t^0|\psi_t^1\rangle + \langle\psi_t^0|d\psi_t^1\rangle + \langle d\psi_t^0|d\psi_t^1\rangle = \langle d\psi_t^0|\psi_t^1\rangle + \langle\psi_t^0|d\psi_t^1\rangle\\ &= \imath\left(\langle\psi_t^0|\hat{H}(a_t)|\psi_t^1\rangle\,dt - \langle\psi_t^0|\hat{H}(a_t)|\psi_t^1\rangle\,dt - \langle\psi_t^0|\hat{G}(a_t)|\psi_t^0\rangle\,dy_t\right)\\ &= -\imath\langle\psi_t^0|\hat{G}(a_t)|\psi_t^0\rangle\,dy_t\end{aligned}$$

Adding that $|\psi_0^1\rangle = 0$, the hermitianity of $\hat{G}$ implies that $\imath\langle\psi_\tau^0|\psi_\tau^1\rangle \in \mathbb{R}$, so $\Re(\langle\psi_\tau^0|\psi_\tau^1\rangle) = 0$.

$$\begin{aligned}d\langle\psi_t^0|\psi_t^2\rangle &= \langle d\psi_t^0|\psi_t^2\rangle + \langle\psi_t^0|d\psi_t^2\rangle + \langle d\psi_t^0|d\psi_t^2\rangle = \langle d\psi_t^0|\psi_t^2\rangle + \langle\psi_t^0|d\psi_t^2\rangle\\ &= \imath\left(\langle\psi_t^0|\hat{H}(a_t)|\psi_t^2\rangle\,dt - \langle\psi_t^0|\hat{H}(a_t)|\psi_t^2\rangle\,dt - \langle\psi_t^0|\hat{G}(a_t)|\psi_t^1\rangle\,dy_t\right)\\ &= -\imath\langle\psi_t^0|\hat{G}(a_t)|\psi_t^1\rangle\,dy_t = \imath\langle\psi_t^1|\hat{H}(a_t)|\psi_t^1\rangle - \langle d\psi_t^1|\psi_t^1\rangle\end{aligned}$$

The hermitianity of $\hat{H}$ implies that $\imath\langle\psi_t^1|\hat{H}(a_t)|\psi_t^1\rangle \in \mathbb{R}$ so $\Re(\langle\psi_t^1|\hat{H}(a_t)|\psi_t^1\rangle) = 0$.

$$\begin{aligned}\Re(d\langle\psi_t^0|\psi_t^2\rangle) &= -\Re(\langle d\psi_t^1|\psi_t^1\rangle) = -\frac{1}{2}(\langle d\psi_t^1|\psi_t^1\rangle + \langle\psi_t^1|d\psi_t^1\rangle + \langle d\psi_t^1|d\psi_t^1\rangle - \langle d\psi_t^1|d\psi_t^1\rangle)\\ &= -\frac{1}{2}(d\langle\psi_t^1|\psi_t^1\rangle - \langle d\psi_t^1|d\psi_t^1\rangle)\end{aligned}$$

Suppose further that y is Markovian and satisfies a SDE:

$$dy_t = f(y_t)dt + g(y_t)dW_t$$

This hypothesis is justified below. In infinitesimal terms, this gives:

$$\langle d\psi_t^1|d\psi_t^1\rangle = g(y_t)^2\langle\psi_t^1|\hat{G}(a_t)^2|\psi_t^1\rangle\,dt$$

Adding that $|\psi_0^1\rangle = |\psi_0^2\rangle = 0$:

$$\Re(\langle\psi_\tau^0|\psi_\tau^2\rangle) = \frac{1}{2}\left(\int_0^\tau g(y_t)^2\langle\psi_t^1|\hat{G}(a_t)^2|\psi_t^1\rangle\,dt - \langle\psi_\tau^1|\psi_\tau^1\rangle\right)$$

Thus ψ^2 is removed from the picture, consequently:

$$|\langle\psi_\tau^0|\psi(\tau;\varepsilon)\rangle|^2 = 1 + \varepsilon^2\left(\int_0^\tau g(y_t)^2\,\langle\psi_t^1|\hat{G}(a_t)^2|\psi_t^1\rangle\,dt + \Re(\langle\psi_\tau^0|\psi_\tau^1\rangle)^2 - \langle\psi_\tau^1|\psi_\tau^1\rangle\right) + O(\varepsilon^3)$$

In a robust control perspective, we seek to minimize the second order term of this expression, while imposing a constraint $\psi_\tau^0 = \psi_{\text{targ}}^0$ through a penalizing infidelity term tied to a large positive constant $A > 0$. The eventual formulation of the stochastic control problem is then:

$$\begin{aligned}
&\text{minimize } \mathbb{E}\left(\int_0^\tau g(y_t)^2\,\langle\psi_t^1|\hat{G}(a_t)^2|\psi_t^1\rangle\,dt + \Re(\langle\psi_\tau^0|\psi_\tau^1\rangle)^2 - \langle\psi_\tau^1|\psi_\tau^1\rangle + A(1 - |\langle\psi_\tau^0|\psi_{\text{targ}}\rangle|^2)\right)\\
&\text{subject to } (|\psi_0^0\rangle, |\psi_0^1\rangle, y_0) = (|\psi^{\text{init}}\rangle, 0, 0)\\
&\qquad a_t \in \mathcal{A}\\
&\qquad \imath d|\psi_t^0\rangle = \hat{H}(a_t)\,|\psi_t^0\rangle\,dt\\
&\qquad \imath d|\psi_t^1\rangle = \hat{H}(a_t)\,|\psi_t^1\rangle\,dt + \hat{G}(a_t)\,|\psi_t^0\rangle\,dy_t\\
&\qquad dy_t = f(y_t)dt + g(y_t)dW_t
\end{aligned}$$

Formulated in a stochastic control setup, this problem can be treated with the previously mentioned control randomization algorithm, outlined in [20].

7.3.4 SDE representations of pink noise

A stochastic process is called a *pink noise* or $1/f$-noise process when its power spectral density (PSD) [2] is proportional to $f \mapsto 1/f$ in a fixed finite frequency window $[f_{\min}, f_{\max}] \subsetneq]0, +\infty[$. Pink noise is omnipresent in control eletronics, also known as 'flicker noise' in this context, so it is expected to play a role in the optical apparatus of atomic quantum computers as well. It can subtly be approximated by a long-range markovian continuous process through nonlinear multiplicative SDE's, such as this one coined in [25] (here $\eta > 1/2$ is a constant parameter):

$$dy_t = \left(\eta - \frac{\nu}{2}\right) y^{2\eta-1}dt + y^\eta dW_t$$

We expect the contructions in [26] to be generalizable to other spectra, so this could be used in quantum control by converting generic noise models to SDE. Deeper mathematical details about this SDE are found in [27]. An open question here is how "unique" that representation is.

[2] The PSD S of a stochastic process y is technically defined as: $S(\nu) = \lim_{B\to\infty} \frac{1}{B}\mathbb{E}\left(\left|\int_{-B}^{B} y_t e^{\imath 2\pi t\nu} dt\right|^2\right)$.

7.3.5 Discussion

The preparation of specific quantum states under a perturbed hamiltonian is a recurring problem in quantum engineering, appearing e.g through the design of robust quantum logic gates in various hardware types. Statistical assumptions on the noise process considerably ease the derivation of optimal motion equations, as seen in the previous case of a stationary process. On another leaf of this tree, the noise process can instead be considered as continuous in time and markovian, thus expressed through a stochastic differential equation. This time, the state preparation problem is a purely *stochastic control problem*, whose solution by control discretization would again incur a number of variables detrimental to its efficiency, and might be trapped in local minima if resorting to gradient descent algorithms. The *control randomization method* solves both those issues by computing the *value function* of the control problem along an optimal path, retrieving the optimal steering in the process. Its main ingredient is a reformulation of the value function as a maximum over Poisson jump processes replacing the controls and independent from the trajectory, simplifying the calculations, benefitting from numerical schemes and convergence bounds of backward stochastic differential equation. The main advantage of this framework is to compute an ensured approximation of the **global optimum** of the OC problem, within controllable accuracy and certainty. The control can be later refined to the genuine target through a gradient descent algorithm.

The usage of this approach in neutral-atomic pulse shaping is vindicated by the derivation of stochastic differential equations for pink noise, commonly featured in the electro-optical machinery of this quantum hardware platform. An interesting research perspective would be to construct SDEs for other recurring noise spectra in quantum hardware.

A computational drawback of this method is its open-loop character: the optimal control is computed at each time-step, requiring each time both a forward and a backward loops. Therefore, this problem would benefit from future tighter convergence bounds on the stochastic numerical scheme.

7.4 Conclusion

The design of quantum gates robust to control noise and excited state decay can be formulated through the lenses of optimal control (OC) theory, which provides a remarkable structure to this problem.

First, when considering a stationary gaussian control noise, the robust pulse

shaping problem is converted to a deterministic OC problem with additional trajectories corresponding to individual frequency rays in the discretized noise spectrum. The Pontryagin principle then computes the optimal light pulse through auxiliary costate trajectories. However, the high nonlinearity of the induced boundary value problem and the high number of frequency rays imply that initializing those costates is a hard root-finding problem, arguably affected by the curse of dimensionality. This is answered by a deeper characterization of the OC problem, its value function (VF), whose gradient is the sought costate. Under reasonable conjectures about its vanishing viscosity approximation and numerical schemes for backward stochastic differential equations (SDE), the nonlinear Feynman-Kac theorem provides a Monte Carlo algorithm to approximate the VF and its gradient that converges independently of the dimensionality of its variables. This allows to solve the OC problem while avoiding the curse of dimensionality.

Second, when considering the control as a time-continuous and markovian process thus represented by a SDE, the robust pulse shaping is a stochastic OC problem. Optimizing the objective cost in terms of a time-discretized control in many steps puts the derived pulse at risk of getting trapped in local minima. Again, this issue is solved by computing the value function through the control randomization method, guaranteeing a globally optimal pulse in the end. This method is vindicated by the derivation of SDE for pink noise, which is commonly featured in the opto-electronic control machinery in neutral-atomic quantum hardware computing. Formulating SDE for other common spectra would further enhance the generalizability of that method in other quantum hardware challenges.

Bibliography

[1] Koch, C.P., Boscain, U., Calarco, T. et al. Quantum optimal control in quantum technologies. Strategic report on current status, visions and goals for research in Europe. EPJ Quantum Technol. 9, 19 (2022). https://doi.org/10.1140/epjqt/s40507-022-00138-x (Cited on pages 9 and 106.)

[2] Khaneja, N., Reiss, T., Kehlet, C., Schulte-Herbrüggen, T., & Glaser, S. J. (2005). Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms. Journal of magnetic resonance, 172(2), 296-305. https://doi.org/10.1016/j.jmr.2004.11.004, https://www.ch.nat.tum.de/fileadmin/w00bzu/ocnmr/pdf/94_GRAPE_JMR_05_.pdf (Cited on pages 106 and 108.)

[3] Sikorski, K. and Trojan, G.M., 1990. Asymptotic near optimality of the bisection method. Numerische Mathematik, 57, pp.421-433. https://core.ac.uk/download/pdf/276277391.pdf (Cited on page 110.)

[4] Jandura, S. and Pupillo, G., 2022. Time-optimal two-and three-qubit gates for Rydberg atoms. Quantum, 6, p.712. (Cited on pages xxiv, 26, 53, 54, 61, 79, 81 and 114.)

[5] Tran, H.V., 2021. Hamilton–Jacobi equations: theory and applications (Vol. 213). American Mathematical Society.. (Not cited.)

[6] Cirant, M., & Goffi, A. (2025). Convergence rates for the vanishing viscosity approximation of Hamilton-Jacobi equations: the convex case. arXiv preprint arXiv:2502.15495. https://cvgmt.sns.it/media/doc/paper/7004/RateHJviscous4.pdf (Cited on page 111.)

[7] Camilli, F., Goffi, A., & Mendico, C. (2023). Quantitative and qualitative properties for Hamilton-Jacobi PDEs via the nonlinear adjoint method. arXiv preprint arXiv:2307.12932. https://journals.sns.it/index.php/annaliscienze/article/view/6681/2147 (Cited on page 111.)

[8] Chang, D.E., 2011. A simple proof of the Pontryagin maximum principle on manifolds. Automatica, 47(3), pp.630-633. (Cited on page 109.)

[9] Mangasarian, O.L., 1966. Sufficient conditions for the optimal control of nonlinear systems. SIAM Journal on control, 4(1), pp.139-152. (Cited on page 109.)

[10] Barron, E.N. and Jensen, R., 1986. The Pontryagin maximum principle from dynamic programming and viscosity solutions to first-order partial differential equations. Transactions of the American Mathematical Society, 298(2), pp.635-641. (Cited on page 109.)

[11] Cannarsa, P. and Sinestrari, C., 2004. Semiconcave functions, Hamilton-Jacobi equations, and optimal control (Vol. 58). Springer Science & Business Media. https://www.ceremade.dauphine.fr/~pbernard/enseignement/M2/Cannarsa%20P.,%20Sinestrari%20C.%20Semiconcave%20functions,%20Hamilton-Jacobi%20equations,%20and%20optimal%20control%20(Birkhauser,%202004)(ISBN%200817643362)(311s)_MCv_.pdf (Cited on pages 111 and 118.)

[12] Vrahatis, M.N. and Iordanidis, K.I., 1986. A rapid generalized method of bisection for solving systems of non-linear equations. Numerische Mathematik, 49, pp.123-138. (Cited on page 110.)

[13] Vrahatis, M.N., 2020. Generalizations of the intermediate value theorem for approximating fixed points and zeros of continuous functions. In Numerical Computations: Theory and Algorithms: Third International Conference, NUMTA 2019, Crotone, Italy, June 15–21, 2019, Revised Selected Papers, Part II 3 (pp. 223-238). Springer International Publishing. (Cited on page 110.)

[14] Vrahatis, M.N., 1988. Algorithm 666: Chabis: a mathematical software package for locating and evaluating roots of systems of nonlinear equations. ACM Transactions on Mathematical Software (TOMS), 14(4), pp.330-336. (Cited on page 110.)

[15] Brockwell, P.J., Davis, R.A. (1987). The Spectral Representation of a Stationary Process. In: Time Series: Theory and Methods. Springer Series in Statistics. Springer, New York, NY. https://doi.org/10.1007/978-1-4899-0004-3_4 (Cited on page 116.)

[16] Pham, H. (2015). Feynman-Kac representation of fully nonlinear PDEs and applications. Acta Mathematica Vietnamica, 40, 255-269. (Cited on pages 112, 113, 118 and 119.)

[17] Bouchard, B., Ekeland, I., and Touzi, N. (2004). On the Malliavin approach to Monte Carlo approximation of conditional expectations. Finance and Stochastics, 8, 45-71. (Cited on pages 112 and 119.)

[18] Kharroubi, I. and Pham, H., 2015. Feynman–Kac representation for Hamilton–Jacobi–Bellman IPDE. The Annals of Probability, pp.1823-1865. (Cited on page 118.)

[19] Kharroubi, I., Langrené, N. and Pham, H., 2015. Discrete time approximation of fully nonlinear HJB equations via BSDEs with nonpositive jumps. (Cited on page 118.)

[20] Kharroubi, I., Langrené, N. and Pham, H., 2014. A numerical algorithm for fully nonlinear HJB equations: an approach by control randomization. Monte Carlo Methods and Applications, 20(2), pp.145-165. (Cited on pages 112, 118, 119 and 121.)

[21] Camilli, F., Goffi, A., & Mendico, C. (2023). Quantitative and qualitative properties for Hamilton-Jacobi PDEs via the nonlinear adjoint method. arXiv preprint arXiv:2307.12932. (Not cited.)

[22] Aronna, M.S., Tonon, D., Boccia, A., Campos, C.M., Mazzola, M., Nguyen, L.V., Palladino, M., Scarinci, T. and Silva, F.J., 2017. Optimality conditions (in pontryagin form). Optimal Control: Novel Directions and Applications, pp.1-125. (Not cited.)

[23] Zheng, W., Zhang, Y., Dong, Y. et al. Optimal control of stimulated Raman adiabatic passage in a superconducting qudit. npj Quantum Inf 8, 9 (2022). https://doi.org/10.1038/s41534-022-00521-7 (Not cited.)

[24] Hu, Y., Nualart, D. and Song, X., 2011. Malliavin calculus for backward stochastic differential equations and application to numerical solutions. (Cited on pages 112 and 113.)

[25] Ruseckas, J. and Kaulakys, B., 2010. 1/f noise from nonlinear stochastic differential equations. Physical Review E—Statistical, Nonlinear, and Soft Matter Physics, 81(3), p.031105. (Cited on page 121.)

[26] Kaulakys, B. and Ruseckas, J., 2004. Stochastic nonlinear differential equation generating 1/f noise. Physical Review E—Statistical, Nonlinear, and Soft Matter Physics, 70(2), p.020101. (Cited on page 121.)

[27] Ruseckas, J. and Kaulakys, B., 2011. Tsallis distributions and 1/f noise from nonlinear stochastic differential equations. Physical Review E—Statistical, Nonlinear, and Soft Matter Physics, 84(5), p.051125. (Cited on page 121.)

Chapter 8

Discussion and perspectives

Sommaire

8.1 Discussion

General optimal control (OC) problems are solved by either:

- Discretizing the control signal, differentiating the loss function with regard to the control signal based on the controlled trajectory evolution, and running a gradient descent algorithm to locate the minimum of the problem.
- Constructing a 'costate' dual trajectory (corresponding to the Lagrange multiplier of the primal trajectory constraint) that satisfies a certain backwards differential equation. This costate is initialized by solving a root-finding problem associated to a certain function accounting for the forward-backward system of ordianry differential equations of the primal and dual trajectories.

In early chapters, we have tackled the problem of designing two-qubit gates for Rydberg-atomic platforms, which are robust with regard to a stochastic time-dependent control noise expressed in harmonic sum decomposition. But even with simplification of the noise, and accounting for a finite number of noise samples, we reasonably expect the usual optimal control-theoretic tools not to work there, for two reasons:

- Due to the stochasticity of the problem, control discretization might result in an even less convex loss function of the control signal, so that gradient-based updates would tend even more to be trapped in local minima far from the global one.
- Costate initialization does not work efficiently because here the trajectory dimension is proportional to the number of samples considered in this stochastic problem, a quantity that must be close to zero in order to estimate the control loss function with high accuracy, and therefore the problem's dimensionality blows up.

Therefore, being intrinsically stochastic with a seemingly broken differential structure, this problem seems, at first glance, a good testing ground for artificial evolution algorithms, and in particular to the Covariance Matrix Adapatation Evolution Strategy. Effectively, applying this algorithm gave good results, and with deeper computational power, better results are expected to come in various other stochastic noise settings.

Importantly, we also noticed that the design of robust such two-qubit gates showcases other capital criteria which are conflicting with noise robustness. So, under the previous observation, this would make the problem further compatible with multi-objective artificial evolution, and in this vein we successfully applied the Non-Dominated Sorting Genetic Algorithm III to identify a surface of Pareto-optimal quantum gate setups, with regard to control noise robustness and sensitivity to Rydberg state decay. Similarly to the previous prospect, it is expected that with a deeper knowledge of atomic-optical dynamics, computational power, and more adapted algorithms,

this approach would identify other optimal robustness tradeoffs including other criteria.

Recent advances in hardware control for neutral atoms led us to rethink our physical models, and notably the noise levels considered. While we initially worked such levels between 0.1 and 0.3 (with regard to the maximal power of the control electro-magnetic field), counting on an increase of the quality of the optical control apparatus, lower levels closer to 0.01 should be considered instead. At this scale, perturbation-theoretic methods become relevant, and the control problems retrieves its full differential structure, which questions the relevance of artificial evolution in that context. Indeed, we even observed that in perturbation theory, much more could be said about the geometry of quantum errors stemming from control noise in the Rydberg blockade setup.

By leveraging pertubation-theoretic expansions of the propagator of a noisy light-atom hamiltonian, we could derive the full error channels of the control noise on the atomic setup for both frequency and intensity noises, separately so. The key analytical ingredient was the statiscally reasonable and common assumption that the noise process satisfies a Langevin equation. More precisely, for each noise type (frequency or intensity) three error scenarii were discussed, depending on wether the atoms unwantedly left in a Rydberg state after the faulty gate are forcefully projected to one of qubit states or an equal statistical mixture thereof.

With these six different types of error channels in hand, we took the opportunity to test them against a frequent quantum many-body problem, the simulation of a quantum transverse Ising model, either antiferromagnetic or ferromagnetic. Therein, in each noise scenario and post-gate optical projection, the Pauli-ZZ two-body operations are bracketed with the adequate two-qubit error channels. For each of the six scenarii and on different sizes of atomic arrays, we computed relevant metrics such as circuit fidelity, Pauli two-body correlations and operator entanglement entropy. Numerical simulations showed that both the hamiltonian types, noise types and optical projection schemes significantly impact all those metrics. In particular, the best state to project after the noisy gate strongly depends on wether the emulated hamiltonian is antiferromagnetic or ferromagnetic. The bottom line of this work in chapter 6, is that the perturbation-theoretic approximation of the propagator gives a handful of relevant informations for numerical experiments.

With perturbation theory and this previous work on error channels in hand, we reconsidered our position on the opportunity of the optimal control structure for robust gate control problems, with regard a to stochastic time-dependent control noise. In fact, leveraging the noise spectral decomposition already used in chapter 4, the studied problem can be reformulated in an optimal control perspective, where the overall tra-

jectory dimension is driven by the number of relevant frequencies in the noise spectrum (instead of the number of trajectory samples). However, this number is still quite high to address realistic spectra, keeping the problem in the range of the curse of dimensionality. To address this issue, we have proposed in chapter 7 an algorithm to initialize the dynamical Lagrange multipliers $p_0 \in \mathbb{R}^d$ of general optimal control-theory that bypasses the computational bottleneck of worst-case root-finding procedures, which are seemingly necessary in optimal control. This approach feeds on those ingredients:

- The OC problem is encapsulated by a value function v_0, whose gradient (at the trajectory starting point $x_0 \in \mathbb{R}^d$ and time t_0) is the Lagrange multiplier p_0 sought to compute (in the conventional setting): $\nabla v_0(x_0, t_0) = p_0$.
- This value function v_0 satisfies the first-order nonlinear 'Hamilton-Jacobi-Bellman' PDE.
- In turn, this PDE allows to formulate a 'vanishing viscosity' v_σ approximation to the value function v_0. v_σ satisfies a second order semi-linear PDE.
- The stochastic control interpretation and various PDE properties of v_σ bound not only the value discrepancy $\|v_\sigma - v_0\|$ but also the gradient discrepancy $\|\nabla v_\sigma - \nabla v_0\|$ as $\sigma \to 0$.
- The nonlinear Feynman-Kac formula allows to compute, pointwise, the solution to a general semi-linear elliptic second-order PDE through a Forward-Backward-Stochastic Differential Equations system.
- In particular, this system also features both the value and the gradient of the function under study, through an adjoint process.
- Numerical solutions to BSDES allow to approximate both within a overall computational complexity polynomially dependent on the desired accuracy, and independently of the forward trajectory dimension.
- Putting all pieces together, by applying the previous points to v_σ and its semi-linear PDE with a carefully chosen small viscosity parameter σ, one gets an approximation of $p_0 \nabla v_0(x_0, t_0)$ independent of d. Due to previous convergence arguments, this approximation is expected to be accurate enough to be fed into a Newton method (running polynomially in d) to retrieve the true Lagrange multiplier and solve the control problem.

While its theoretical underpinnings must still be refined, this method addresses the dimensionality bottleneck of optimal control (i.e usual root-finding shooting methods scaling exponentially in d without a good initial guess). We propose to apply it on the robust control problem where the control noise is assumed to be a stationary stochastic

process, admitting therefore a harmonic expansion. To find that we previously wanted to answer with artificial evolution. Assuming the

An alternative formulation of this problem is to assume that instead of being stationary, the control noise is a time-continuous Markovian stochastic process, therefore satisying a stochastic differential equation. For example, this is the case for pink noise, also known as '1/f noise' in reference to the hyperbolic shape of its power spectral density. Being frequently featured in the electronics omnipresent in optical apparatii, this noise spectrum is expected to widely corrupt the control of neutral-atomic qubits and thus falls within the scope of this thesis. Also, the synthesis of a given noise process into a SDE is not exclusive to that and is applicable to other spectra.

Consequently, in this formulation the robust control problem can be solved through stochastic optimal control algorithms based on the value function and Hamilton-Jacobi-Bellman equation frameworks. In particular, we propose in chapter 7 to tackle it with control randomization, an algorithm whose main ingredient is to reformulate the control as a Poisson jump process, and the value function as an infimium over the Poisson measures supporting the control jump processes. The value function is computed by a stochastic differential equation and the optimal policy is updated at each step through conditional expectations, down to the starting time where this approach outputs an approximation to the optimal control. Being thoroughly grounded in stochastic control theory, this method reasonably outputs an approximation to the global minimum of the control problem, then to be refined by local gradient updates. This answers the issue of outputting the global minimum.

So, while the problem of robust gate synthesis under stochastic control noise seemed at first glance a good testing ground for artificial evolution, the deeper details of stochastic optimal control theory contradict the two main justifications of this position. Indeed, for each of those two arguments, we have formulated algorithms inspired by the mathematical finance litterature to solve the problem, for an adapted formulation of the noise process though. The existence of such reductions is not exclusive to physical control problems and we expect it to be a common feature of continuously-formulated computational problem in general: for example, it has been shown in machine learning that replacing residual neural networks with continuous layers, led to a dramatic decrease of the number of necessary parameters in the model. Therefore, faced with this persitence of mathematical structure and deeper conventional algorithms for our problem, it seems that artificial evolution is not the best figher in hand.

An argument against this reasoning, is that the algorithms considered actually depend on specific formulations of the noise process, each corresponding to a mathematical

property thereof: either stochastic harmonic decomposition (for stationary processes) or stochastic differential equations (for continuous markovian processes). In other words, the OC structure would persist precisely and solely because of the mathematical assumptions on the noise processes and a consequence would be that for other noise generation methods corresponding to loosened mathematical conditions, the OC structure could really break for good, genuinely leaving room for artificial evolution. However, to the best of our knowledge, even though they are tied to narrow mathematical assumptions, the methods we used are actually the only ones to emulate noise processes with a given power spectrum anyway. In other words, we are not aware of any computationally efficient way to sample non-stationary and non-markovian noise with a given power spectral density. So, given our knowledge, any efficient noise generation method (in fact any method *at all*) results in fairly structured problems because we can't generate relevant noise beyond that. Simulating non-stationary and non-markovian noise with a given power spectral density is quite difficult. So the design of robust gates is fundamentally limited by the ability to simulate realistic noise, which itself has its own standing assumptions that may not correspond to actual realsitic setups. In other words, the problem of synthesing gates is already structured enough to be tackled by the tools from optimal control and dynamic programming:

- In the deterministic case, the reduction to a shooting problem fatally reduces the number of relevant parameters and we have developed a Monte Carlo method for this. Plus, even when accounting for multi-objectives
- In the noisy case, if the noise process is expressed through:
  - A randomly weighted harmonic decomposition as posited in chapter 4, then the very assumptions behind this decomposition also allows to formulate the robustness problem in a deterministic control setup, although with many frequencies. In this case, our Monte-Carlo method in chapter 7 is a good solution candidate.
  - A stochastic differential equation, then the control robustness problem can be solved through the control randomization method as pointed in chapter 7

8.2 Perspectives

While the physics of the gate synthesis problem allow for sophisticated control algorithms seen before, there is in fact still room for artificial evolution (AE) in another unavoidable segment of quantum computing. As already said in the introduction, *Quantum Error Correction* (QEC) is necessary to one day implement genuine quantum computation without errors, and numerical optimization is expected to play a big role in

this field. For example, some QEC code words and spaces are constructed with the help of numerical optimization algorithms (see the content and references in [1]), and in some other QEC setups the recovery map optimization step of the decoding phase can be cast as a convex optimisation problem, more precisely semi-definite programming [4, 5], see also chapter 13 in [2]. It is possible to optimize the QEC setup beyond this structured setting, for example the scheme in [3] features the Matrix Rank Minimization Problem (RMP) [3]. The latter is a NP-hard combinatorial problem generally tackled with heuristic approaches, so this would be a good ground for artificial evolution algorithms to play a role in QEC. In another unstructured setting, reinforcement learning (RL) has showcased remarkable abilities to construct efficient QEC codes [6, 7, 8]. On the other hand, Evolution Strategies (ES, a subset of AE) have been proposed as scalable alternatives to reinforcement learning (RL) [9] by experts at OpenAI (a company known to usually promote and rely heavily on RL). An interesting research perspective would be to check wether the construction of QEC codes counts among the RL applications benefitting from a transition from RL to ESs, showcasing the probable contribution of AE to QEC in particular and the developement of quantum computing hardware in general.

Bibliography

[1] "Numerically optimized bosonic code", The Error Correction Zoo (V. V. Albert and P. Faist, eds.), 2022. https://errorcorrectionzoo.org/c/numopt (Cited on page 133.)

[2] Lidar, D.A. and Brun, T.A. eds., 2013. Quantum error correction. Cambridge university press. (Cited on pages 5 and 133.)

[3] Taghavi, S., Kosut, R.L. and Lidar, D.A., 2010. Channel-optimized quantum error correction. IEEE transactions on information theory, 56(3), pp.1461-1473. (Cited on page 133.)

[4] Taghavi, S., Brun, T.A. and Lidar, D.A., 2010. Optimized entanglement-assisted quantum error correction. Physical Review A—Atomic, Molecular, and Optical Physics, 82(4), p.042321. (Cited on page 133.)

[5] Kosut, R.L. and Lidar, D.A., 2009. Quantum error correction via convex optimization. Quantum Information Processing, 8, pp.443-459. (Cited on page 133.)

[6] Nautrup, H.P., Delfosse, N., Dunjko, V., Briegel, H.J. and Friis, N., 2019. Optimizing quantum error correction codes with reinforcement learning. Quantum, 3, p.215. https://doi.org/10.22331/q-2019-12-16-215 (Cited on page 133.)

[7] Olle, J., Zen, R., Puviani, M. et al. Simultaneous discovery of quantum error correction codes and encoders with a noise-aware reinforcement learning agent. npj Quantum Inf 10, 126 (2024). https://doi.org/10.1038/s41534-024-00920-y (Cited on page 133.)

[8] Su, V.P., Cao, C., Hu, H.Y., Yanay, Y., Tahan, C. and Swingle, B., 2025. Discovery of optimal quantum codes via reinforcement learning. Physical Review Applied, 23(3), p.034048. https://doi.org/10.1103/PhysRevApplied.23.034048, https://arxiv.org/abs/2305.06378. (Cited on page 133.)

[9] Salimans, T., Ho, J., Chen, X., Sidor, S., & Sutskever, I. (2017). Evolution strategies as a scalable alternative to reinforcement learning. arXiv preprint arXiv:1703.03864. https://arxiv.org/abs/1703.03864. See the blog post on this: https://openai.com/index/evolution-strategies/. (Cited on page 133.)

Chapter 9

Conclusion

Sommaire

9.1 General remarks

This work started with the goal of designing two-qubit gate pulses for Rydberg-atomic-based hardware, that are robust with regard to a control noise that is stochastic and time-dependent, and with regard to other error sources such as the sensitivity to photon spontaneous emission. In the beginning of this work, we identified several reasonable arguments for the inefficiency of standard optimal control (OC) methods to treat this problem. The first one was that the high stochasticity in trajectories ramping up in the loss function would likely make the control loss function highly non convex with regard to the discretized control signal, and so gradient descent methods would trap candidate controls in sub-optimal local minima of the robust control problem. Most importantly, the second argument was that in the nonperturbative regime then considered, the problem's dimensionality was dominated by the number of random samples, which must be high enough to reliably estimate the noisy quality metrics such as gate fidelity. But since continuous OC methods induce a root-finding step with regard to the evolved trajectories, their worst-case complexity scales defavorably with the trajectory dimensionality; and so in our case these methods would not work.

From those arguments, this problem, inherently stochastic and lacking some formal structure, can be considered as a black-box problem in which one cannot extract much information about the loss function beyond its values. Therefore we deducted that is a good fit for artificial evolution algorithms, which are known to work well in such black-box settings. By applying the Covariance-Matrix-Adaptation Evolution Strategy (CMA-ES) algorithm on it, we could derive two-qubit gate pulses showcasing a decent gate fidelity while faced with a stochastic control noise (both in laser frequency and intensity) of relatively high magnitude. Moreover, we also identified conflicting robustness criteria : for example, control noise robustness is easier with longer pulses, but this raises the exposure to spontaneous photon emission, therefore those two criteria are antagonistic.

AE algorithms admit natural generalizations to such *multi-objective optimisation problems*, notably by replacing their usual ranking mechanisms with Pareto dominance rankings, so this make our gate robustness problem even more adapted to the AE framework, as long as it's formulated in a black-box setting. In this vein, using the third Non-dominating Sorting Genetic Algorithm (NSGA-III), we obtained a variety of pulses achieving optimal trade-offs between robustness to control noise and sensitivity. Those results confirmed the capability of AE to solve stochastic control problems in quantum computing hardware.

However, upon a careful examination of the litterature on quantum control, we

realized that the magnitudes of the considered control noise should be lowered down to the level of reliable treatment by perturbative methods. When treating the control noise perturbatively, it is possible to separate the gate trajectory from the robustness terms. Considering that the effect of control noise is to leave some atoms in the Rydberg manifold, three choices are left to the user: either forcing the atom back to a computational basis state or to a uniform statistical mixture thereof. Then, by adding reaonable statstical assumptions on the noise stochastic process, either in frequency or intensity, the complete quantum error channel of the control noise can be derived. Upon constructing those channels, we plugged them into a common quantum many-body problem, the simulations of the kicked transverse Ising model, respectively in antiferromagnetic or ferromagnetic orientations.

Upon the calculation, significant metrics tied to the quality of the simulation were gathered, namely: the nearest-neighbour and next-nearest-neighbour two-body Pauli ZZ correlations (respectively C1 and C2) at the middle of the spin chain, the circuit fidelity (CF) and the operator space entanglement entropy (OE). The different optical noise compensation schemes (i.e the post-gate projection into the computational basis) could thus be compared on those quantities, and so for each noise type (frequency or intensity), leading to the question: which scheme preserves best a given metric ? We found that schemes preserving best CF also preserved best C1 and C2, and that the order of efficiency of schemes depended pretty much on the nature of the simulated hamiltonian: the antiferromagnetic model 'prefered' to project the atom back to $|1\rangle$ while the ferromagnetic model 'prefered' the projection to $|0\rangle$. We also found that the discrepancy between schemes were sharper in frequency noise than in intensity noise. Finally, the most interesting observation was that, for each noise and hamiltonian type, projecting the atom to $|1\rangle$ dealt the highest OE; suggesting that the classical simulation of a noisy hamiltonian compensated in this way would asymptotically require more computational resources. However, despite this higher proximity to quantum dynamics, this scheme did not always give the highest circuit fidelity. This chapter highlighted the impact of 'hardware' error mitigation schemes at the physical level, and showcased the theoretical power of perturbation theory.

Looking back at the gate robustness problem, we saw that by using perturbation theory it could be integrally converted to a standard optimal control problem, with an integral term corresponding to sensitivity to photon spontaneous emission, and many addititional robustness trajectories: one such for each frequency ray in the noise spectrum. This new formulation completely bypasses any reference to random samples of noisy trajectories. However, this optimal control problem still requires root-finding methods scaling bad in such highly-dimensional settings. This calls for an alternative preconditionner avoiding this combinatorial explosion.

Beyond constrained optimization, optimal control problems are equivalently formulated in terms of a *value function*, that satisfies the nonlinear Hamilton-Jacobi-Bellman (HJB) partial differential equation (PDE) and whose gradient is the costate variable, i.e the sought solution to the root-finding step aforementioned. The value function admits a 'vanishing viscosity' approximation that follows a second-order semi-linear PDE, and whose gradient is also an approximation to the sought costate; out of stochastic control and PDE arguments. On the other hand, the nonlinear Feynman-Kac theorem relates second-order semi-linear PDEs to systems of forward-backward stochastic differential equations (SDE), so that the former can be pointwise solved with Monte Carlo methods independently of its intrisic dimensionality. Moreover, the adequate numerical schemes also approximate its gradient. Joining these two points, we get a Monte Carlo method to initialize the costate variable of an abstract optimal control problem independently of its dimension, solving it while avoiding the curse of dimensionality affecting its conventional numerical solution. This method is expected to be particularly suited to our gate robustness problem where the many frequency rays of the control noise spectrum increase the dimensionality of the system.

Instead of a harmonic decomposition stemming from a given frequency spectrum, the noise process can be formulated as a nonlinear SDE, so the gate robustness is formulated as a stochastic control problem. This vision came quite lately in the thesis as we learned more about the deep details of stochastic control theory. While there is no clear generalization of the costate variable to reconstruct the control in a closed-loop setting, the value function framework nevertheless therein provides an open-loop method to find the pulse achieving the global minimum of the problem, called the 'control randomization' method. Like previously, it is a Monte Carlo method based on an approximation of the value function by backward SDE through the HJB PDE. This setup fits well into our problem because '1/f noise' has been approximated using SDEs, and this noise is common in the opto-electronic setups featured in neutral-atomic quantum computing hardware.

Therefore, using perturbation theory, the full derivation of quantum error channels associated to stochastic control noise led us rethink our initial view on how standard optimal control theory could solve the problem of optimizing gate pulses of Rydberg-atomic quantum computers. Indeed, we proposed two methods adressing each argument vindicating the use of artificial evolution in this problem. Since these recent proposals conserve structural properties of optimal control theory (compact representation of the problem ensured global minima), they should either replace or complete artificial evolution in this problem. However, this is not the end for artificial evolution in quantum computing hardware.

We opine that this resilience of mathematical structure is mainly a consequence of the time-continuous formulation of the gate robustness problem, as indicated by the reliance of our proposals on stochastic calculus. Therefore, artificial evolution would be much more welcome in time-discrete less structured problems, that are rather found in quantum error correction (QEC). Elaborating QEC strategies, e.g code words and space construction, recovery map optimization; features many poorly structured combinatorial problems currently tackled with heuristic techniques such as reinforcement learning. On the other hand, AE algorithms such as Evolution Strategies are known to perform in those problems, so applying them in such QEC setting would showcase a tangible contribution of artificial evolution to the development of quantum computing hardware.

9.2 Scientific communications

The reasearch in this thesis has been communicated in the following instances:

- At the start of the PhD, the content of chapter 3, genetic algorithms, applied this time to reproduce quantum annealing benchmarks, was featured in a poster, presented at the "European Spring School on Quantum Technologies" held in April 2022 in Strasbourg.
- The related topic of optimal control of quantum adiabatic paths for combinatorial optimization was the object of a presentation at the 2022 BinGI Days seminar.
- The work of chapter 4 was presented at the 'Complex Computational Ecosystems' conference held in Baku (Azerbaidjan) in April 2023. Subsequently, as part of the conference proceedings it was published as a peer-reviewed paper in:

  Ley, T., Leonteva, A.O., Schachenmayer, J., Collet, P.
  Evolutionary Reduction of the Laser Noise Impact on Quantum Gates. Complex Computational Ecosystems 2023. Lecture Notes in Computer Science, vol 13927. Springer, Cham. https://doi.org/10.1007/978-3-031-44355-8_5

- The work in chapter 7, more precisely the part about robust optimal control, was presented as a poster at the 54th conference of the European Group on Atomic Systems, held in Strasbourg in June 2023. In July 2023, it was presented at the 'BinGI Days' seminar held in Strasbourg too.

Besides those communications, the author of those lines also worked on quantum computing training programs for finance professionals for the start-up QuantFi. He also taught introductory Quantum Machine Learning classes in the QUANTUM track of the engineering school EPITA.