The disparate impact of noise on quantum learning algorithms

Armando Angrisani

The disparate impact of noise on quantum learning algorithms

By

ARMANDO ANGRISANI

Laboratoire d'Informatique de Paris 6
SORBONNE UNIVERSITÉ

A dissertation submitted to Sorbonne Université in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY, under the supervision of Elham Kashefi (*directrice de thèse*) and Vincent Cohen-Addad (*co-encadrant*).

The PhD defense was held publicly on December 11, 2023, with the following thesis committee:

- TOM GUR, *Rapporteur*
  Associate Professor, University of Cambridge (United Kingdom)
- RIK SARKAR, *Rapporteur*
  Associate Professor, University of Edinburgh (United Kingdom)
- IORDANIS KERENIDIS, *Examinateur*
  Directeur de Recherche, Université Paris Cité
- ELHAM KASHEFI, *Directrice de thèse*
  Directrice de Recherche, Sorbonne Université
  Professor, University of Edinburgh (United Kingdom)

FOREWORD

This thesis encapsulates a three-year journey immersed in continuous learning. I have been fortunate to thrive in a supportive environment, surrounded by brilliant researchers who played a pivotal role in shaping my development as an independent thinker.

First and foremost, my gratitude extends to my supervisor, Elham Kashefi. She welcomed me into her group during my master's thesis and has been a constant source of inspiration throughout this journey. Elham's unwavering support and guidance significantly contributed to my growth and accomplishments.

Special appreciation goes to Vincent Cohen-Addad for gently introducing me to research on algorithms and theoretical computer science. He not only shared his passion but also demonstrated unwavering dedication to the field.

I extend my sincere gratitude to the referees Tom Gur and Rik Sarkar for their meticulous examination of this thesis, valuable insights, and constructive feedback.

I am also thankful for the valuable opportunities to explore other quantum groups and appreciate Anna Pappa, Jens Eisert, Vedran Dunjko, Daniel Stilck França, Omar Fawzi, and Zoe Holmes for welcoming me into their respective groups, offering new perspectives.

A heartfelt thanks goes to Daniel for mentoring me during my stay in Lyon and the subsequent months. Our discussions played a crucial role in my research interests and understanding of quantum information processing.

Collaborating with amazing co-authors – Antonio, Brian, Daniel, Jens, Elham, Mina, Soumik, Sumeet, and Yihui – has been a privilege, each providing valuable insights.

Many thanks to Christophe Pere for engaging discussions on the quantum ecosystem, providing valuable feedback on my work.

I am grateful to all who provided useful feedback on my manuscripts – special mentions to Daniel Stilck França and Christoph Hirche, as well as the anonymous reviewers.

Being part of an exceptional research group spanning Paris and Edinburgh has proven to be an immensely enriching experience. My sincere thanks to my fellow colleagues, whose guidance and insights have significantly contributed to my academic growth. Special thanks to Abbas, Adriano, Alisa, Alexandru, Alex, Andrea, Bo, Brian, Chirag, Damian, Damien, Debasis, Dominik, Eleni, Eliott, Elham, Federico, Frédéric, Gizem, Grégoire, Harold, Hela, Ioannis, Ivan, James, Joannis, Jonathan, Jonas, Kaushik, Kim, Laura, Léo Colisson, Léo Monbroussou, Luka, Luís, Majid, Manon, Mario, Marine, Matilde, Matteo, Michael, Mina, Mounir, Nathan, Nicolas, Paolo, Pascal, Paul Hermouet, Paul Hilaire, Pierre-Emmanuel, Raja, Robert, Santiago, Sean, Slimane, Ulysse, Uta, Valentina, Vanessa, Verena, Victor, Yao, Yoann, and many others whom I may have inadvertently forgotten.

I extend special thanks to the residents of the *Maison des Étudiants Canadiens*, whose warmth made my first doctoral year remarkably pleasant, even in the face of numerous constraints imposed by the pandemic. Heartfelt appreciation is also due to my beloved Asmaa for the unwavering love,

support, and the joyous moments we have shared throughout our Parisian – and more recently, Montrougian – life.

I extend my gratitude to Margherita, Eugenio, Mattia, Ludovico, Roberta, and Andrea for being incredibly supportive friends during my time in Paris.

Lastly, my deepest gratitude goes to my family for their enduring support and encouragement in pursuing my aspirations and intellectual interests.

SOMMAIRE

L'informatique quantique est l'un des voyages scientifiques les plus passionnants de notre époque. Les algorithmes quantiques offrent un potentiel remarquable en promettant de résoudre plusieurs problèmes computationnels de manière exponentiellement plus rapide que leurs homologues classiques. Cependant, la mise en œuvre pratique de ces algorithmes représente un défi immense. L'objectif insaisissable d'atteindre un ordinateur quantique universel et tolérant aux erreurs reste inatteint, et la chronologie précise de son arrivée est incertaine.

Actuellement, nous assistons à l'émergence de plusieurs dispositifs quantiques à court terme. Ces dispositifs, cependant, font face à des limitations substantielles, notamment des niveaux élevés de bruit et une capacité d'intrication limitée. Bien qu'ils puissent avoir le potentiel de fournir un avantage quantique pour des tâches spécifiques, leur utilité pratique fait l'objet de débats. L'impact du bruit quantique, en particulier, soulève des questions sur leur efficacité.

Motivée par cette situation actuelle, cette thèse se plonge dans l'impact profond du bruit sur les algorithmes d'apprentissage quantique, explorant trois dimensions clés de ce problème.

Tout d'abord, elle se concentre sur l'influence du bruit sur les algorithmes quantiques variationnels, en particulier les méthodes quantiques "à noyaux". Alors que la recherche précédente considérait principalement un modèle de bruit idéalisé, tel que le bruit local de Pauli, les dispositifs quantiques du monde réel font face à un spectre plus large de bruit, y compris des composants de bruit "non-unital". Étonnamment, nos résultats révèlent des disparités marquées dans le comportement des noyaux quantiques projetés sous un bruit unital et non-unital. Sous un bruit unital, ils subissent une concentration exponentielle à une profondeur linéaire, tandis que le bruit non unital empêche une concentration exponentielle à n'importe quelle profondeur. Comme la concentration exponentielle entrave la possibilité d'entraînement, ce contraste frappant remet en question les conclusions antérieures sur les algorithmes quantiques variationnels bruyants et souligne la nécessité d'explorer des modèles de bruit plus réalistes.

Ensuite, nous abordons le problème de l'apprentissage des dynamiques quantiques avec des mesures binaires bruyantes de l'état de Choi-Jamiolkowski. À cette fin, nous adoptons le modèle précédemment défini des requêtes statistiques quantiques. Nous montrons que l'algorithme quantique Goldreich-Levin peut être mis en œuvre avec des requêtes statistiques quantiques, alors que la version antérieure de l'algorithme implique un accès oracle à l'unitaire et à son inverse. De plus, nous prouvons que les $\mathcal{O}(\log(n))$-juntas et les fonctions booléennes quantiques avec une influence totale constante sont efficacement apprenables dans notre modèle, et les circuits de profondeur constante sont apprenables de manière efficace avec des requêtes statistiques quantiques.

Enfin, nous apportons plusieurs contributions au domaine émergent de la confidentialité différentielle quantique, éclairant la manière dont le bruit quantique et classique peut être exploité pour offrir une sécurité statistique et une robustesse contre les attaques adverses. Alors que des travaux antérieurs ont proposé plusieurs extensions quantiques de la confidentialité différentielle, chacune reposant sur des notions substantiellement différentes d'états quantiques voisins, nous proposons

une définition nouvelle et générale d'états quantiques voisins. Nous démontrons que cette définition capture la structure sous-jacente des encodages quantiques et peut être utilisée pour fournir des garanties de confidentialité exponentiellement plus strictes pour les mesures quantiques. De plus, nous explorons également la confidentialité différentielle quantique dans le modèle local. Nous établissons une équivalence entre les requêtes statistiques quantiques et la confidentialité différentielle quantique dans le modèle local, étendant un résultat classique célèbre au cadre quantique. De plus, nous dérivons des inégalités de traitement des données fortes pour l'entropie relative quantique sous la confidentialité différentielle locale et appliquons ce résultat à la tâche de test d'hypothèse asymétrique avec des mesures restreintes. À titre de preuve de principe, nous démontrons que les fonctions de parité sont efficacement apprenables dans ce modèle, tandis que la tâche classique correspondante nécessite un nombre exponentiel d'échantillons.

ABSTRACT

Quantum computing is one of the most exciting scientific journeys of our times. Quantum algorithms offer remarkable potential, promising to solve several computational problems exponentially faster than their classical counterparts. However, the practical implementation of these algorithms poses an immense challenge. The elusive goal of achieving a fault-tolerant, universal quantum computer remains unattained, and the precise timeline for its arrival is uncertain.

Currently, we are witnessing the emergence of several near-term quantum devices. These devices, however, grapple with substantial limitations, including high levels of noise and limited entangling capacity. While they may hold the potential for delivering quantum advantage for specific tasks, their practical utility is a subject of debate. The impact of quantum noise, in particular, raises questions about their effectiveness.

Motivated by this current scenario, this thesis delves into the profound impact of noise on quantum learning algorithms, exploring three key dimensions of this issue.

First, it focuses on the influence of noise on variational quantum algorithms, specifically quantum kernel methods. While prior research primarily considered an idealized noise model, such as local Pauli noise, real-world quantum hardware contends with a broader spectrum of noise, including non-unital noise components. Surprisingly, our findings reveal stark disparities in the behavior of projected quantum kernels under unital and non-unital noise. Under unital noise, they incur in exponential concentration at linear depth, whereas non-unital noise prevents exponential concentration at any depth. Since exponential concentration hinders trainability, this stark contrast challenges prior findings on noisy variational quantum algorithms and underscores the necessity of exploring more realistic noise models.

Second, we consider the problem of learning quantum dynamics with noisy single-copy binary measurements of the Choi-Jamiolkowski state. To this end, we adopt the previously defined model of quantum statistical queries. We show that the quantum Goldreich-Levin algorithm can be implemented with quantum statistical queries, whereas the prior version of the algorithm involves oracle access to the unitary and its inverse. Moreover, we prove that $\mathcal{O}(\log n)$-juntas and quantum Boolean functions with constant total influence are efficiently learnable in our model, and constant-depth circuits are learnable sample-efficiently with quantum statistical queries.

Finally, we provide several contributions to the emerging field of quantum differential privacy, shedding light on how quantum and classical noise can be harnessed to provide statistical security and robustness to adversarial attacks. Whereas prior works proposed several quantum extensions of differential privacy, each of them built on substantially different notions of neighboring quantum states, we propose a novel and general definition of neighboring quantum states. We demonstrate that this definition captures the underlying structure of quantum encodings and can be used to provide exponentially tighter privacy guarantees for quantum measurements. Moreover, we also investigate quantum differential privacy in the local model. We establish an equivalence between quantum statistical queries and quantum differential privacy in the local model, extending a celebrated

classical result to the quantum setting. Furthermore, we derive strong data processing inequalities for the quantum relative entropy under local differential privacy and apply this result to the task of asymmetric hypothesis testing with restricted measurements. As a proof of principle, we demonstrate that parity functions are efficiently learnable in this model, whereas the corresponding classical task requires exponentially many samples.

CONTENTS

CHAPTER 1

INTRODUCTION

Tracing the historical roots of computer science is a challenging task. While digital computers emerged as a defining invention of the 20th century, the practice of computation by humans has an enduring history spanning thousands of years. Throughout ancient civilizations, we find evidence of step-by-step procedures for solving mathematical problems, showcasing the timeless human quest for efficient problem-solving. The very term "algorithm" finds its origins in the latinization of the last name of Muhammad ibn Musa al-Khwarizmi, a 12th-century Abbasid polymath whose systemation algebra significantly advanced the field of mathematics. This intertwined narrative of early algebra and the emergence of computer science underscores the profound connection between the two disciplines. Alan Cobham notably emphasized this association when he postulated that the complexity class of problems decidable in polynomial time served as an apt descriptor for the set of problems feasibly computable.

> *The subject of my talk is perhaps most directly indicated by simply asking two questions: first, is it harder to multiply than to add? and second, why?...I (would like to) show that there is no algorithm for multiplication computationally as simple as that for addition, and this proves something of a stumbling block.*
>
> - Alan Cobham, *The intrinsic computational difficulty of functions* [Cob64]

In the first half of the 20th century, pioneers such as Alonzo Church and Alan Turing made key contributions to the establishment of theoretical computer science. Several foundational models of computation have been proposed over time, with notable examples including Turing machines [Tur36],

lambda calculus [Chu32], and cellular automata [Tur52, VNB$^+$66]. These models have played a pivotal role in shaping the field of computer science.

Fast forward almost a century, and the world has witnessed an unprecedented transformation. Digital computers have become ubiquitous, leaving an indelible mark on our modern societies. They have revolutionized fields such as healthcare, finance, communication, and logistics, powering complex algorithms that have unlocked new frontiers in data analysis, artificial intelligence, and beyond.

To a certain degree, the present state of quantum computing resembles that of classical computer science during the 1940s. A collaborative endeavor involving physicists, computer scientists, and mathematicians is gradually unraveling the computational potential of quantum systems. However, the most promising applications remain on the horizon, awaiting practical implementation.

The birth of quantum computing is conventionally attributed to two pivotal events. Firstly, in May 1981, Richard Feynman delivered a seminal lecture titled "Simulating Physics with Computers" [F$^+$18, Pre23]. This talk popularized the concept of harnessing the computational power of quantum mechanisms for simulating quantum systems. It became evident that conventional digital computers were ill-suited for this task, as classical descriptions of quantum systems necessitated a number of variables that is exponential in the number of particles. Notably, Yuri Manin and Paul Benioff arrived at similar conclusions almost simultaneously in 1980.

The second milestone was the breakthrough achieved by Peter Shor in 1994, who introduced a polynomial-time quantum algorithm for factorizing integers – a task conjectured to require exponential time on classical computers [Sho97]. Although conceptually tantalizing, Shor's algorithm initially faced widespread skepticism regarding its practical implementation. Quantum systems are notoriously susceptible to noise, which leads to the detrimental effects of decoherence, rendering the system classically simulatable [CLSZ95]. Shor, together with Robert Calderbank, addressed these concerns with the introduction of quantum error correction [CS96], enabling the dependable execution of quantum algorithms on noisy quantum devices, granted the noise rate remains below a specific threshold. Nevertheless, the current noise rates in quantum devices surpass this threshold, thus delaying the advent of fault-tolerant quantum computers to a more distant future.

While our primary focus revolves around examining the influence of hardware noise on quantum computation, particularly within the realm of quantum learning algorithms, this thesis will not delve into error correction techniques. Instead, our scope focuses to the existing and near-future family of quantum devices.

In the recent years, theoretical advances have been accompanied by significant improvements in hardware capabilities. While the new generation of Noisy Intermediate-Scale Quantum (NISQ) devices can manipulate hundreds of physical qubits, their performance is tainted by a number of limitations, including noise, reduced entangling capacity, and limited quantum memory [Pre18a]. Given these constraints, it is not immediately clear whether these devices are of practical utility.

Recent research has revealed that noise can prevent quantum advantage for specific tasks [SFGP21,

DPMRF23, WFC+21], or even allow classical computers to efficiently sample from the output distribution of a quantum circuit measured in the computational basis [AGL+23a]. Additionally, many quantum algorithms require quantum data as input [AA23, Aar18, HKP20, RF21], which may also be corrupted by noise, making a thorough understanding of quantum noise fundamental, even in the presence of fault-tolerant quantum computers.

In this context, we consider the following overarching questions:

What is the effect of noise on near-term quantum learning algorithms? To what extent can quantum speed-ups be achieved despite the presence of noise?

While these questions possess the breadth to encompass numerous research directions, in the following we will narrow our focus to specific issues explored within this thesis. It is worth noting that, despite significant efforts, the scientific community has yet to attain a comprehensive understanding of the repercussions of noise on quantum algorithms. From the experimental side, the significant strides in quantum hardware [AAB+19, KEA+23] have been swiftly matched by advances in quantum simulation techniques [BC23], especially those based on tensor networks [PZ22]. Moreover, many theoretical results suggest that the presence of noise renders quantum circuits ineffective, even at very shallow depths [QFK+22, DPMRF23, AGL+23b].

Particularly, in the realm of variational quantum algorithms, which encompass approaches based on cost functions and quantum kernel methods, noise emerges as a primary barrier to quantum advantage. Its existence gives rise to several insurmountable challenges, including the well-known issue of barren plateaus [WFC+21]. However, we argue that prior investigations explored a rather idealized model of noise, that goes under the name of local Pauli noise, which does not account for "non-unital" perturbations , such as those present in superconductive quantum circuits [KSW20, FGG+23].

Question 1. *What is the impact of more realistic sources of noise on variational quantum algorithms?*

It could be tempting to assume that, even if there exist minor sources of noise apart from local Pauli noise, a slight modification to the model would not substantially impact the performances of variational quantum algorithms. However, we will prove that this intuition is not correct, and those "non-unital" perturbations may lead to qualitatively different scenarios.

In the context of variational quantum algorithms, noise is usually modeled as a series of local channels interspersing the layers of a quantum circuit. On the other hand, it's equally crucial to investigate the influence of noise on quantum measurements, and address additional limitations, such as the absence of quantum memory and the limited entangling capacity. These constraints find their unifying framework in the model of "quantum statistical queries" [AGY20] . Intriguingly, quantum statistical queries exhibit a significantly higher degree of computational power compared to their classical counterparts. For instance, quantum statistical queries enable the efficient learning of parity functions from uniform quantum examples, whereas their classical counterparts are believed to entail an exponentially larger sample complexity. While earlier research on quantum statistical

queries has predominantly focused on learning quantum states, we take a step forward by raising the following inquiry.

Question 2. *Can we employ quantum statistical queries to learn quantum dynamics?*

Shifting the focus, a vast body of literature suggests that noise can offer notable benefits for specific computational tasks. Particularly, noise holds the potential to ensure diverse notions of statistical security, thereby enhancing adversarial robustness and generalization in various settings [CMS11a, DF18, CRK19, LAG+19]. In this context, the comprehensive framework of differential privacy emerges as a unifying approach for understanding the role of noise in machine learning and statistics [DMNS06, DR14, CDE+23]. Driven by these insights, we pose the following question.

Question 3. *Can we leverage quantum noise to guarantee properties like differential privacy and robustness to adversarial attacks?*

Differential privacy comes in several flavors. Particularly, a distinction arises between standard differential privacy and local differential privacy. Notably, the latter model offers a more robust notion of security, as it treats even the curator (i.e., the analyst who accesses the raw input data) as untrusted [KLN+11a, DJW13, AAC21a]. On the other hand, the local model entails the injection of a considerable amount of noise, which may hinder the computational power of quantum algorithms, motivating the following question.

Question 4. *Can we attain an exponential quantum speed-up under the stringent constraint of local differential privacy?*

1.1 Summary of contents

We offer an overview of the thesis's contributions, driven by the consideration of the four compelling questions mentioned earlier. In addition to presenting the main results, we provide the reader with crucial background information on quantum computing, with a focus on models of quantum noise and differential privacy.

Introduction to quantum computing on noisy devices. This thesis commences with a gentle introduction to the fundamental concepts of quantum computing in Chapter 2. This introduction is designed to be accessible to a broad audience with no prior background in physics. Following this, in Chapter 3, we delve into essential quantum information materials. This includes discussions on quantum states and channels, information-theoretic divergences, and the Haar measure.

Chapter 4 is dedicated to exploring the impact of noise in quantum circuits. We examine previous findings on the decay of purity, both within circuits interspersed by local noise and in the context of unitary sampled from a 2-design followed by an arbitrary noise channel. Additionally, we offer an alternative proof for the exponential concentration of the cost function induced by unital noise,

previously demonstrated in [WFC+21]. Furthermore, we present a novel result concerning the contraction coefficient of any local channel in relation to the quantum Wasserstein distance of order 1. In this analysis, we unify the actions of such a channel and a random local unitary, expanding upon the prior worst-case analysis conducted in [DPMTL21a].

Exponential concentration and lack thereof in quantum kernel methods. Chapter 5 is devoted to the limitations of variational quantum algorithms on noisy near-term devices. In particular, we explore the phenomenon of noise-induced concentration [WFC+21, TWH22]. First, we relate the exponential concentration of the fidelity quantum kernels to the purity decay, providing an exponentially tighter lower bounds on their sample complexity under both unital and non-unital noise. Second, we discuss the impact of non-unital noise on another family of quantum kernel methods, namely the projected quantum kernels. Surprisingly, projected quantum kernels do not exhibit exponential concentration under non-unital noise. However, we argue that this phenomenon does not imply the trainability of the entire circuit, but solely of its final layers. Thus, we conjecture that the early layers of a super-logarithmic depth circuit bear little influence on the final output, and we prove this statement in the high-noise regime, hinging on novel techniques based on the contraction of the quantum Wasserstein distance of order 1 [DPMTL21a].

Learning unitaries with quantum statistical queries. In Chapter 6, we propose a model for learning unitary operators from quantum statistical queries (QSQs) with respect to their Choi-Jamiolkowski state. Our model is a natural extension of a previous model for learning classical Boolean functions from quantum statistical queries with respect to quantum examples [AGY20, AHS23]. Quantum statistical queries capture the capabilities of a learner with limited quantum resources, which receives as input only noisy estimates of expected values of measurements. Particularly, we prove that quantum $\mathcal{O}(\log n)$-juntas, quantum Boolean functions with constant total influence and constant-depth circuits are efficiently learnable in our model, while previous algorithms required direct access to the Choi-Jamiolkowski state or oracle access to the unitary and its inverse. We also demonstrate that, despite these positive results, quantum statistical queries lead to an exponentially larger sample complexity for certain tasks, compared to separable measurements to the Choi-Jamiolkowski state.

Background on differential privacy. Moving on to Chapter 7, we provide an overview of differential privacy and its interaction with competing privacy-preserving techniques. We review the mathematical foundations of differential privacy in both the standard and local models, as well as prior research in the domain of quantum differential privacy. Then, we introduce the concept of "neighboring-preserving quantum encodings" and showcase how quantum differential privacy can be employed to safeguard the privacy of the underlying classical input data, particularly within the context of hybrid classical-quantum algorithms. Finally, we explore the connections between privacy, robustness to adversarial attacks, and generalization in machine learning.

A unifying framework for differentially private quantum algorithms. In Chapter 8 we revisit the notion of quantum differential privacy [ZY17a, AR19] and demonstrate that quantum noise can enhance the privacy of classical data embedded in quantum states. To this end, we give a novel and general definition of neighbouring quantum states. We demonstrate that this definition captures the underlying structure of quantum encodings and can be used to provide exponentially tighter privacy guarantees for quantum measurements. Our approach exploits both classical and quantum noise and is motivated by the noisy nature of near-term quantum devices. Finally, we complement our theoretical findings with an empirical estimation of the certified adversarial robustness ensured by differentially private measurements. Our results hinges on a novel result on quantum divergences, namely the advanced joint convexity of the quantum hockey-stick divergence, whose proof is delayed to Chapter 11.

Quantum differential privacy in the local model. Chapter 9 focuses on quantum differential privacy in the local model. We establish an equivalence between quantum statistical queries and quantum differential privacy in the local model, extending a celebrated classical result to the quantum setting [KLN$^+$11b]. Furthermore, we derive strong data processing inequalities for the quantum relative entropy under local differential privacy and apply this result to the task of asymmetric hypothesis testing with restricted measurements. Finally, we consider the task of quantum multi-party computation under local differential privacy. As a proof of principle, we demonstrate that parity functions are efficiently learnable in this model, whereas the corresponding classical task requires exponentially many samples.

In the concluding Chapter 10, we present a series of open questions that emerge from our findings, with the aspiration of igniting future research endeavors.

1.2 Additional remarks

The thesis is based on the following articles.

- [ADK23] – **Armando Angrisani**, Mina Doosti and Elham Kashefi. "A unifying framework for differentially private quantum algorithms." arXiv preprint arXiv:2307.04733
  (previous version: [ADK22] – **Armando Angrisani**, Mina Doosti and Elham Kashefi. "Differential privacy amplification in quantum and quantum-inspired algorithms" arXiv:2203.03604)
  This work overlaps with Chapters 7 and 8.

- [AK22a] – **Armando Angrisani** and Elham Kashefi. "Quantum local differential privacy and quantum statistical query model." arXiv preprint arXiv:2203.03591
  This work overlaps with Chapter 9.

- [Ang23] – **Armando Angrisani**. “Learning unitaries from quantum statistical queries.” arXiv preprint arXiv:2310.02254
  This work overlaps with Chapter 6.

- [MAE+23] – Antonio Anna Mele, **Armando Angrisani**, Jens Eisert, Soumik Ghosh, Sumeet Khatri, Yihui Quek and Daniel Stilck França. “Noise-induced absence of barren plateaus: Non-unital noise can be a friendly foe.”
  This work overlaps with Chapters 4 and 5.

Armando Angrisani is the leading author and main contributor of [AK22a] and [ADK23]. Armando Angrisani contributed to [MAE+23] by conceiving and proving results concerning quantum kernels and by proving results concerning the limitations of variational quantum algorithms in the the high-noise regime, supporting the “effective depth” picture.

The following first-author article is excluded from the present thesis.

- [ACK21] – **Armando Angrisani**, Brian Coyle, and Elham Kashefi. “Probably approximately correct quantum source coding.” arXiv preprint arXiv:2112.06841.

CHAPTER 2

A GENTLE START TO QUANTUM COMPUTING

Quantum mechanics is a beautiful generalization of the laws of probability: a generalization based on the 2-norm rather than the 1-norm, and on complex numbers rather than nonnegative real numbers. It can be studied completely separately from its applications to physics (and indeed, doing so provides a good starting point for learning the physical applications later). This generalized probability theory leads naturally to a new model of computation – the quantum computing model – that challenges ideas about computation once considered a priori, and that theoretical computer scientists might have been driven to invent for their own purposes, even if there were no relation to physics. In short, while quantum mechanics was invented a century ago to solve technical problems in physics, today it can be fruitfully explained from an extremely different perspective: as part of the history of ideas, in math, logic, computation, and philosophy, about the limits of the knowable.

-Scott Aaronson, *Quantum Computing Since Democritus*

This thesis delves into three core domains of computer science: quantum computing, learning theory, and differential privacy. Here, we present a concise introduction to quantum com-

puting, designed with a computer science audience in mind. Our objective is to facilitate the connection between the classical and quantum communities.

In this context, it is possible to introduce quantum computation and information without making direct reference to physics, as demonstrated, for example, in [Wat18]. However, we contend that achieving a balance between physical intuition and mathematical tools is a more desirable approach. Indeed, a grasp of the following foundational physical concepts is sufficient to comprehend the workings of quantum algorithms.

1. **Quantization of physical parameters**: some physical attributes, such as the energy and momentum of elementary particles, exhibit quantization, meaning they can only assume values from a discrete set. Consider the model of the atom illustrated in Figure 2.1. When electrons are measured, they can be found only in a finite set of orbits.

2. **Superposition**: the state of a quantized physical parameter can be expressed through a concept known as a "quantum superposition". In the context of electronic levels within an atom, this notion corresponds to the idea that an electron can "exist in multiple electronic states simultaneously", each state having a specific amplitude. However, it is important to note that this intuition, while helpful, lacks precision; expressing it rigorously requires the formalism of linear algebra and complex numbers.

3. **Probabilistic nature of quantum states**: the amplitude of a quantum superposition comes with an associated probability distribution, which we denote as $p(\cdot)$. Prior to measurement, the electron's state exists in a superposition, spanning the possible orbits. Upon measurement, the electron collapses into the i-th orbit with a probability of $p(i)$.

The detection of these intriguing effects might not be immediately apparent. The reader might wonder why we should be concerned if an electron is in a superposition of states, as long as we ultimately observe it in a single state from a finite set.

However, it is crucial to recognize that classical and quantum states evolve differently over time. Ignoring quantum effects can lead to seemingly paradoxical phenomena. The celebrated double-slit experiment is a prime illustration of wave-particle duality: photons behave as waves when unobserved, yet exhibit particle-like behavior upon measurement. This peculiar behavior underscores the significance of understanding quantum effects in various physical phenomena.

2.1 Qubits

The foundation of quantum computing lies in qubits, which are the quantum counterparts of classical bits. Consider a 2-dimensional Hilbert space $\mathbb{C}^2$ over the field $\mathbb{C}$. In this space, we have a physical system capable of assuming two mutually exclusive classical states, such as those illustrated in Figure 2.1. These states can be represented by the orthonormal vectors $|0\rangle := (1,0)^\intercal$ and $|1\rangle := (0,1)^\intercal$. This pair, $\{|0\rangle, |1\rangle\}$, serves as a basis for $\mathbb{C}^2$ and is commonly referred to as the "computational basis".

While a classical bit can only be in either state $|0\rangle$ or $|1\rangle$, a qubit can exist in any normalized complex superposition of $|0\rangle$ and $|1\rangle$. This superposition can be expressed as:

$$|\psi\rangle := \alpha|0\rangle + \beta|1\rangle = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}, \tag{2.1}$$

where α and β are complex numbers, and $|\alpha|^2 + |\beta|^2 = 1$.

In the context that follows, we will refer to vectors of the form $\alpha|0\rangle + \beta|1\rangle$ as "quantum states" or simply "states". As the normalization condition holds, a qubit's state induces a probability distribution, denoted as $p : \{|0\rangle, |1\rangle\} \to [0,1]$, where:

$$p(|0\rangle) = |\alpha|^2$$

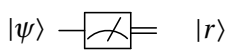
and

$$p(|1\rangle) = |\beta|^2.$$

2.2 Measurements

Compared to classical bits, qubits contain a wealth of information, represented by two real parameters. A state like $\alpha|0\rangle + \beta|1\rangle$ is described by two complex numbers, which correspond to four real numbers. Importantly, the normalization is fixed, and the external phase does not produce any physical effect, i.e., $|\psi\rangle = e^{i\alpha}|\psi\rangle$ for any $\alpha \in \mathbb{R}$. Consequently, two real parameters are sufficient to describe a qubit.

However, this information is not directly accessible. Given a qubit in the state $\alpha|0\rangle + \beta|1\rangle$, we can perform a *measurement* in the computational basis. As a result, the qubit *collapses* to $|0\rangle$ with probability $|\alpha|^2$ and to $|1\rangle$ with probability $|\beta|^2$, in accordance with Born's rule. Should we measure the qubit again, it will be found in the same state. The measurement process is depicted in the circuit below:



Where

$$|r\rangle = \begin{cases} |0\rangle & \text{with probability } |\alpha|^2 \\ |1\rangle & \text{with probability } |\beta|^2 \end{cases}$$

As depicted in Figure 2.1, an electron orbiting an atom can represent a physical system with two classical states. When measured, the electron can be found either in the "ground" state $|0\rangle$ or the "excited" state $|1\rangle$. However, quantum mechanics allows intermediate states, such as $|+\rangle := (|0\rangle + |1\rangle)/\sqrt{2}$, to exist as well. Furthermore, quantum mechanics permits a broader class of measurements known as "positive operator-valued mesure" (POVM) measurements, which we define in Chapter 3.

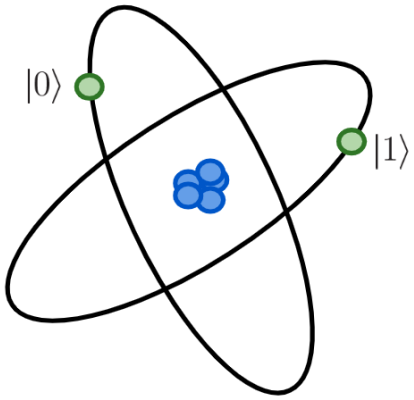


Figure 2.1: An atom with 2 electronic orbits, denoted as the states $|0\rangle$ and $|1\rangle$. This 2-level system is a physical realization of the qubit.

2.3 Multipartite Systems

Now, let's consider a system consisting of two distinct qubits. Each qubit resides in its respective Hilbert space, denoted as $\mathscr{H}^{(1)}$ for the first qubit and $\mathscr{H}^{(2)}$ for the second. These spaces can be combined using the tensor product operation, resulting in a new Hilbert space $\mathscr{H}^{(1)} \otimes \mathscr{H}^{(2)}$. The elements of $\mathscr{H}^{(1)} \otimes \mathscr{H}^{(2)}$ are superpositions in the form:

$$\phi_1 |00\rangle + \phi_2 |01\rangle + \phi_3 |10\rangle + \phi_4 |11\rangle ,$$

where $\forall i : \phi_i \in \mathbb{C}$ and $\sum_i |\phi_i|^2 = 1$. Similarly, the tensor product of two quantum states, $\alpha_1 |0\rangle + \beta_1 |1\rangle \in \mathscr{H}^{(1)}$ and $\alpha_2 |0\rangle + \beta_2 |1\rangle \in \mathscr{H}^{(2)}$, is:

$$(\alpha_1 |0\rangle + \beta_1 |1\rangle) \otimes (\alpha_2 |0\rangle + \beta_2 |1\rangle) = \alpha_1\alpha_2 |00\rangle + \alpha_1\beta_2 |01\rangle + \beta_1\alpha_2 |10\rangle + \beta_1\beta_2 |11\rangle .$$

Notably, there are elements in $\mathscr{H}^{(1)} \otimes \mathscr{H}^{(2)}$ that cannot be decomposed into tensor products of states in $\mathscr{H}^{(1)}$ and $\mathscr{H}^{(2)}$. These non-decomposable states are called "entangled" and exhibit a "non-local" behavior, which is a distinctively quantum phenomenon.

Suppose two parties, Alice and Bob, share an entangled state, such as the Einstein-Podolsky-Rosen (EPR) pair:

$$\frac{1}{\sqrt{2}} |00\rangle + \frac{1}{\sqrt{2}} |11\rangle .$$

If Alice possesses the first qubit, and Bob holds the second one, measuring Alice's qubit and obtaining $|1\rangle$ results in the collapse of Bob's qubit to $|1\rangle$, even if Alice and Bob are arbitrarily distant from each other. This effect is considered non-local. However, Alice's measurement cannot transmit any information to Bob, as dictated by the "No-communication theorem". Nonetheless, sharing an EPR pair can offer advantages for various problems, with the most famous being the "CHSH game",

introduced in [CHTW04]. In this game, Alice and Bob are given input bits x and y, and their objective is to output bits a and b such that:

$$a \oplus b \tag{2.2}$$

In the absence of communication and without any prior entanglement between the parties, their chances of winning are limited to a maximum probability of just 3/4. However, when they do share an EPR pair, their winning probability significantly improves, reaching approximately $\cos(\pi/8)^2 \simeq 0.85$.

2.4 Quantum gates

Quantum gates, the quantum counterparts of classical logic gates, play a pivotal role in quantum computing. A quantum gate, denoted as $\mathscr{U}$, is responsible for defining a unitary transformation that operates on a set of d qubits. This transformation is described as:

$$\mathscr{U} : \mathscr{H}^{(1)} \otimes \cdots \otimes \mathscr{H}^{(d)} \rightarrow \mathscr{H}^{(1)} \otimes \cdots \otimes \mathscr{H}^{(d)}$$

Typically, d is a small number, most commonly 1, 2, or 3. The fundamental property of unitarity ensures that quantum gates preserve the normalization of quantum states.

In the realm of quantum mechanics, only linear operations are permissible when manipulating quantum states. These unitary transformations are accurately depicted using unitary matrices. It's important to note that the inverse of a unitary matrix, denoted as U^{-1}, corresponds to its conjugate matrix, denoted by $U^{\dagger}$. As a result, quantum gates are reversible operations.

In stark contrast to this, the only irreversible operation in quantum computing is the measurement. Consequently, there exist classical gates, such as the AND and OR gates, which lack direct quantum counterparts due to their inherently irreversible nature.

We provide here some examples of single-qubit quantum gates. I, X, Y and Z are known as the Pauli gates, while H is the Hadamard gate.

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

We introduce the following 2-qubit gates, known as Controlled Z (CZ) and Controlled X (CX).

$$CZ = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \qquad CX = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

The CX gate, also called *control-NOT* is usually denoted with the following symbol.

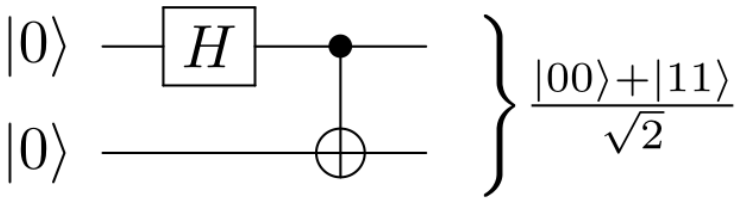


Figure 2.2: Quantum circuit to create the EPR pair.

In general, a controlled unitary CU acts as follows: it leaves unchanged the first qubit; if the first qubit is $|0\rangle$, it leaves unchanged also the second qubit, otherwise it applies U to the second qubit.

We introduce as well the *phase shifter* gate P,

$$P(\alpha) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{bmatrix}.$$

The parametric gates presented below play a variational quantum algorithms.

$$R_z(\alpha) = \begin{bmatrix} e^{-i\frac{\alpha}{2}} & 0 \\ 0 & e^{i\frac{\alpha}{2}} \end{bmatrix} \qquad R_x(\alpha) = \begin{bmatrix} \cos\left(\frac{\alpha}{2}\right) & -i\sin\left(\frac{\alpha}{2}\right) \\ -i\sin\left(\frac{\alpha}{2}\right) & \cos\left(\frac{\alpha}{2}\right) \end{bmatrix}$$

$$CZ(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-i\frac{\alpha}{2}} & 0 \\ 0 & 0 & 0 & e^{i\frac{\alpha}{2}} \end{bmatrix} \qquad CX(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\left(\frac{\alpha}{2}\right) & -i\sin\left(\frac{\alpha}{2}\right) \\ 0 & 0 & -i\sin\left(\frac{\alpha}{2}\right) & \cos\left(\frac{\alpha}{2}\right) \end{bmatrix}$$

Observe that $iR_x(\pi) = X$, $iR_z(\pi) = Z$, $iCX(\pi) = CX$ and $iCZ(\pi) = CZ$.

2.5 The circuit model

A quantum algorithm is usually described by a quantum circuit, that is a finite directed acyclic graph whose nodes are either input/output nodes or *quantum gates*. As an illustrative example, we present in Figure 2.2 the circuit producing the EPR pair.

Notably, the extensive array of existing quantum gates need not be exhaustively considered. The Solovay-Kitaev Theorem ([CLSZ95], Appendix 3) provides a crucial insight, indicating that our focus can be narrowed down to a small set of gates. Specifically, it implies the universality of the gate set $\mathcal{G} = \{CX, H, P(\pi/8)\}$ for quantum computation. This universality guarantees that any quantum circuit can be approximated by another circuit using gates from $\mathcal{G}$ with only a logarithmic slowdown.

CHAPTER 3

FOUNDATIONS OF QUANTUM INFORMATION THEORY

Quantum information science has experienced a rapid development over the last three decades. A comprehensive introduction to this discipline would undoubtedly exceeds the space of this chapter. Given this fact, we cover selected topics and tools that are used during this thesis, in order to provide a self-contained exposition. For an extensive treatment of the subject, we refer instead to [NC10, Wil13, Wat18].

3.1 Preliminaries

We start by introducing the mathematical notation. For a vector $\boldsymbol{x} = (x_1, \ldots, x_n)$, we denote as $\|\boldsymbol{x}\|_p$ its p-norm, where $\|\boldsymbol{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $1 \leq p < \infty$ and $\|\boldsymbol{x}\|_\infty = \max_i \|x_i\|$. It is convenient to introduce also the 0-norm (which is technically not a norm): $\|\boldsymbol{x}\|_0 = |\{i : x_i \neq 0\}|$, which is the number of the non-zero entries of $\boldsymbol{x}$. For $n \geq 1$, we will write $[n] = \{1, 2, \ldots, n\}$. Given $T \subseteq [n]$, we will write $\overline{T} := [n] \setminus T$. We will denote the $2^n \times 2^n$ identity matrix as I_n and we may omit the index n when is clear from the context. For a matrix A, we will denote as $A[i, j]$ or A_{ij} the entry corresponding to the i-th row and the j-th column. We will use the indicator string $\mathscr{S} = (x_1, x_2, \ldots, x_k, *, * \ldots, *)$ to denote the set of n-bit strings whose first k elements are $x_1, x_2, \ldots, x_k$, i.e. $\mathscr{S} = \{(t_1, t_2, \ldots, t_n) | \ \forall i \in [k] : x_i = t_i)\}$. Given a random variable X sampled according to a distribution ν, we will denote by $\mathbb{E}_\nu[X]$ its expected value and its variance by $\mathbb{V}_\nu[X]$, and omit the index ν when it's clear from the context.

For two probability distributions P, Q over a domain $\mathscr{X}$, we denote their total variation distance as $|P-Q|_{\mathrm{tv}} = \frac{1}{2}\sum_{x\in\mathscr{X}} |P(x)-Q(x)|$.

Basic definitions. Let $\{|0\rangle, |1\rangle\}$ be the canonical basis of $\mathbb{C}^2$, and $\mathscr{H}_n = (\mathbb{C}^2)^{\otimes n}$ be the Hilbert space of n qubits. For $x = x_1 x_2 \dots x_n \in \{0,1\}^n$, we denote the computational state $|x\rangle = \bigotimes_{i=1}^n |x_i\rangle$ and, in particular, we write $|0^n\rangle = |00\dots0\rangle$. We use the bra-ket notation, where we denote a vector $v \in (\mathbb{C}^2)^{\otimes n}$ using the ket notation $|v\rangle$ and its adjoint using the bra notation $\langle v|$. For $u, v \in \mathscr{H}_n$, we will denote by $\langle u|v\rangle$ the standard Hermitian inner product $u^\dagger v$. A quantum (pure) state is a normalized vector $|v\rangle$, i.e. $|\langle v|v\rangle| = 1$. Let $\mathscr{L}_n$ be the subset of linear operators on $\mathscr{H}_n$, with I representing the identity operator, and let $\mathscr{O}_n \subset \mathscr{L}_n$ be the subset of self-adjoint linear operators on $\mathscr{H}_n$. We denote by $\mathscr{O}_n^T \subset \mathscr{O}_n$ be the subset of traceless self-adjoint linear operators on $\mathscr{H}_n$, by $\mathscr{O}_n^+ \subset \mathscr{O}_n$ the subset of the positive semidefinite linear operators on on $\mathscr{H}_n$ and by by $\mathscr{S}_n \subset \mathscr{O}_n^+$ the set of the quantum states of $\mathscr{H}_n$, i.e. $\mathscr{S}_n := \{\rho \in \mathscr{L}_n : \rho \geq 0, \mathrm{Tr}[\rho] = 1\}$. We denote by $\mathscr{U}_n$ the unitary group, that is the set linear operators $U \in \mathscr{L}_n$ satisfying $UU^\dagger = U^\dagger U = I$, and we denote by $\mathrm{Id} : \mathscr{L}_n \to \mathscr{L}_n$ the identity map. For any operators $A, B \in \mathscr{L}_n$, let $\langle A, B\rangle$, denote the normalized Hilbert-Schmidt inner product,

$$\langle A, B\rangle = \frac{1}{2^n}\mathrm{Tr}\left[A^\dagger B\right] = \frac{1}{2^n}\sum_{i,j\in\{0,1\}^n} A^*_{i,j} B_{i,j}. \tag{3.1}$$

We define the canonical maximally entangled state as $|\Omega\rangle = \frac{1}{\sqrt{2^n}}\sum_{i,j\in\{0,1\}^n} |i,i\rangle$.

Operators on tensor spaces. We also introduce a further notation for tensor products of k Hilbert spaces. We define by $\mathscr{L}_n^k$ be the subset of linear operators on $\mathscr{H}_n^{\otimes k}$. In particular, for $k = 2$, the *identity* $\mathbb{I}$ and the *Flip operator* $\mathbb{F}$ associated to a tensor product of two Hilbert spaces are defined as

$$\mathbb{I} := \sum_{i,j\in\{0,1\}^n} |i,j\rangle\langle i,j|, \qquad \mathbb{F} := \sum_{i,j\in\{0,1\}^n} |i,j\rangle\langle j,i|. \tag{3.2}$$

Notably, they satisfy the following properties:

$$\mathbb{I}\left(|\psi\rangle\otimes|\phi\rangle\right) = |\psi\rangle\otimes|\phi\rangle, \qquad \mathbb{F}\left(|\psi\rangle\otimes|\phi\rangle\right) = |\phi\rangle\otimes|\psi\rangle, \tag{3.3}$$

for all $|\psi\rangle, |\phi\rangle \in \mathscr{H}_n$.

Let X_{AB} be an operator acting on a tensor product Hilbert space $\mathscr{H}_A \otimes \mathscr{H}_B$, and let $\{|l\rangle_B\}$ be an orthonormal basis for $\mathscr{H}_B$. Then the partial trace over the Hilbert space $\mathscr{H}_B$ is defined as follows:

$$\mathrm{Tr}_B[X_{AB}] = \sum_l (I_A \otimes \langle l|_B) X_{AB} (I_A \otimes |l\rangle_B). \tag{3.4}$$

Measurements. The most general class of measurements that we can perform on mixed states are the POVM (Positive Operator Valued Measure) measurements. Although they can be represented as channels, it is convenient to define them separately. In the POVM formalism, a measurement $\mathscr{M}$ is

given by a list of $d \times d$ positive semidefinite matrices $(\mathscr{M}_1, \ldots, \mathscr{M}_k)$, which satisfy $\sum_{i=1}^{k} \mathscr{M}_i = I$. Each $\mathscr{M}_i$ is called POVM element. The measurement rule is:

$$\Pr[\mathscr{M} \text{ returns outcome } i \text{ on input } \rho] = \mathrm{Tr}(\mathscr{M}_i \rho).$$

We denote as $\mathscr{M}(\rho)$ the distribution over $[k]$ induced by performing $\mathscr{M}$ on the state ρ. Thus we have $\mathbb{E}[\mathscr{M}(\rho)] = \sum_{i=1}^{k} i \cdot \mathrm{Tr}(\mathscr{M}_i \rho)$. Moreover, we denote by $\mathsf{range}(M)$ the set of possible outcomes of M.

3.2 Ensembles of states and unitaries

In this section, we will introduce ensembles of states and unitaries of particular interest. Throughout this thesis, we will use the terms "distribution" and "ensemble" interchangeably. We will sometimes define an ensemble as a set of parametrized unitaries, for instance $\{U(\boldsymbol{\theta})\}_{\boldsymbol{\theta} \in \Theta}$. In this case, the associated distribution is the one obtained by sampling $\boldsymbol{\theta}$ uniformly at random from Θ.

3.2.1 Haar measure and t-designs

We start by providing some rudimentary notions about the Haar measure μ_n, which can be thought as the uniform distribution over the unitary group $\mathscr{U}_n$. For a comprehensive introduction to the Haar measure and its properties, we refer to [Mel23].

Definition 3.1 (Haar measure)**.** The Haar measure on the unitary group $\mathscr{U}_n$ is the unique probability measure μ_n that is both left and right invariant over the set $\mathscr{U}_n$, i.e., for all integrable functions f and for all $V \in \mathscr{U}_n$, we have:

$$\int_{\mathscr{U}_n} f(U) d\mu_n(U) = \int_{\mathscr{U}_n} f(UV) d\mu_n(U) = \int_{\mathscr{U}_n} f(VU) d\mu_n(U). \tag{3.5}$$

Given a state $|\phi\rangle$, we denote the k-th moment of a Haar random state as

$$\mathbb{E}_{|\psi\rangle \sim \mu_n} \left[|\psi\rangle \langle\psi|^{\otimes k} \right] := \mathbb{E}_{U \sim \mu_n} \left[U^{\otimes k} |\phi\rangle \langle\phi|^{\otimes k} U^{\dagger \otimes k} \right]. \tag{3.6}$$

Note that the right invariance of the Haar measure implies that the definition of $\mathbb{E}_{|\psi\rangle \sim \mu_n} \left[|\psi\rangle \langle\psi|^{\otimes k} \right]$ does not depend on the choice of $|\phi\rangle$.

In numerous scenarios, random unitaries and states are drawn from distributions that effectively capture solely the lower-order statistical properties of the Haar measure. A prime example of this is the exploration of the well-documented barren plateaus phenomenon [MBS+18]. This naturally brings us to the concept of (unitary) k-designs, for integers $k \geq 1$ [DCEL09].

Definition 3.2 (Unitary k-design)**.** Let ν be a probability distribution over the unitary group $\mathscr{U}_n$. The distribution ν is unitary k-design if

$$\mathbb{E}_{V \sim \nu} \left[V^{\otimes k} X V^{\dagger \otimes k} \right] = \mathbb{E}_{V \sim \mu_n} \left[V^{\otimes k} X V^{\dagger \otimes k} \right], \tag{3.7}$$

for all linear operator $X \in \mathscr{L}_n^k$.

Informally, we say that ν and μ_n "agree" up to the first k-moments. State k-designs are defined analogously.

Definition 3.3 (State k-design)**.** Let ν be a probability distribution over the set of quantum states $\mathscr{S}_n$. The distribution ν is said to be a state k-design if

$$\mathbb{E}_{|\psi\rangle\sim\nu}\left[|\psi\rangle\langle\psi|^{\otimes k}\right] = \mathbb{E}_{|\psi\rangle\sim\mu_n}\left[|\psi\rangle\langle\psi|^{\otimes k}\right]. \tag{3.8}$$

We will now give the expressions of the first two moments of the Haar measure in terms of the identity and swap operator. Given $X \in \mathscr{L}_n$, we have

$$\mathbb{E}_{U\sim\mu_n}\left[UXU^\dagger\right] = \mathrm{Tr}[X]\frac{I}{2^n}. \tag{3.9}$$

Given $X \in \mathscr{L}_n^2$, we have

$$\mathbb{E}_{U\sim\mu_n}\left[U^{\otimes 2}XU^{\dagger\otimes 2}\right] = \frac{\mathrm{Tr}[X] - 2^{-n}\mathrm{Tr}[\mathbb{F}X]}{2^{2n}-1}\mathbb{I} + \frac{\mathrm{Tr}[\mathbb{F}X] - 2^{-n}\mathrm{Tr}[X]}{2^{2n}-1}\mathbb{F}. \tag{3.10}$$

The first two moments of the Haar measure occur in many calculations involving random states sampled from a 2-design, particularly in Chapters 4, 5 and 6 of this thesis.

Example 3.1. The Pauli group $\mathscr{P}_n$ forms a 1-design. This can be checked by expanding an arbitrary operator $X \in \mathscr{L}_n$ in the Pauli basis:

$$\frac{1}{4^n}\sum_{P\in\mathscr{P}_n} PXP^\dagger = \frac{1}{8^n}\sum_{Q\in\mathscr{P}_n}\sum_{P\in\mathscr{P}_n} PQP^\dagger\,\mathrm{Tr}[XQ] = \mathrm{Tr}[X]\frac{I}{2^n}, \tag{3.11}$$

where we used the fact the all non-identity Pauli strings commute with half Pauli strings and anti-commutes with the other halfs, therefore $\sum_{P\in\mathscr{P}_n} PQP^\dagger = 4^n$ if $Q = I$ and 0 otherwise.

Example 3.2 (Overlap of random states)**.** From Equation 3.9, we can immediately compute the overlap between two random states. Let ν be a state 1-design. We have,

$$\mathbb{E}_{\rho\sim\nu,\sigma\sim\nu}\mathrm{Tr}[\rho\sigma] = \mathrm{Tr}\left[\left(\frac{I}{2^n}\right)^2\right] = \frac{1}{2^n}. \tag{3.12}$$

3.2.2 Locally scrambled ensembles

Along with t-designs, another important family of unitaries (and states) is the one of *locally scrambled ensembles*, introduced in [CHE+23].

Definition 3.4 (Locally scrambled ensembles)**.** An ensemble of n-qubit unitaries is called locally scrambled if it is invariant under pre-processing by tensor products of arbitrary local unitaries. That is, a unitary ensemble $\mathscr{U}_{\mathrm{LS}}$ is locally scrambled if for $U \sim \mathscr{U}_{\mathrm{LS}}$ and for any fixed $U_1,\ldots,U_n \in \mathscr{U}_1$ also $U(\bigotimes_{i=1}^n U_i) \sim \mathscr{U}_{\mathrm{LS}}$. Accordingly, an ensemble $\mathscr{S}_{\mathrm{LS}}$ of n-qubit quantum states is locally scrambled if it is of the form $\mathscr{S}_{\mathrm{LS}} = \mathscr{U}_{\mathrm{LS}}|0^n\rangle$ for some locally scrambled unitary ensemble $\mathscr{U}_{\mathrm{LS}}$.

Notable examples of locally scrambled ensembles are the products of random single-qubit stabilizer states and the products of Haar random k-qubit states, which, in particular, include Haar random n-qubit states the products of Haar random single-qubit states. We emphasize that the above families include both product states and highly entangled states. This definition is motivated by a phenomenon referred as *out-of-distribution* generalization. First, we briefly discuss the intuition behind *in-distribution* generalization. In most learning tasks, an agent, or learner, is provided some data sampled from a distribution $\mathscr{P}$ used during the a *training* phase, and subsequently tested according the same distribution $\mathscr{P}$ during the *testing* phase. This corresponds to the intuition that a fair examination of a student should adhere to the materials she encountered during the course, rather than covering a totally unrelated topic. Thus, if the testing distribution is $\mathscr{Q} \neq \mathscr{P}$, providing an accurate learning algorithm often becomes an insurmountable task. In this scenario, say that a learner that produces an accurate prediction has achieved *out-of-distribution* generalization. Instances of this problems have been addressed in (classical) machine learning with "transfer learning" techniques [PY10]. In the quantum setting, *out-of-distribution* generalization is achievable when $\mathscr{P}, \mathscr{Q}$ are locally scrambled distribution over states. To state this result, we first need to introduce the following notion of expected risk, also employed in Chapter 6. For $U, V \in \mathscr{U}_n$, we define

$$\mathscr{R}_\nu(U,V) := \mathbb{E}_{|\psi\rangle\sim\nu}\left[\left\| U|\psi\rangle\langle\psi|U^\dagger - V|\psi\rangle\langle\psi|V^\dagger \right\|_{\mathrm{tr}}^2\right], \tag{3.13}$$

where we can think U as an unknown unitary, V as the unitary output by the learner and ν as the testing distribution. Then the goal of the learner is to output the unitary V minimizing the expected risk. Surprisingly, the expected risks with respect to all locally scrambled ensembles are within a constant multiplicative factor.

Lemma 3.1 ([CHE+23], Lemma 1)**.** *For any $\nu \in \mathscr{S}_{\mathrm{LS}}$ and $U, V \in \mathscr{U}_n$,*

$$\frac{1}{2}\mathscr{R}_{\mu_n}(U,V) \leq \frac{2^n}{2^n+1}\mathscr{R}_\nu(U,V) \leq \mathscr{R}_{\mu_n}(U,V). \tag{3.14}$$

As also argued in [CHE+23], this result holds for a larger family of ensemble agrees with a locally scrambled one up to and including its (complex) second moments.

3.2.3 Collision probability and anticoncentration

Given an ensemble of states ν, it is fruitful to look at the distribution induced by a computational measurement. In particular, we define the *scaled* collision probability of ν as

$$\mathscr{Z}(\nu) := 2^n \cdot \mathbb{E}_{\rho\sim\nu}\left[\sum_{x\in\{0,1\}^n} \mathrm{Tr}[\rho|x\rangle\langle x|]^2\right] - 1. \tag{3.15}$$

Let Z^y be Pauli strings in $\{I,Z\}^{\otimes n}$ associated to n-bit the binary string y, i.e. $Z^y = \bigotimes_{i\in\{0,1\}^n} Z^{y_i}$, where $Z^0 = I$ and $Z^1 = Z$. We can expand each computational basis state in the Pauli basis:

$$2^n \sum_{x\in\{0,1\}^n} \mathrm{Tr}[\rho|x\rangle\langle x|]^2 = \frac{1}{2^n} \sum_{x\in\{0,1\}^n} \mathrm{Tr}\left[\sum_{y\in\{0,1\}^n} (-1)^{x\cdot y}\rho Z^y\right]^2 \tag{3.16}$$

$$= \frac{1}{2^n} \sum_{y,z\in\{0,1\}^n} \mathrm{Tr}[\rho Z^y]\,\mathrm{Tr}[\rho Z^z]\left(\sum_{x\in\{0,1\}^n} (-1)^{x\cdot(y+z)}\right) = \sum_{y\in\{0,1\}^n} \mathrm{Tr}[\rho Z^y]^2, \tag{3.17}$$

where we used the identity $\sum_{x\in\{0,1\}^n}(-1)^{x\cdot(y+z)} = \delta_{yz}2^n$. This immediately implies the following characterization of the scaled collision probability:

$$\mathcal{Z}(\nu) = 2^n \cdot \mathbb{E}_{\rho\sim\nu} \sum_{P\in\{I,Z\}^{\otimes n}} \mathrm{Tr}[\rho P]^2 - 1. \tag{3.18}$$

We say that ν exhibits the *anti-concentration* property if

$$\mathcal{Z}(\nu) = O(1). \tag{3.19}$$

3.2.4 Approximate scrambling

A further property of quantum ensembles is the presence (or lack) of scrambling. The notion of scrambling also arises in the black-hole information paradox [LSH+13] and refers to the process of mapping most initial pure product states to states that are highly entangled [BF12, BF15]. From the standpoint of Pauli basis, scrambling tends to reduce the mass of low-weight Pauli strings, as demonstrated in the example below. We revisit the definition of scrambling given in [HM23].

Definition 3.5 (Approximate scrambler)**.** Let $k \le n$ a positive integer. An ensemble of unitaries μ is an (ε,k)-approximate scrambler if for any density matrix $\rho \in \mathcal{S}_n$ and subset S of qubits with $|S| \le k$.

$$\mathbb{E}_{U\sim\mu} \left\| \rho_S(U) - \frac{I}{2^{|S|}} \right\|_1^2 \le \varepsilon, \tag{3.20}$$

where $\rho_S(U) = \mathrm{Tr}_{\overline{S}} U|0^n\rangle\langle 0^n|U^\dagger$

Example 3.3 (Scrambling wipes out low-weight Paulis)**.** We will now derive an implication of the definition of ε-approximate scrambler, restating an argument given in [BF12, BF15]. Let $k = |S|$. We first rewrite the squared 2-distance in terms of the Pauli strings.

$$\left\| \rho_S(U) - \frac{I}{2^k} \right\|_2^2 = \left(\mathrm{Tr}[\rho_S(U)^2] - \frac{1}{2^k} \right) \tag{3.21}$$

$$= \frac{1}{2^k}\left(\frac{1}{2^k} \mathrm{Tr}\left[\sum_{P\in\mathscr{P}_k} \mathrm{Tr}[P\rho_S(U)]P \right]^2 - 1 \right) = \tag{3.22}$$

$$= \frac{1}{2^k} \sum_{P\in\mathscr{P}_k} \mathrm{Tr}[P\rho_S(U)]^2 - \frac{1}{2^k} = \frac{1}{2^k} \sum_{P\in\mathscr{P}_k\setminus\{I\}} \mathrm{Tr}[P\rho_S(U)]^2. \tag{3.23}$$

By the Cauchy-Schwartz inequality, the squared 1 and 2-distances are within a factor 2^k:

$$\left\|\rho_S(U) - \frac{I}{2^k}\right\|_1^2 \leq 2^k \left\|\rho_S(U) - \frac{I}{2^k}\right\|_2^2 \leq 2^k \left\|\rho_S(U) - \frac{I}{2^k}\right\|_1^2. \tag{3.24}$$

Thus,

$$\sum_{P\in\mathscr{P}_k\setminus\{I\}} \mathrm{Tr}[P\rho_S(U)]^2 \leq 2^k \cdot \varepsilon. \tag{3.25}$$

Thus the contribution of the Paulis with weight smaller than k is at most $2^k \cdot \varepsilon$.

3.2.5 Pauli invariant ensembles

We conclude this section with a property of measures over n-qubit unitaries that are invariant under right or left multiplication of random Pauli. So, in particular, this includes circuits whose initial layer consists in Haar-random single-qubit gates. In particular, 2-design property implies Pauli invariance, but not the converse. The following is a variant of the results of Ref. ([AGL+23a], Lemma 2).

Lemma 3.2 (Pauli invariant distributions)**.** *Let $\mathscr{D}$ be any distribution over n-qubit unitaries that is invariant under right-multiplication of random Pauli, i.e., for any measurable function F,*

$$\mathbb{E}_{U\sim\mathscr{D}}[F(U)] = \frac{1}{4^n}\sum_{P\in\mathscr{P}_n} \mathbb{E}_{U\sim\mathscr{D}}[F(UP)]. \tag{3.26}$$

Then for any $P, Q \in \mathscr{P}_n$ such that $P \neq Q$, we have

$$\mathbb{E}_{U\sim\mathscr{D}}[UPU^\dagger \otimes UQU^\dagger] = 0. \tag{3.27}$$

Similarly, assuming left-invariance instead of right-invariance, for any $P, Q \in \mathscr{P}_n$ such that $P \neq Q$, we obtain

$$\mathbb{E}_{U\sim\mathscr{D}}[U^\dagger PU \otimes U^\dagger QU] = 0. \tag{3.28}$$

Proof. We will prove only the first statement, as the proof of the second is analogous. First, let us use the invariance under right-multiplication of random Pauli operators

$$\begin{aligned}\mathbb{E}_{U\sim\mathscr{D}}[UPU^\dagger \otimes UQU^\dagger] &= \frac{1}{4^n}\sum_{R\in\mathscr{P}_n} \mathbb{E}_{U\sim\mathscr{D}}\left[URPRU^\dagger \otimes URQRU^\dagger\right] \\ &= \frac{1}{4^n}\mathbb{E}_{U\sim\mathscr{D}}\left[U^{\otimes 2}\left(\sum_{R\in\mathscr{P}_n} RPR\otimes RQR\right)(U^\dagger)^{\otimes 2}\right].\end{aligned} \tag{3.29}$$

It suffices to show that

$$\sum_{R\in\mathscr{P}_n} RPR \otimes RQR = 0. \tag{3.30}$$

Let $\langle\!\langle P, Q\rangle\!\rangle := 1[P \text{ and } Q \text{ anticommute}]$, so that

$$\begin{aligned}\sum_{R\in\mathscr{P}_n} RPR\otimes RQR &= \sum_{R\in\mathscr{P}_n} (-1)^{\langle\!\langle P,R\rangle\!\rangle + \langle\!\langle Q,R\rangle\!\rangle} P\otimes Q \\ &= \sum_{R\in\mathscr{P}_n} (-1)^{\langle\!\langle PQ,R\rangle\!\rangle} P\otimes Q = 0,\end{aligned} \tag{3.31}$$

where the last line follows from the fact that PQ is not identity, and therefore commutes with half Paulis and anticommutes with the other half. ■

3.3 Distances and divergences over quantum states

We provide a concise introduction to different measures of distance and divergences used for comparing quantum states in various contexts.

3.3.1 Schatten *p*-norms

Schatten p-norm can be used to define distances between linear operators. The Schatten p-norm of an operator $A \in \mathscr{L}_n$ is given by

$$\|A\|_p := [\mathrm{Tr}\left\{|A|^p\right\}]^{1/p},$$

where $|A| := \sqrt{A^\dagger A}$ and $p \geq 1$. For each $p \in [1,\infty]$, we consider the dual index q such that $\frac{1}{p} + \frac{1}{q} = 1$. The Hölder inequality gives:

$$\mathrm{Tr}\left\{A^\dagger B\right\} \leq \|A\|_p \|B\|_q. \tag{3.32}$$

For two quantum states $\rho, \sigma \in \mathscr{S}_n$, the *trace distance* is defined as follows:

$$\|\rho - \sigma\|_{\mathrm{tr}} := \frac{1}{2}\|\rho - \sigma\|_1. \tag{3.33}$$

Notably, the trace distance admits the following variational characterization:

$$\|\rho - \sigma\|_{\mathrm{tr}} = \max_{0 \leq M \leq I} \mathrm{Tr}[M(\rho - \sigma)]. \tag{3.34}$$

The above maximization is with respect to all positive semi-definite operators $M \in \mathscr{L}_n$ that have their eigenvalues bounded from above by one. Another key feature of the trace distance is its invariance under unitary evolution.

$$\|U\rho U^\dagger - U\sigma U^\dagger\|_{\mathrm{tr}} = \|\rho - \sigma\|_{\mathrm{tr}}. \tag{3.35}$$

Moreover, it is convenient to write the following spectral decomposition

$$\rho - \sigma = \sum_i \lambda_i |i\rangle\langle i| = X^+ - X^-,$$

where X^+ and X^- denote respectively the positive part and the negative part of $\rho - \sigma$, i.e.

$$X^+ := \sum_{\lambda_i > 0} \lambda_i |i\rangle\langle i|, \quad X^- := \sum_{\lambda_i < 0} \lambda_i |i\rangle\langle i|.$$

In particular, the following identities can be easily verified:

$$\|\rho - \sigma\|_{\mathrm{tr}} = \mathrm{Tr}(X^+) = \mathrm{Tr}(X^-). \tag{3.36}$$

3.3.2 The quantum Wasserstein distance of order 1

We adopt the definition of quantum Wasserstein distance of order 1 proposed in [DPMTL21a]. This is based on the following notion of neighbouring quantum states, which also arises in the context of differentially private measurements [AR19]. We say that ρ and $\sigma \in \mathscr{S}_n$ are neighbouring if they

coincide after discarding one qubit, i.e., if $\mathrm{Tr}_i \rho = \mathrm{Tr}_i \sigma$ for some $i \in [n]$. The quantum W_1 distance between the quantum states ρ and σ of $\mathscr{H}_n$ is defined as

$$W_1(\rho,\sigma) = \min\Big(\sum_{i=1}^n c_i : c_i \geq 0, \rho - \sigma = \sum_{i=1}^n c_i \left(\rho^{(i)} - \sigma^{(i)}\right), \tag{3.37}$$

$$\rho^{(i)}, \sigma^{(i)} \in \mathscr{S}_n, \mathrm{Tr}_i \rho^{(i)} = \mathrm{Tr}_i \sigma^{(i)}\Big). \tag{3.38}$$

Intuitively, the distance $W_1(\rho,\sigma)$ is associated to the number of local operations needed to turn ρ into a state close to σ in trace distance. The W_1 distance is induced by the associated quantum W_1 norm. For $X \in \mathscr{O}_n^T$, we define

$$\|X\|_{W_1} = \frac{1}{2}\left(\sum_{i=1}^n \|X^{(i)}\|_1 : X^{(i)} \in \mathscr{O}_n^T, \mathrm{Tr}_i X^{(i)} = 0, X = \sum_{i=1}^n X^{(i)} \right). \tag{3.39}$$

The quantum W_1 norm and the trace norm are always within a factor n,

$$\frac{n}{2}\|X\|_1 \leq \|X\|_{W_1} \leq \frac{n}{2} \cdot \|X\|_1. \tag{3.40}$$

We also need the following technical lemma that can be used to upper bound the quantum W_1 distance under the action of a local evolution.

Lemma 3.3 (Proposition 5, [DPMTL21a])**.** *Let $\mathscr{I} \subseteq [n]$, and let $\rho, \sigma \in \mathscr{S}_n$ such that $\mathrm{Tr}_{\mathscr{I}} \rho = \mathrm{Tr}_{\mathscr{I}} \sigma$,*

$$W_1(\rho,\sigma) \leq |\mathscr{I}| \frac{d^2-1}{d^2} \|\rho - \sigma\|_1. \tag{3.41}$$

We will employ the *contraction coefficient* of a channel Φ with respect to the quantum W_1 distance, defined as

$$\|\Phi\|_{W_1 \to W_1} := \max_{\rho \neq \sigma \in \mathscr{S}(\mathbb{C}^{2^n})} \frac{W_1(\Phi(\rho), \Phi(\sigma))}{W_1(\rho,\sigma)} = \max_{\substack{X \in \mathscr{O}_n^T, \\ \|X\|_{W_1}=1}} \|\Phi(X)\|_{W_1}. \tag{3.42}$$

Note that contraction coefficient $\|\cdot\|_{W_1 \to W_1}$ is not in general bounded by 1, as the W_1 does not satisfy a data-processing inequality for all channels. However, Φ is a layer of of k-qubit gates, the contraction coefficient of Φ can be bounded by light-cone argument as follows

$$\|\Phi\|_{W_1 \to W_1} \leq \begin{cases} 1 & \text{if } k = 1, \\ \frac{3}{2}k & \text{if } k > 1 \ \text{([DPMTL21a], Proposition 13).} \end{cases} \tag{3.43}$$

And thus a layer of two qubit gates has contraction coefficient at most 3. For instance, consider a local channel $\mathscr{N}(X) = (1-p)X + \mathrm{Tr}[X]\sigma$ for a qubit state $\sigma \in \mathscr{S}(\mathbb{C}^2)$. As proven in [DPMTL21a], the contraction coefficient of corresponding tensor power channel can be computed exactly,

$$\|\mathscr{N}^{\otimes n}\|_{W_1 \to W_1} = (1-p). \tag{3.44}$$

For other kinds of local noise, computing the exact expression of the contraction coefficient can be a complicated task, and thus we need to resort to coarse upper bounds. If $\mathscr{N}$ is a single-qubit channel, the contraction coefficient of the tensor power channel $\mathscr{N}^{\otimes n}$ can be upper bounded by the diamond distance between $\mathscr{N}$ and a suitable 1-qubit channel $\mathscr{E}$. In particular, we recall a result given in Ref. [DPMTL21a].

Proposition 3.1 (Proposition 11, [DPMTL21a]). *Let Φ be a quantum channel on $\mathbb{C}^d$ with fixed point $\omega \in \mathscr{S}_1$ and let $\mathscr{E}$ the quantum channel on $\mathbb{C}^d$ that replaces the input state with ω. Then,*

$$\frac{1}{2}\|\Phi - \mathscr{E}\|_{1\to 1} \leq \|\Phi^{\otimes n}\|_{W_1 \to W_1} \leq \|\Phi - \mathscr{E}\|_\diamond \leq 2\|\Phi - \mathscr{E}\|_{1\to 1}, \tag{3.45}$$

where we recall that for any single-qubit linear map $\mathscr{F}$,

$$\|\mathscr{F}\|_{1\to 1} = \max_{\rho \in \mathscr{S}_1} \|\mathscr{F}(\rho)\|_1, \tag{3.46}$$

$$\|\mathscr{F}\|_\diamond = \max_{\rho \in \mathscr{S}_2} \|\mathscr{F} \otimes I_2(\rho)\|_1. \tag{3.47}$$

We also define the quantum Lipschitz constant of a self-adjoint linear operator $H \in \mathscr{O}_n$:

$$\|H\|_L = \max_{i\in[n]}(\max(\mathrm{Tr}[H(\rho-\sigma)] : \rho,\sigma \in \mathscr{S}_n, \mathrm{Tr}_i\,\rho = \mathrm{Tr}_i\,\sigma)). \tag{3.48}$$

From the definition of W_1 distance, we can readily derive that

$$\mathrm{Tr}[H(\rho-\sigma)] \leq \|H\|_L W_1(\rho,\sigma). \tag{3.49}$$

The quantum Lipschitz constant is particularly useful to determine concentration inequalities of noisy states [DPMRF23]. In particular, the maximally mixed state satisfies the following *Gaussian concentration inequality* for every observable O,

$$\Pr_{I/2^n}\left[\left|O - \frac{\mathrm{Tr}[O]}{2^n}\right| \geq an\right] \leq \exp\left(-\frac{a^2 n}{\|O\|_L}\right). \tag{3.50}$$

3.3.3 Rényi divergences

In the classical setting, for two probability measures P, Q the Rényi divergences of order $\alpha \in (1,\infty)$ are defined as

$$D_\alpha(P\|Q) = \frac{1}{\alpha-1}\log \mathbb{E}_{x\sim Q}\left(\frac{P(x)}{Q(x)}\right)^\alpha,$$

where we adopt the conventions that $0/0 = 0$ and $z/0 = \infty$ for $z > 0$. In the limit $\alpha \to 1$, the Rényi divergence reduces to the relative entropy, also known as the Kullback-Leibler divergence, i.e. $\lim_{\alpha\to 1} D_\alpha(P\|Q) = D(P\|Q) = \mathbb{E}_{x\sim P}\log\frac{P(x)}{Q(x)}$. Moreover, by taking the limit $\alpha \to \infty$, we obtain the max-divergence

$$D_\infty(P\|Q) = \sup_{S\subseteq \mathsf{supp}(Q)} \log\frac{P(S)}{Q(S)}.$$

We will also need the related smooth max-divergence,

$$D_\infty^\delta(P\|Q) = \sup_{S\subseteq \mathsf{supp}(Q): P(S)\geq\delta} \log\frac{P(S)-\delta}{Q(S)}.$$

We emphasise that $D_\infty^\delta(P\|Q) \leq \varepsilon$ if and only if for every subset S,

$$P(S) \leq e^\varepsilon Q(S) + \delta.$$

Notably, the (smooth) max-divergence occurs in the definition of differential privacy.

Now we introduce divergences for quantum states. We make use of the quantum Petz-Rényi divergences [MH11, MLDS+13] of order $\alpha \in (1,\infty)$. For two states ρ,σ such that the support of ρ is included in the support of σ, they are defined as

$$D_\alpha(\rho\|\sigma) = \frac{1}{\alpha-1}\log \mathrm{Tr}[\rho^\alpha \sigma^{1-\alpha}].$$

In case the support of ρ is not contained in that of σ, all the divergences above are defined to be $+\infty$. In the limit $\alpha \to 1$, the quantum Petz-Rényi divergence reduces to the quantum relative entropy, i.e., $\lim_{\alpha\to 1} D_\alpha(\rho\|\sigma) = D(\rho\|\sigma) = \mathrm{Tr}[\rho(\log\rho - \log\sigma)]$. We also consider the divergence obtained by taking the limit $\alpha \to \infty$, known as quantum max-divergence,

$$D_\infty(\rho\|\sigma) = \inf\{\lambda : \rho \le e^\lambda \sigma\},$$

and the related quantum smooth max-divergence [HRF23],

$$D^\delta_\infty(\rho\|\sigma) = \inf_{\overline{\rho}\in B_\delta(\rho)} D_\infty(\overline{\rho}\|\sigma),$$

where $B^\delta(\rho) = \{\overline{\rho} : \overline{\rho}^\dagger = \overline{\rho} \ge 0 \wedge \|\rho - \overline{\rho}\|_1 < 2\delta\}$. Similarly to its classical counterpart, the quantum (smooth) max-divergence plays a central role in this thesis as it occurs in the definition of differentially private quantum channels.

The (standard) joint convexity of the Rényi divergence for $\alpha \in [0,\infty]$ is proven in [vEH14] (Theorem 13). For the max divergence have

$$D_\infty(\sum_i \lambda_i P_i \| \sum_i \lambda_i Q_i) \le \max_i D_\infty(P_i\|Q_i).$$

For the smooth max divergence, we can easily prove the statement from scratch. Assume $P_i(x) \le e^\varepsilon Q_i(x) + \delta$:

$$\sum_i \lambda_i P_i(x) \le \sum_i \lambda_i (e^\varepsilon Q_i(x) + \delta) = e^\varepsilon \left(\sum_i \lambda_i Q_i(x)\right) + \delta.$$

3.3.4 The measured Pinsker's inequality

We provide an alternative version of the popular Pinsker's inequality [HOT81], where the quantum relative entropy is replaced by the measured relative entropy. As the proof is almost identical to the one of the (standard) quantum Pinsker's inequality, this can be regarded as a folklore result. We include it here since we were unable to find an appropriate reference.

Lemma 3.4 (Measured Pinsker's inequality)**.** *For ρ,σ quantum states, the following inequality holds:*

$$\|\rho - \sigma\|^2_{\mathrm{tr}} \le \frac{1}{2} D_M(\rho\|\sigma),$$

where $D_M(\rho\|\sigma)$ is the measured relative entropy.

Proof. Recall the variational interpretation of the trace distance as a probability difference:

$$\|\rho-\sigma\|_{\mathrm{tr}} = \max_{0\le\Lambda\le I} \mathrm{Tr}[\Lambda(\rho-\sigma)] = |\mathscr{M}(\rho)-\mathscr{M}(\sigma)|_{\mathrm{tv}},$$

where $\mathscr{M} = (\Lambda^*, I-\Lambda^*)$ and $\Lambda^* = \arg\max_{0\le\Lambda\le I}\mathrm{Tr}[\Lambda(\rho-\sigma)]$. The classical Pinsker's inequality yields:

$$|\mathscr{M}(\rho)-\mathscr{M}(\sigma)|_{\mathrm{tv}}^2 \le \frac{1}{2}D(\mathscr{M}(\rho)\|\mathscr{M}(\sigma)) \le \frac{1}{2}D_M(\rho\|\sigma),$$

where the second inequality follows from the definition of measured relative entropy. This proves the lemma. ■

The standard inequality can be deduced by noting that $D_M(\rho\|\sigma) \le D(\rho\|\sigma)$.

3.3.5 The quantum hockey-stick divergence

The quantum hockey-stick divergence was first introduced in [SW12], in the context of exploring strong converse bounds for the quantum capacity, and further investigate in [HRF23] in the context of quantum differential privacy. It is defined as

$$E_\gamma(\rho\|\sigma) := \mathrm{Tr}(\rho-\gamma\sigma)^+, \tag{3.51}$$

for $\gamma \ge 1$. Here X^+ denotes the positive part of the eigendecomposition of a Hermitian matrix $X = X^+ - X^-$. In [SW12] it was noted that this quantity is closely related to the trace norm via

$$E_\gamma(\rho\|\sigma) = \frac{1}{2}\|\rho-\gamma\sigma\|_1 + \frac{1}{2}(\mathrm{Tr}(\rho)-\gamma\,\mathrm{Tr}(\sigma)), \tag{3.52}$$

so for ρ,σ quantum states, $E_1(\rho\|\sigma) = \frac{1}{2}\|\rho-\sigma\|_1$ equals the trace distance. We also state some useful properties of the hockey-stick divergence proven in ([HRF23], Proposition II.5).

- (Triangle inequality) For $\gamma_1,\gamma_2 \ge 1$ and $\rho,\sigma \in \mathscr{S}_n$, we have

$$E_{\gamma_1\gamma_2}(\rho\|\sigma) \le E_{\gamma_1}(\rho\|\tau) + \gamma_1 E_{\gamma_2}(\tau\|\sigma). \tag{3.53}$$

- (Convexity) Let $\gamma_1,\gamma_2 \ge 1$, $\rho = \sum_x p(x)\rho_x$ and $\sigma = \sum_x q(x)\sigma_x$ with $\rho_x,\sigma_x \in \mathscr{S}_n$, we have

$$E_{\gamma_1\gamma_2}(\rho\|\sigma) \le \sum_x p(x)E_{\gamma_1}(\rho_x\|\sigma_x) + \gamma_1 E_{\gamma_2}(\tilde{p}\|\tilde{q}), \tag{3.54}$$

  where $\tilde{p}$ and $\tilde{q}$ are non-normalised distributions $\tilde{p}(x) = p(x)\mathrm{Tr}\sigma_x$ and $\tilde{q}(x) = q(x)\mathrm{Tr}\sigma_x$, respectively. This also implies convexity and joint convexity.

- (Stability) For $\gamma \ge 1$ and $\rho,\sigma,\tau \in \mathscr{S}_n$, we have

$$E_\gamma(\rho\otimes\tau\|\sigma\otimes\tau) = \mathrm{Tr}\,[\tau]\,E_\gamma(\rho\|\sigma). \tag{3.55}$$

3.4 Quantum channels

An n-qubit ideal quantum circuit can be represented by a unitary operator $U \in \mathscr{U}_n$. However, this representation does not captures the imperfections of real devices, which often manifests as incoherent noise and hence irreversible operations. General evolutions of quantum states can be represented as quantum channels. Quantum channels bridge in a unified formalism both unitary operators and classical channels, similary to how density matrices incorporate both quantum pure states and classical probability distributions. We define a *quantum channel* $\mathscr{N} : \mathscr{L}_n \to \mathscr{L}_m$ as a linear, completely positive and trace-preserving map. Complete positivity means that for all positive operators $\sigma \in \mathscr{L}((\mathbb{C}^2)^{\otimes n} \otimes \mathbb{C}^d)$, for any $d \in \mathbb{N}$, the operator $\mathscr{N} \otimes \mathrm{Id}(\sigma)$ is positive. Trace-preservation, i.e. $\mathrm{Tr}(\mathscr{N}(A)) = \mathrm{Tr}(A)$ for any $A \in \mathscr{L}_n$, corresponds to the conservation of probabilities. We say that a quantum channel $\mathscr{N}$ is unital if it preserves the identity, i.e. if $\mathscr{N}(I) = I$, and non-unital otherwise. Any quantum channel $\mathscr{N}$ over n-qubit can be represented in terms of $2^n \times 2^n$ Kraus operators $K_1, K_2, \dots K_{4^n} \in \mathscr{L}_n$, i.e.,

$$\mathscr{N}(\cdot) = \sum_{i=1}^{4^n} K_i (\cdot) K_i^\dagger, \tag{3.56}$$

where the condition $\sum_{i=1}^{4^n} K_i^\dagger K_i = I$ is needed to satisfy trace-preservation. For a unitary channel $U(\cdot)U^\dagger$, all Kraus operators but one are zeros.

Example 3.4 (Depolarizing channel)**.** The n-qubit depolarizing channel $\mathscr{N}_p^{(\mathrm{dep})}$ is a channel acting as identity with probability $1-p$ and returning the maximally mixed state with the remaining probability. Thus for $A \in \mathscr{L}_n$, we have

$$\mathscr{N}_p^{(\mathrm{dep})}(A) = (1-p)A + p \cdot \mathrm{Tr}[A] \frac{I}{2^n}. \tag{3.57}$$

Since the Pauli group $\mathscr{P}_n$ forms a 1-design (Example 3.1), we can rewrite the channel as follows

$$\mathscr{N}_p^{(\mathrm{dep})}(A) = (1-p) IAI + p \cdot \frac{1}{4^n} \sum_{P \in \mathscr{P}_n} PAP \tag{3.58}$$

$$= \left(1 - p + \frac{p}{4^n}\right) IAI + \frac{p}{4^n} \sum_{P \in \mathscr{P}_n \setminus \{I\}} PAP, \tag{3.59}$$

which immediately yields the expression of the 4^n Kraus operators,

$$K_i = \begin{cases} \sqrt{1 - p + \frac{p}{4^n}}\, I & \text{for } i = 1, \\ \frac{\sqrt{p}}{2^n} P_i & \text{for } i > i, \text{ where } P_i \in \mathscr{P}_n \setminus \{I\}. \end{cases} \tag{3.60}$$

* * *

In this thesis we will also employ two alternative representations of quantum channels, based respectively on the *Pauli Transfer Matrix* and the *Choi-Jamiolkowski isomorphism*.

3.4.1 The Pauli Transfer Matrix

As previously discussed, the Pauli strings form an orthonormal basis for the (scaled) Hilbert-Schimdt inner product, therefore we can represent states as a linear combination of Pauli strings. As a consequence, a quantum channel $\mathcal{N}$ is fully determined by its action on Pauli strings. Then the Pauli Transfer Matrix (PTM) of $\mathcal{N}$ is a $4^n \times 4^n$ matrix whose entries are

$$t_{P,Q} = \frac{1}{2^n} \mathrm{Tr}[Q\mathcal{N}(P)]. \tag{3.61}$$

Then the action of $\mathcal{N}$ on $\mathscr{P} \in \mathscr{P}_n$ can be expressed concisely as $\mathcal{N}(P) = \sum_{Q \in \mathscr{P}_n} t_{P,Q} Q$. Trace-preservation implies that

$$t_{P,I} = \frac{1}{2^n} \mathrm{Tr}[\mathcal{N}(P)] = \frac{1}{2^n} \mathrm{Tr}[P] = \begin{cases} 1 & \text{if } P = I \\ 0 & \text{otherwise.} \end{cases} \tag{3.62}$$

Example 3.5 (Depolarizing channel)**.** To compute the Pauli Transfer Matrix of the n-qubit depolarizing channel $\mathcal{N}_p^{(\mathrm{dep})}$, it suffices to plug a non-identity Pauli string $P \in \mathscr{P}_n \setminus \{I\}$ in the definition of the channel and observe that

$$\mathcal{N}_p^{(\mathrm{dep})}(P) = (1-p)P, \tag{3.63}$$

and therefore

$$\forall P, Q \in \mathscr{P}_n \setminus \{I\} : t_{P,Q} = \delta_{P,Q}(1-p), \tag{3.64}$$

where $\delta_{P,Q}$ is the Kronecker's delta.

Example 3.6 (Amplitude damping channel)**.** The single-qubit amplitude damping channel $\mathcal{N}_q^{(\mathrm{amp})}$ is defined by the following Kraus operators:

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-q} \end{pmatrix}, \; K_1 = \begin{pmatrix} 0 & \sqrt{q} \\ 0 & 0 \end{pmatrix}. \tag{3.65}$$

Therefore a single-qubit linear operator X undegoes the following transformation:

$$X = \begin{pmatrix} x_{00} & x_{01} \\ x_{10} & x_{11} \end{pmatrix} \mapsto \mathcal{N}_q^{(\mathrm{amp})}(X) = \begin{pmatrix} x_{00} + q x_{11} & \sqrt{1-q}\, x_{01} \\ \sqrt{1-q}\, x_{10} & (1-q) x_{11} \end{pmatrix}. \tag{3.66}$$

Replacing X with the single-qubit Pauli operators yields the entries of the Pauli Transfer Matrix. Other than $t_{I,I} = 1$, the non-zero PTM elements are

$$t_{I,Z} = q, \quad t_{X,X} = t_{Y,Y} = \sqrt{1-q}, \quad t_{Z,Z} = 1-q. \tag{3.67}$$

Example 3.7 (Purity of the evolution of the single-qubit maximally mixed state)**.** We will now compute the purity of the state obtained by applying the single-qubit channel $\mathcal{N}$ on the maximally mixed state.

$$\frac{1}{4} \mathrm{Tr}[\mathcal{N}(I)^2] = \frac{1}{4} \mathrm{Tr}[(I + t_{I,X} X + t_{I,Y} Y + t_{I,Z} Z)^2] = \frac{1}{2}(1 + t_{I,X}^2 + t_{I,Y}^2 + t_{I,Z}^2). \tag{3.68}$$

3.4.2 The Choi-Jamiolkowski isomorphism

Furthermore, we can represent a channel with its dual state, known as Choi-Jamiolkowski state, or simply Choi state [Cho75, Jam72]. This will play a central role in several algorithms proposed in Chapter 6. The Choi state $\mathscr{J}(\mathcal{N})$ can be prepared by first creating the maximally entangled state on $2n$ qubits, which we denoted by $|\Omega\rangle$, and then applying $\mathcal{N}$ on half of the maximally entangled state. This is equivalent to preparing n Einstein–Podolsky–Rosen (EPR) pairs $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ (which altogether forms $2n$ qubits) and applying the channel $\mathcal{N}$ to the n qubits coming from the second half of each of the EPR pairs. We have

$$\mathscr{J}(\mathcal{N}) = \mathrm{Id} \otimes \mathcal{N}(|\Omega\rangle\langle\Omega|) = \frac{1}{2^n} \sum_{i,j \in \{0,1\}^n} |i\rangle\langle j| \otimes \mathcal{N}(|i\rangle\langle j|). \tag{3.69}$$

We emphasize that the n EPR pairs may be prepared with a constant depth circuit. If $\mathcal{N} = U(\cdot)U^\dagger$ is a unitary channel, the Choi state $\mathscr{J}(\mathcal{N})$ is pure and we denote it by $\mathscr{J}(\mathcal{N}) = |v(U)\rangle\langle v(U)|$.

Example 3.8 (Choi states of Paulis)**.** The Choi states of Pauli strings are of particular interest. First, we note that the Choi states of the single-qubit Pauli operators are proportional to the Bell basis:

$$|v(I)\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle),\ |v(X)\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \tag{3.70}$$

$$i\,|v(Y)\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle),\ |v(Z)\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle). \tag{3.71}$$

Hence, the set $\{|v(I)\rangle, |v(X)\rangle, |v(Y)\rangle, |v(Z)\rangle\}^{\otimes n}$ forms an orthonormal basis for $2n$-qubit pure states with respect to the inner product $|\langle\cdot|\cdot\rangle|$.

Example 3.9 (Purity of the Choi state of a single-qubit channel)**.** It is convenient to express the purity of the Choi state of a single-qubit channel $\mathcal{N}$ in terms of the elements of the Pauli Transfer Matrix. First, recall that Choi state can be expressed as follows:

$$J(\mathcal{N}) := \mathcal{N} \otimes I\left(\frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|)\right) \tag{3.72}$$

$$= \frac{1}{4}(\mathcal{N}(I) \otimes I + \mathcal{N}(X) \otimes X - \mathcal{N}(Y) \otimes Y + \mathcal{N}(Z) \otimes Z), \tag{3.73}$$

where we expanded the Bell state in the Pauli basis. Now recall that in general $\mathcal{N}(Q) = \sum_{P \in \{I,X,Y,Z\}} t_{Q,P} P$. Hence we can express the purity of the Choi state as:

$$\mathrm{Tr}[J(\mathcal{N})^2] = \frac{1}{16} \sum_P \mathrm{Tr}[\mathcal{N}(P)^2 \otimes I] = \frac{1}{16} \sum_{P,Q} t_{P,Q}^2 \,\mathrm{Tr}[I \otimes I] = \frac{1}{4} \sum_{P,Q} t_{P,Q}^2. \tag{3.74}$$

* * *

As the name suggests, the mapping from channels to states $\mathcal{N} \leftrightarrow \mathscr{J}(\mathcal{N})$ is an isomorphism and therefore it induces a distance over channels, previously introduced in [BY23]. In particular, we define

$$D(\mathcal{M}, \mathcal{N}) := \frac{1}{\sqrt{2}} \| \mathscr{J}(\mathcal{N}) - \mathscr{J}(\mathcal{M}) \|_2 \tag{3.75}$$

When the channels $\mathcal{N} = U(\cdot)U^\dagger$ and $\mathcal{M} = V(\cdot)V^\dagger$ are unitary, we simply write $D(U,V)$ instead of $D(U(\cdot)U^\dagger, V(\cdot)V^\dagger)$. Since for pure states the 1-distance and the 2-distance are equal up to a scaling factor, we also obtain,

$$D(U,V) = \| |v(U)\rangle\langle v(U)| - |v(V)\rangle\langle v(V)| \|_{\mathrm{tr}} = \sqrt{1 - |\langle v(U)|v(V)\rangle|^2}. \tag{3.76}$$

We remark that closely related distance have also appeared in other works. In particular, the pseudo-distance $\mathrm{dist}(U,V)$ of [CNY23] and $D(U,V)$ are within a constant factor $\sqrt{2}$. We now state a useful result relating $D(U,V)$ to the expected risk $\mathscr{R}_\nu(U,V)$ introduced in Section 3.2.2. First, we generalize the definition from unitaries to quantum channels.

$$\mathscr{R}_\nu(\mathcal{M},\mathcal{N}) := \frac{1}{2}\mathbb{E}_{|\psi\rangle\sim\nu}\left[\left\|\mathcal{M}(|\psi\rangle\langle\psi|) - \mathcal{N}(|\psi\rangle\langle\psi|)\right\|_2^2\right], \tag{3.77}$$

It is immediate to see that $\mathscr{R}_\nu(U(\cdot)U^\dagger, V(\cdot)V^\dagger) = \mathscr{R}_\nu(U,V)$, thus this generalization is consistent with the definition given for unitaries. We now rephrase a result of [BY23] according to our notation.

Lemma 3.5 ([BY23], Proposition 15)**.** *For quantum channels $\mathcal{M},\mathcal{N}$, it holds that*

$$\mathscr{R}_{\mu_n}(\mathcal{M},\mathcal{N}) = \frac{2^n}{2^n+1}D(\mathcal{M},\mathcal{N})^2 + \frac{1}{2^n(2^n+1)}\|\mathcal{M}(I) - \mathcal{N}(I)\|_2^2 \tag{3.78}$$

Note that the last term is 0 if $\mathcal{M},\mathcal{N}$ are unital. Therefore, $D(\mathcal{M},\mathcal{N})$ is an "average-case" measure of the distance between quantum channels, and it is closely related to task of learning the action of a channel on a Haar-random state. For unitary channels, Lemma 3.1 swiftly extends this guarantee to all locally scrambled ensembles of states.

3.4.3 Adjoint channels

Every quantum channel admits a dual transformation referred as the adjoint, or dual, channel. While quantum channels model evolution of quantum states, the adjoint channels capture the evolution of observables in the so-called *Heisenberg picture*. For a quantum channel $\mathcal{N}$, we define its adjoint $\mathcal{N}^\dagger$, as the unique linear map satisfying the following for all $A,B \in \mathscr{L}_n$:

$$\mathrm{Tr}[\mathcal{N}(A)B] = \mathrm{Tr}[A\mathcal{N}^\dagger(B)] \tag{3.79}$$

Since quantum channels are trace-preserving, adjoint channels are always unital. This can be seen with simple manipulations

$$\mathrm{Tr}[A] = \mathrm{Tr}[\mathcal{N}(A)] = \mathrm{Tr}[I\mathcal{N}(A)] = \mathrm{Tr}[\mathcal{N}^\dagger(I)A] \implies \mathcal{N}^\dagger(I) = I. \tag{3.80}$$

In general, an adjoint channel may not been trace-preserving. Moreover, the adjoint channel $\mathcal{N}^\dagger$ is trace-preserving if and only if the channel $\mathcal{N}$ is unital. This comes as an immediate consequence of the following identity,

$$\mathrm{Tr}[\mathcal{N}^\dagger(A)] = \mathrm{Tr}[\mathcal{N}(I)A]. \tag{3.81}$$

Given this definition, it's natural to ask whether the Pauli Transfer Matrices of $\mathcal{N}$ and $\mathcal{N}^\dagger$ are related. We can easily verify that each matrix is the transpose of the other one. In other words, we have, for all $P \in \mathscr{P}_n$,

$$\mathcal{N}(P) = \sum_{Q \in \mathscr{P}_n} t_{P,Q} Q \quad \text{and} \quad \mathcal{N}^\dagger(P) = \sum_{Q \in \mathscr{P}_n} t_{Q,P} Q. \tag{3.82}$$

It is easy to see that $\mathcal{N}_p^{(\text{dep})} = \mathcal{N}_p^{(\text{dep})\dagger}$, i.e. the depolarizing channel and its adjoint coincide. This is a consequence of the fact that the Pauli Transfer Matrix of the depolarizing channel is diagonal.

CHAPTER 4

MODELING NEAR-TERM NOISY QUANTUM DEVICES

-

This Chapter introduces several technical tools for the analysis of variational quantum algorithms on noisy near-term devices, which will be employed throughout the rest of the thesis. We start by considering the combined effect of a random unitary followed by an arbitrary noise channels, which has been extensively studied in [QFK+22]. Subsequently, we examine a model of noisy circuit, where the noise acts as a tensor power of local channels interspersing the unitary layers, and we discuss the decay of purity within this model [HRF22, DPMRF23]. We also consider the evolution of the quantum Wasserstein distance of order 1 in noisy random circuits, providing novel upper bounds on the average contraction coefficients. Finally, give a concise proof of noise-induced cost concentration under unital noise, which was previously studied in [WFC+21].

4.1 Purity and overlap change after one noisy gate

It is insightful to consider the intertwined effect of random unitaries and noisy channels. We consider a n-qubit random unitary U sampled from a 2-design ν, followed from an arbitrary channel $\mathcal{N}$,

representing the action of noise. For two arbitrary input states ρ_0, σ_0, we have

$$\rho = \mathcal{N}\left(U\rho_0 U^\dagger\right), \tag{4.1}$$

$$\sigma = \mathcal{N}\left(U\sigma_0 U^\dagger\right). \tag{4.2}$$

We are interested in the purity of those states and their overlap, which will play a central role in several technical results of Chapter 5. Their values have been computed in prior work, hinging on Eq. 3.10. We state the final result and refer to ([QFK+22], Section VII) for further details. We have

$$\mathbb{E}_{U\sim\mu_n} \operatorname{Tr}\left[\rho\sigma\right] = \left(\frac{2^{2n}}{2^{2n}-1} - \frac{2^n}{2^{2n}-1}\operatorname{Tr}\left[\rho_0\sigma_0\right]\right)\operatorname{Tr}\left[\mathcal{N}\left(\frac{I}{2^n}\right)^2\right] + \left(\frac{2^{2n}}{2^{2n}-1}\operatorname{Tr}[\rho_0\sigma_0] - \frac{2^n}{2^{2n}-1}\right)\operatorname{Tr}\left[J(\mathcal{N})^2\right]. \tag{4.3}$$

And thus the expression of the purity readily follows,

$$\mathbb{E}_{U\sim\mu_n} \operatorname{Tr}\left[\rho^2\right] = \left(\frac{2^{2n}}{2^{2n}-1} - \frac{2^n}{2^{2n}-1}\operatorname{Tr}\left[\rho_0^2\right]\right)\operatorname{Tr}\left[\mathcal{N}\left(\frac{I}{2^n}\right)^2\right] + \left(\frac{2^{2n}}{2^{2n}-1}\operatorname{Tr}[\rho_0^2] - \frac{2^n}{2^{2n}-1}\right)\operatorname{Tr}\left[J(\mathcal{N})^2\right]. \tag{4.4}$$

The above expressions can be combined together to upper bound the expected 2-distance between ρ and σ. By ([QFK+22], Proposition 3), we have

$$\left\|\rho-\sigma\right\|_2^2 = \operatorname{Tr}\left[\rho^2\right] + \operatorname{Tr}\left[\sigma^2\right] - 2\operatorname{Tr}\left[\rho\sigma\right] \tag{4.5}$$

$$= \frac{2^n}{2^{2n}-1}\left(2^n \operatorname{Tr}[J(\mathcal{N})^2] - \operatorname{Tr}\left[\mathcal{N}\left(\frac{I}{2^n}\right)^2\right]\right)\left\|\rho_0-\sigma_0\right\|_2^2. \tag{4.6}$$

Example 4.1 (Average distance under the amplitude damping noise)**.** We evaluate the upper bound above for the single-qubit case, assuming that $\mathcal{N} = \mathcal{N}_q^{(\text{amp})}$ is the amplitude damping channel of noise rate q. From Examples 3.6, 3.7, 3.9, we obtain

$$\operatorname{Tr}\left[\mathcal{N}_q^{(\text{amp})}\left(\frac{I}{2}\right)^2\right] = \frac{1}{2}(1+q^2), \tag{4.7}$$

$$\operatorname{Tr}\left[J\left(\mathcal{N}_q^{(\text{amp})}\right)^2\right] = 1 - q + \frac{q^2}{2}. \tag{4.8}$$

By plugging these values inside Equation 4.6, we have

$$\left\|\rho-\sigma\right\|_2^2 = \frac{1}{3}(q-1)(q-3)\left\|\rho_0-\sigma_0\right\|_2^2. \tag{4.9}$$

4.2 The interspersed model

Let $\mathcal{U}_i = U_i(\cdot)U_i^\dagger$

$$\Phi_{\boldsymbol{\theta}} := \mathcal{N}^{\otimes n} \circ \mathcal{U}_L \circ \mathcal{N}^{\otimes n} \circ \cdots \circ \mathcal{N}^{\otimes n} \circ \mathcal{U}_1(|0^n\rangle\langle 0^n|) \tag{4.10}$$

V is a 2-qubit parameterized unitary $V(\theta_l) := e^{-i\theta_l H_l}$ with $l \in [m]$.

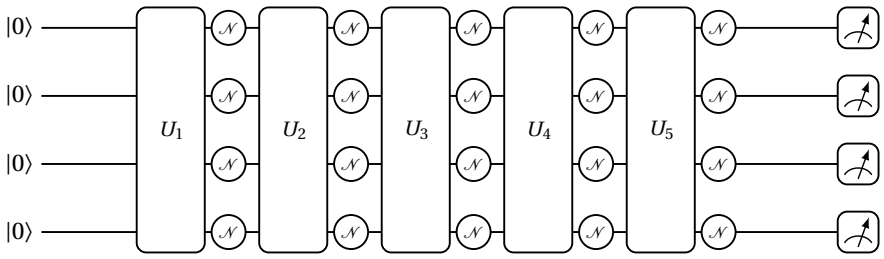


Figure 4.1: Example of a noisy quantum circuit on $n = 4$ qubits

Cost functions and derivatives. We now define $j \in [L]$, and define

$$\Phi_A := \mathcal{N}^{\otimes n} \circ \mathcal{U}_L \circ \cdots \circ \mathcal{N}^{\otimes n} \circ \mathcal{U}_j, \tag{4.11}$$

$$\Phi_B := \mathcal{N}^{\otimes n} \circ \mathcal{U}_{j-1} \circ \cdots \circ \mathcal{N}^{\otimes n} \circ \mathcal{U}_1. \tag{4.12}$$

Therefore, we have $\Phi_{\boldsymbol{\theta}} = \Phi_A \circ \Phi_B$. Let $\mu \in [m]$ a parameter corresponding to a 2-qubit gates in the j-th layer of the circuit. We now give the following lemma, which is also implicit in the proof of ([WFC+21], Theorem 1).

Lemma 4.1 (Derivative)**.** *Let us denote the partial derivative with respect the parameter θ_μ as $\partial_\mu := \frac{\partial}{\partial \theta_\mu}$, then we have*

$$\partial_\mu C(\boldsymbol{\theta}) = i \operatorname{Tr}\left[\Phi_B(\rho_0)\left[H_\mu, \Phi_A^\dagger(H)\right]\right]. \tag{4.13}$$

Proof. We can represent the cost function as

$$C(\boldsymbol{\theta}) := \operatorname{Tr}\left[\Phi_{\boldsymbol{\theta}}(\rho_0) H\right] = \operatorname{Tr}\left[\Phi_A \circ \Phi_B(\rho_0) H\right] = \operatorname{Tr}\left[\Phi_B(\rho_0)\Phi_A^\dagger(H)\right]. \tag{4.14}$$

Here, the adjoint of $\Phi_A(H)$ can be expressed as

$$\Phi_A^\dagger(H) = U_j^\dagger\left(\mathcal{N}^{\dagger\otimes n} \circ \cdots \circ \mathcal{U}_D^\dagger \circ \mathcal{N}^{\dagger\otimes n}(H)\right) U_j, \tag{4.15}$$

where U_j is the unitary layer of brickwork circuit that contains the 2-qubit unitary $UV(\theta_\mu)$ where U is a 2-qubit gate and V is a 2-qubit parametrized unitary of the form $V(\theta_\mu) = e^{-i\theta_\mu H_\mu}$. By taking the partial derivative with respect the parameter θ_μ, we have

$$\begin{aligned} \partial_\mu C(\boldsymbol{\theta}) &= \operatorname{Tr}\left[\Phi_B(\rho_0)\partial_\mu(\Phi_A^\dagger(H))\right] \\ &= i \operatorname{Tr}\left[\Phi_B(\rho_0) H_\mu \Phi_A^\dagger(H)\right] - i \operatorname{Tr}\left[\Phi_B(\rho_0)\Phi_A^\dagger(H) H_\mu\right] \\ &= i \operatorname{Tr}\left[\Phi_B(\rho_0)\left[H_\mu, \Phi_A^\dagger(H)\right]\right], \end{aligned} \tag{4.16}$$

where we have used the fact that $\partial_\mu e^{-i\theta_\mu H_\mu} = -i H_\mu e^{-i\theta_\mu H_\mu}$. ■

4.2.1 Purity decay in noisy circuits

In this section we consider a circuit interspersed by local noise as in Figure 4.1 and we study the decay of purity of the output state. We will deal with the cases of unital and non-unital noise separately.

Unital noise. We will restrict our attention to the local Pauli noise, though similar results also hold for more general families of unital noise. Following [WFC+21], the coefficients of the Pauli Transfer Matrix of a local Pauli noise channel $\mathcal{N}$ are of the form

$$t_{PQ} = \begin{cases} c_P & \text{if } P = Q \\ 0 & \text{if } P \neq Q, \end{cases} \tag{4.17}$$

where $c_I = 1$ by unitality, and $c_X, c_Y, c_Z \in (-1, 1)$. The noise strength, or noise rate, is characterized by the following parameter,

$$c = \sqrt{\max\{|c_X|, |c_Y|, |c_Z|\}}. \tag{4.18}$$

Thus, the decay 2-Rényi relative entropy (and therefore the purity) can be expressed in terms of the noise strength c.

Lemma 4.2 ((Corollary 5.6, [HRF22]),(Supplementary Lemma 6, [WFC+21]))**.** *Let $\mathscr{C}$ a noisy circuit interspersed by m layers of local Pauli noise of noise rate c Denote by $\rho = \mathscr{C}(\rho_0)$ the output state of the noisy circuit. Then,*

$$D_2\left(\rho \middle\| \frac{I}{2^n}\right) \leq c^{2m} D_2\left(\rho_0 \middle\| \frac{I}{2^n}\right) \leq c^{2m} n. \tag{4.19}$$

This readily implies the following upper bound on the purity

$$\mathrm{Tr}[\rho^2] \leq 2^{n(c^{2m}-1)}. \tag{4.20}$$

We also note that the depolarizing noise $\mathcal{N}_p^{(dep)}$ of noise rate p can be recovered as a special case by setting $c_X = c_Y = c_Z = (1-p)^2$.

Non-unital noise. We will now upper bound the purity of the state produced by a circuit intersperse by non-unital noise. In particular, by mean of the *data-processed triangle inequality* ([CMH17], Theorem 3.1), the authors of [SFGP21, DPMRF23], obtained an upper bound on the purity of the output of a non-unital channel, which is exponentially small in n when the unital component of the noise "dominates" the unital one. In particular, we derive the following corollary for the special case of the amplitude-depolarizing noise model.

Corollary 4.1 (Corollary of ([SFGP21], Lemma 1) or ([DPMRF23], Lemma C.1))**.** *Let $\mathscr{C}$ a noisy circuit interspersed by m layers of local noise, either of the form $\mathcal{N}_{p,q}^{(\mathrm{dep,amp}),\otimes n}$ or $\mathcal{N}_{q,p}^{(\mathrm{amp,dep}),\otimes n}$. Denote by $\rho = \mathscr{C}(\rho_0)$ the output state of the noisy circuit. Then,*

$$D_2\left(\rho \middle\| \frac{I}{2^n}\right) \leq n\left((1-p)^{2m} + q\frac{1-(1-p)^{2m}}{2p-p^2}\right) := n \cdot \delta_m. \tag{4.21}$$

This readily implies the following upper bound on the purity

$$\mathrm{Tr}[\rho^2] \leq 2^{n(\delta_m - 1)}. \tag{4.22}$$

Proof. We first recall that $\mathrm{Tr}[\rho^2] \leq 2^{-n+D_2(\rho\|I/2^n)}$, then first bound implies the second. We note the following

$$D_\infty\left(\mathcal{N}_q^{(\mathrm{amp})\otimes n}\left(\frac{I}{2^n}\right)\middle\|\frac{I}{2^n}\right) = nD_\infty\left(\mathcal{N}_q^{(\mathrm{amp})}\left(\frac{I}{2}\right)\middle\|\frac{I}{2}\right) = nq. \tag{4.23}$$

Then the bound $D_2\left(\rho\|I/2^n\right)$ follows from a direct application of ([DPMRF23], Lemma C.1). ■

Analogous results can be proved by replacing the depolarizing noise with Pauli noise, or substituting the amplitude-damping noise with other kinds of non-unital noise. Note that the term δ_m converges exponentially fast to $q/(2p-p^2)$, and thus in this regime the bound is non-trivial if $q \leq 2p-p^2$. Moreover, for this regime of the noise, if $m \geq c\cdot\log n$ for a sufficiently large constant c, we obtain that $\mathrm{Tr}[\rho^2] = 2^{-\Omega(n)}$. Moreover, if $p>0$ and q is a sufficiently small constant, the purity is exponentially small even after a single layer of noise.

4.3 Average contraction coefficients for the W_1 distance

We will now consider a more general case, where the noise is modeled an arbitrary local channel preceded by a single-qubit gate drawn from a 2-design. The presence of local 2-design is a minimal assumption in our setting and it has the advantage of simplifying the analysis thanks to the Schur-Weyl duality.

Proposition 4.1. *Let $\mathcal{N}$ be a local channel and let $U_1, U_2, \dots U_n$ be random single-qubit gates drawn from a local 2-design. We will denote $U = \bigotimes_{i=1}^n U_i$. Then the average contraction coefficient of $\mathcal{N}^{\otimes n}\circ U(\cdot)U^\dagger$ can be upper bounded as follows:*

$$\mathbb{E}_U\|\mathcal{N}^{\otimes n}\circ U(\cdot)U^\dagger\|_{W_1\to W_1} \leq \min\left\{1, \sqrt{\frac{4}{3}\sum_{P,Q\in\{X,Y,Z\}} t_{P,Q}^2}\right\}. \tag{4.24}$$

Proof. We first need to extend Proposition 3.1 to encompass the presence of random gates. Define $\mathscr{E}$ as the channel mapping any single-qubit states to $\mathcal{N}\left(\frac{I}{2}\right)$. Let $X = \rho - \sigma$ the difference between two states, such that $\mathrm{Tr}_i\,\rho = \mathrm{Tr}_i\,\sigma$, and hence $\mathrm{Tr}_i\,X = 0$. Assume without loss of generality that $i=1$. We notice that $(\mathscr{E}\otimes I_{n-1})(X) = 0$. Therefore,

$$\begin{aligned}
\mathbb{E}_U\|\mathcal{N}^{\otimes n}\circ U(X)U^\dagger\|_{W_1} &= \frac{1}{2}\mathbb{E}_U\|\mathcal{N}^{\otimes n}\circ U(X)U^\dagger\|_1 \leq \frac{1}{2}\mathbb{E}_{U_1}\|(\mathcal{N}\circ U_1(X)U_1^\dagger\otimes I_{n-1})(X)\|_1 && (4.25)\\
&= \frac{1}{2}\mathbb{E}_{U_1}\|((\mathcal{N}\circ U_1(\cdot)U_1^\dagger)\otimes I_{n-1})(X)\|_1 && (4.26)\\
&\leq \frac{1}{2}\mathbb{E}_{U_1}\|(\mathcal{N}\circ U_1(\cdot)U_1^\dagger) - \mathscr{E}\|_\diamond\|X\|_1 \leq \mathbb{E}_{U_1}\|(\mathcal{N}\circ U_1(\cdot)U_1^\dagger) - \mathscr{E}\|_\diamond && (4.27)\\
&\leq 2\mathbb{E}_{U_1}\|(\mathcal{N}\circ U_1(\cdot)U_1^\dagger) - \mathscr{E}\|_{1\to 1}, && (4.28)
\end{aligned}$$

which completes the first part of the proof. Moreover, the induced 1-norm can be upper bounded as follows

$$\mathbb{E}_{U_1}\|\mathcal{N}\circ U_1(\cdot)U_1^\dagger-\mathscr{E}\|_{1\to 1}=\mathbb{E}_{U_1}\max_{\rho\in\mathscr{S}(\mathbb{C}^2)}\|\mathcal{N}\circ U_1(\rho)U_1^\dagger-\mathscr{E}(\rho)\|_1 \tag{4.29}$$

$$=\mathbb{E}_{U_1}\left\|\mathcal{N}\circ U_1(\rho_{U_1})U_1^\dagger-\mathcal{N}\left(\frac{I}{2}\right)\right\|_1\leq\sqrt{2}\,\mathbb{E}_{U_1}\left\|\mathcal{N}\circ U_1(\rho_{U_1})U_1^\dagger-\mathcal{N}\left(\frac{I}{2}\right)\right\|_2 \tag{4.30}$$

$$=\sqrt{2}\,\mathbb{E}_{U_1}\sqrt{\mathrm{Tr}\left[\left(\mathcal{N}\circ U_1(\rho_{U_1})U_1^\dagger\right)^2\right]+\mathrm{Tr}\left[\mathcal{N}\left(\frac{I}{2}\right)^2\right]-2\,\mathrm{Tr}\left[\mathcal{N}\circ U_1(\rho_{U_1})U_1^\dagger\mathcal{N}\left(\frac{I}{2}\right)\right]}, \tag{4.31}$$

where ρ_{U_1} is the state realizing the maximum for a fixed U_1. A well-known consequence of the Schur-Weyl duality is that we can express the the first and second moments of Haar-random state as a weighted sum of identity and SWAP operators (Equations 3.9, 3.10). In particular this holds if we replace the Haar-random unitary with a 2-design. This allows us to write

$$\mathbb{E}_{U_1}\mathrm{Tr}\left[\mathcal{N}\circ U_1(\rho_{U_1})U_1^\dagger\mathcal{N}\left(\frac{I}{2}\right)\right]=2\,\mathrm{Tr}\left[\mathcal{N}\left(\frac{I}{2}\right)^2\right], \tag{4.32}$$

and proceeding as in (Section VII.A, [QFK+22]),

$$\mathrm{Tr}\left[\left(\mathcal{N}\circ U_1(\rho_{U_1})U_1^\dagger\right)^2\right]=\left(\frac{4}{3}-\frac{2}{3}\mathrm{Tr}\left[\rho_{U_1}^2\right]\right)\mathrm{Tr}\left[\mathcal{N}\left(\frac{I}{2}\right)^2\right]+\left(\frac{4}{3}\mathrm{Tr}\left[\mathcal{N}\left(\frac{I}{2}\right)^2\right]-\frac{2}{3}\right)\mathrm{Tr}\left[\mathscr{J}(\mathcal{N})^2\right] \tag{4.33}$$

$$=\frac{2}{3}\left(\mathrm{Tr}\left[\mathcal{N}\left(\frac{I}{2}\right)^2\right]+\mathrm{Tr}\left[\mathscr{J}(\mathcal{N})^2\right]\right), \tag{4.34}$$

where we used the fact that the maximizer of the induced 1-norm is a pure state and $\mathscr{J}(\mathcal{N})$ is the Choi-Jamiolkowski state of the channel $\mathcal{N}$, i.e., $\mathscr{J}(\mathcal{N}):=(\mathcal{N}\otimes I)(|\Omega\rangle\langle\Omega|)$. Putting all together and applying Jensen's inequality, we obtain

$$\mathbb{E}_{U_1}\|\mathcal{N}\circ U_1(\cdot)U_1^\dagger-\mathscr{E}\|_{1\to 1}\leq\sqrt{\frac{4}{3}\mathrm{Tr}\left[\mathscr{J}(\mathcal{N})^2\right]-\frac{2}{3}\mathrm{Tr}\left[\mathcal{N}\left(\frac{I}{2}\right)^2\right]}. \tag{4.35}$$

We can express the purities of $\mathcal{N}(I/2)$ and $\mathscr{J}(\mathcal{N})$ in terms of the parameters of the Pauli Transfer Matrix. By direct calculation (Examples 3.7, 3.9), we get

$$\mathrm{Tr}\left[J(\mathcal{N})^2\right]=\frac{1}{4}\sum_{P,Q\in\mathscr{P}_1}t_{P,Q}^2, \tag{4.36}$$

$$\mathrm{Tr}\left[\mathcal{N}\left(\frac{I}{2}\right)^2\right]=\frac{1}{2}\left(1+t_{I,X}^2+t_{I,Y}^2+t_{I,Z}^2\right). \tag{4.37}$$

and then,

$$\mathbb{E}_{U_1}\|\mathcal{N}\circ U_1(\cdot)U_1^\dagger-\mathscr{E}\|_{1\to 1}\leq\sqrt{\frac{1}{3}\sum_{P,Q\in\{X,Y,Z\}}t_{P,Q}^2}, \tag{4.38}$$

which yields the desired results. ■

We note that this procedure involves several coarse upper bounds and thus it does not improve the existing results on the contraction coefficient of the depolarizing and amplitude damping channels.

However, it produces a general and simple bound which can be of practical utility, since the exact calculation of the contraction coefficients is often not straightforward. This also demonstrates that the effective depth picture is not a feature of depolarizing or amplitude noise, but it may arise under a broad class of noise models.

4.4 A concise proof of noise-induced cost concentration

The most infamous instances of exponential concentration are the so-called *barren plateaus*, which corresponds to a dramatically flat landscape of the cost function, or, alternatively, to the gradient of the cost function being exponentially concentrated around zero with exponentially high probability. Note that two distinct expectations are involved in this definition .

First, we recall the cost function is itself defined as the expectation of an observable O with respect to the output state of a parametrized channel $\mathscr{C}_{\boldsymbol{\theta}}$.

$$C(\boldsymbol{\theta}) = \mathrm{Tr}\left[O\mathscr{C}_{\boldsymbol{\theta}}(\rho_0)\right], \tag{4.39}$$

where ρ_0 is the initial state of the circuit, which is typically $\rho_0 = |0^n\rangle\langle 0^n|$. The goal is then to find the value of the parameter $\boldsymbol{\theta}$ minimizing the cost function. A common avenue consists in performing the popular gradient-descent algorithm [PJSPP21], but analogous trainability issues arise also with alternative optimizers [ACC$^+$21].

Second, we note that the implementation of those optimizers requires an initial choice of $\boldsymbol{\theta}$, that is often set uniformly at random, giving rise to a further expectation.

Given this preliminary remarks, we provide the definition of exponentially concentrated cost function.

Definition 4.1 (Exponential concentration of cost function)**.** We say that a cost function C is exponentially concentrated if

$$\mathbb{V}_{\boldsymbol{\theta}}\left[C(\boldsymbol{\theta})\right] = 2^{-\Omega(n)}. \tag{4.40}$$

We remark that the variance can be expressed as $\mathbb{V}_{\boldsymbol{\theta}}\left[C(\boldsymbol{\theta})\right] = \mathbb{E}_{\boldsymbol{\theta}}\left[C(\boldsymbol{\theta})^2\right] - \mathbb{E}_{\boldsymbol{\theta}}\left[C(\boldsymbol{\theta})\right]^2$, thus upper bounding $\mathbb{E}_{\boldsymbol{\theta}}\left[C(\boldsymbol{\theta})^2\right]$ suffices for our scope. In addition, in absence of noise and under very general assumptions, we have that $\mathbb{E}_{\boldsymbol{\theta}}\left[C(\boldsymbol{\theta})^2\right] = 0$. Moreover, the bound on the variance can be translated to an high probability by means of Chebyshev's inequality. We now give the closely related notion of barren plateaus.

Definition 4.2 (Barren plateaus)**.** A cost function C exhibits barren plateaus if

$$\mathbb{V}_{\boldsymbol{\theta}}\left[\|\nabla_{\boldsymbol{\theta}} C\|_2\right] = 2^{-\Omega(n)}. \tag{4.41}$$

It is fruitful to write the observable O in the expression of the cost function as a sum of Pauli strings, $O = \sum_{P\in\mathscr{P}_n} c_P P$, where $c_P \neq 0$ for at most polynomially many Paulis. In the following lemma, we show that studying the behaviour of the Paulis is enough to prove that the cost function is exponentially concentrated.

Lemma 4.3 (Pauli concentration suffices)**.** *Let $\mathscr{S} \subseteq \mathscr{P}_n$ be of subset of Pauli strings of size $|\mathscr{S}| = k$, ν be a distribution over unitaries and $O = \sum_{P\in\mathscr{P}_n} c_P P$ an observable. Then*

$$\mathbb{E}_{U\sim\nu}\left[\mathrm{Tr}[OU\rho U^\dagger]^2\right] \leq k \cdot \sum_{P\in\mathscr{P}_n} c_P^2 \mathbb{E}_{U\sim\nu}\left[\mathrm{Tr}[PU\rho U^\dagger]^2\right]. \tag{4.42}$$

Moreover, if ν is invariant to left-hand multiplication of random Paulis, we have the the following identity

$$\mathbb{E}_{U\sim\nu}\left[\mathrm{Tr}[OU\rho U^\dagger]^2\right] = \sum_{P\in\mathscr{P}_n} c_P^2 \mathbb{E}_{U\sim\nu}\left[\mathrm{Tr}[PU\rho U^\dagger]^2\right]. \tag{4.43}$$

Proof. The first result is a consequence of the inequality $(\sum_{i=1}^k x_i)^2 \leq k\sum_{i=1}^k x_i^2$, which is a special case of Minkowski's inequality. As for the second result, we can rearrange the expression as follows

$$\mathbb{E}_{U\sim\nu}\left[\mathrm{Tr}[U^\dagger OU\rho]^2\right] = \sum_{P,Q\in\mathscr{P}_n} c_P c_Q \mathbb{E}_{U\sim\nu}\,\mathrm{Tr}[U^\dagger PU\rho]\,\mathrm{Tr}[U^\dagger QU\rho] \tag{4.44}$$

$$= \sum_{P,Q\in\mathscr{P}_n} c_P c_Q \mathbb{E}_{U\sim\nu}\,\mathrm{Tr}\left[U^\dagger PU \otimes U^\dagger QU\rho^{\otimes 2}\right] = \sum_{P\in\mathscr{P}_n} c_P^2\,\mathrm{Tr}[PU\rho U^\dagger]^2, \tag{4.45}$$

where the last identity is a consequence of Lemma 3.2. ■

Lemma 4.4. *For $\rho \in \mathscr{S}_n, P \in \mathscr{P}_n$, we have*

$$\mathrm{Tr}[P\rho]^2 \leq 2D_2(\rho\|I/2^n)$$

Proof. Recall the characterizations of the purity:

$$2^{-n+D_2(\rho\|I/2^n)} = \mathrm{Tr}[\rho^2] = \frac{1}{2^n}\sum_P \mathrm{Tr}[P\rho]^2 = \frac{1}{2^n} + \frac{1}{2^n}\sum_{P\neq I}\mathrm{Tr}[P\rho]^2. \tag{4.46}$$

Hence for $P \neq I$,

$$\mathrm{Tr}[P\rho]^2 + 1 \leq \sum_{P\neq I}\mathrm{Tr}[P\rho]^2 + 1 \leq 2^{D_2(\rho\|I/2^n)}. \tag{4.47}$$

Since $\frac{x}{1+x} \leq \log(1+x)$ and $\mathrm{Tr}[P\rho]^2 \leq 1$,

$$\frac{\mathrm{Tr}[P\rho]^2}{2} \leq \log(\mathrm{Tr}[P\rho]^2 + 1) \leq D_2(\rho\|I/2^n). \tag{4.48}$$

■

Remark that if ρ is a state produced by a circuit interspersed with layers of unital noise, the $D_2(\rho\|I/2^n)$ goes to zero exponentially fast. In particular, for local Pauli noise we have at depth L:

$$\mathrm{Tr}[P\rho]^2 \leq 2n(1-p)^{2L}. \tag{4.49}$$

Combined with Lemma 3.2, this readily gives the desired result.

Theorem 4.1. *Let* $O = \sum_{P \in \mathscr{S} \subseteq \mathscr{P}_n} c_P P$ *with* $\sum_{P \in \mathscr{P}_n} c_P^2 = \mathrm{poly}(n), |\mathscr{S}| = \mathrm{poly}(n)$ *and let* ρ *the output of depth-L circuit interspersed with local Pauli channels* $\mathcal{N}_p^{(\mathrm{Pauli})^{\otimes n}}$. *Then the cost function satisfies the following concentration inequality,*

$$\mathrm{Tr}[O\rho]^2 \leq \mathrm{poly}(n)(1-p)^2 L, \tag{4.50}$$

which for $L \geq c \cdot n$ *for a sufficiently large constant c yields*

$$\mathrm{Tr}[O\rho]^2 \leq 2^{-\Omega(n)}. \tag{4.51}$$

CHAPTER 5

EXPONENTIAL CONCENTRATION AND LACK THEREOF IN QUANTUM KERNEL METHODS

Hofstadter's Law: It always takes longer than you expect, even when you take into account Hofstadter's Law.

-Douglas Hofstadter

Near-term quantum devices are plagued by noise. The presence of multiple sources of errors during the implementation of quantum algorithms brings catastrophic effect, voiding the utility of most algorithms that could achieve exponential quantum advantage on fault-tolerant devices. Given this scenario, variational quantum algorithms are reputedly one of the most promising approach on near-term devices, due to their supposed robustness to hardware errors. [Pre18a, SKCC20, RCA+22]. Yet, they are not immune to a number of trainability issues, which can arise due to different causes. Broadly speaking, variational quantum algorithms are based on parametrized families (or ensembles) of unitaries, which we denote by $\{U(\boldsymbol{\theta})\}_{\boldsymbol{\theta}\in\Theta}$ for a given parameter space Θ. Given a distribution $\mathscr{D}$ over Θ, we also have a conditional distribution over $\mathscr{U}_n$. With an abuse of notation, we denote by $\{U(\boldsymbol{\theta})\}_{\boldsymbol{\theta}\in\Theta}$ both the family of unitaries and the conditional

distribution over $\mathscr{U}_n$. When not otherwise stated, we will implictly assume that the parameter $\boldsymbol{\theta}$ is sampled uniformly at random over Θ.

The ensemble $\{U(\boldsymbol{\theta})\}_{\boldsymbol{\theta}\in\Theta}$ plays a twofold role in variational quantum algorithm. On one hand we need such ensemble to be "rich" enough to hold the potential of a quantum advantage. This can be ensured for instance from the fact that sampling from $\{U(\boldsymbol{\theta})\}_{\boldsymbol{\theta}\in\Theta}$ is classically hard on average. On the other hand, an overly expressive ensemble may render the training process extremely complex, as we will detail in the following of this Chapter. While lack of trainability comes in several flavors, most of these instances are facets of the *exponential concentration* phenomenon. Particularly, we individuate two main families of (exponential) concentration:

- *ensemble*-induced concentration arises due to the properties of the ensemble $\{U(\boldsymbol{\theta})\}_{\boldsymbol{\theta}\in\Theta}$, for instance its closeness to a (global) 2-design, or the output state being highly entangled [MBS$^+$18, MKW21, CSV$^+$21, TWH22, HSCC22]. In this Chapter we will discuss the implications of other properties of the ensemble on the trainability of quantum kernel methods.

- *noise*-induced concentration arises due to the presence of sources of noise in the circuit implementing the ideal unitary $U(\boldsymbol{\theta})$. The catastrophic impact of unital noise on trainability has been previously investigated in [WFC$^+$21, TWH22]. In this Chapter, we will show that a more realistic noise model, accounting for a non-unital perturbation, leads to qualitatively different scenarios, allowing the trainability of certain kinds of variational quantum algorithms.

A possible avenue to mitigate ensemble-induced concentration is to envision a different initialization strategy for the parameter $\boldsymbol{\theta}$, thus implementing a different distribution over unitaries $\{U(\boldsymbol{\theta})\}_{\boldsymbol{\theta}\in\Theta}$ [GWOB19]. Conversely, noise-induced concentration may arise independently of the distribution over $\{U(\boldsymbol{\theta})\}_{\boldsymbol{\theta}\in\Theta}$, making it a more fearsome threat to the training process. In fact, the presence of noise-induced concentration is closely related to the hardness of error-mitigation, which can be a computationally unfeasible task even at very shallow depth [QFK$^+$22].

Exponential concentration. To understand why exponential concentration prevents trainability, assume that two random variables X and Y are exponentially concentrated around the same value μ, with an exponentially high probability. Then there exists $\delta \in 2^{O(-n)}$ such that

$$\Pr[|X-\mu|\geq\delta], \Pr[|Y-\mu|\geq\delta] \in 2^{-\Omega(n)}, \tag{5.1}$$

and therefore

$$\Pr\left[|X-Y|\geq 2\delta\right] \in 2^{-\Omega(n)}. \tag{5.2}$$

For instance, X and Y may represent the norm of the gradient of a cost function evaluated for two different parameters, and one could be interested in finding the parameter optimizing such norm. In the most general case, we need $\Theta\left(\varepsilon^{-2}\right)$ samples of a random variable to recover its value up to

additive error ε. Therefore, distinguishing X from Y is information-theoretically hard, as it would require exponentially many samples.

Whereas exponential concentration is ubiquitous in variational quantum algorithms, throughout this chapter we focus on a particular family of algorithms based on quantum kernels. Despite some encouraging positive results concerning their performance [LAT21, HCT+19], quantum kernel methods are subject to a number of trainability issues. Specifically, the authors of [TWH22] showed that their trainability can be compromised by a number of factors, including unital noise, expressibility and entanglement. Significantly, previous research has not considered the inclusion of non-unital perturbations in the noise model. This leads us to a fundamental question:

Question 1. *What is the impact of more realistic sources of noise on variational quantum algorithms?*

We will address this question by examining its effects on both fidelity and the projected quantum kernels. Furthermore, we will revisit the noiseless scenario, seeking out additional sources of untrainability.

Our contributions. On one hand, we will provide novel concentration bounds for quantum kernels, effectively constraining the potential quantum advantage, and thereby contributing to a more comprehensive understanding of the limitations of these methods.

- In the noiseless case, we show that fidelity quantum kernels exhibit exponential concentration if the associated ensemble $\{U_x\}_{x\in\mathcal{X}}$ is invariant to right-hand multiplication of random Paulis or if its scaled collision probability is constant.

- When the circuit is interspersed by local noise, either unital or non-unital, we demonstrate that the fidelity quantum kernels incur in exponential concentration at any depth, significantly tightening prior results established in the unital noise regime.

On the other hand, we find that projected quantum kernels behave qualitatively differently under non-unital noise.

- Projected quantum kernels do not experience exponential concentration if the circuit is interspersed with local non-unital noise. This is in stark contrast with all prior untrainability results on noisy variational quantum algorithms. In particular, our results holds for a mixture of amplitude damping noise and depolarizing noise, with rates respectively q and p, provided that $q = 1/\operatorname{poly}(n)$.

Our result indicates that non-unital perturbations should be taken into account in the analysis of variational quantum algorithms, and that the model of local Pauli noise may be excessively pessimistic. Moreover, we provide a further conceptual contribution, by conjecturing the existence of an "effective depth" noisy circuit. Given a random noisy circuit with super-logarithmic depth, we conjecture that only the last portion of the circuit of depth $O(\log n)$ bears a significant influence

on the output state. While we do not prove this statement in the most general case, we provide a proof for the high-noise regime, which holds under the assumption that p or q exceeds some fixed constant threshold. Our argument hinges on the contraction coefficients of the quantum Wasserstein distance of order 1.

5.1 The model

Consider an n-qubit data-embedding channel Φ_x parametrized by a point $x \in \mathscr{X}$, so that

$$\rho(x) = \Phi_x(\rho_0), \tag{5.3}$$

where ρ_0 is the initial state of the circuit, usually set as $\rho_0 = |0^n\rangle\langle 0^n|$. A kernel $\kappa : \mathscr{X} \times \mathscr{X} \to \mathbb{R}^+$ is a similarity measure between pair of points $x, y \in \mathscr{X}$. In particular, quantum kernels rely on the quantum embedding scheme described in the Equation 5.3 above. We consider the fidelity quantum kernel [HCT+19, Sch21], defined as

$$\kappa^{FQ}(x, y) = \mathrm{Tr}[\rho(x)\rho(y)] \tag{5.4}$$

The projected quantum kernel [HBM+21] is defined as

$$\kappa^{PQ}(x, y) = \exp\left(-\gamma \sum_{k=1}^{n} \|\rho_k(x) - \rho_k(y)\|_2^2\right), \tag{5.5}$$

where $\rho_k(x) = \mathrm{Tr}_{\overline{k}}\rho(x)$ is the reduced density matrix of the k-th qubit. Kernel-based learning methods are notable for their capacity to transform data from the original space $\mathscr{X}$ into a higher-dimensional feature space, which in our case coincides with the a 2^n-dimensional Hilbert space. In this new feature space, inner products are computed, enabling the training of decision boundaries like support vector machines, as explained in reference [Sch21].

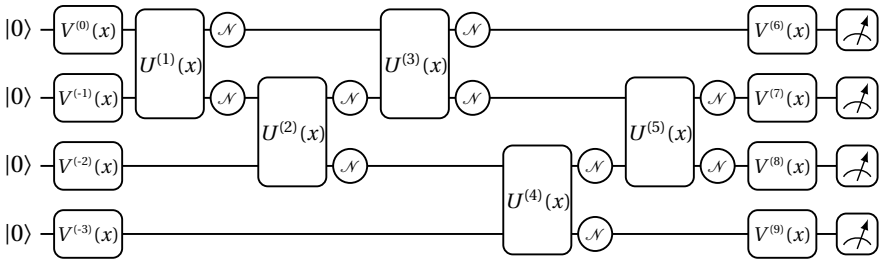


Figure 5.1: Example of a noisy quantum circuit on $n = 4$ qubits with two-qubit and single-qubit gates, parametrized by the input vector $x \in \mathscr{X}$. A pair of single-qubit noise channels $\mathscr{N}$ follow each two-qubit gate. The circuit begins and ends with a layer of noiseless single-qubit gates. This model contains minimal assumptions on the circuit architecture, and it coincides with the one adopted in [DHJB21].

Kernel-based supervised learning

To better suite our results, we sketch how kernel methods can be used to perform supervised learning. We consider a training set of labelled inputs $\mathscr{S} = \left\{x^{(i)}, f\left(x^{(i)}\right)\right\}_{i\in[m]}$, where $f(\cdot)$ is some unknown function that we want to learn. Thus our goal is to find a function h approximating f. Thanks to the Representer Theorem (see, for instance, [SSBD14], Theorem 16.1), the optimal function can be expressed as follows

$$h(z) = \sum_{i=1}^{m} a_i \kappa\left(x^{(i)}, z\right), \tag{5.6}$$

where the $\boldsymbol{a} = (a_1, a_2, \ldots, a_m)$ is a vector of parameters to be optimized with respect to a suitable loss function.

Then, to enable the implementation of kernel methods, it is necessary to estimate the Gram matrix. This matrix, denoted as $\mathscr{G}$, comprises the kernels derived from pairs of inputs within the training set $x^{(1)}, x^{(2)}, \ldots, x^{(m)}$, and is defined as:

$$\forall\, i \in [m] : \mathscr{G}[i,j] = \kappa\left(x^{(i)}, x^{(j)}\right) \tag{5.7}$$

We recall that kernels exhibit exponential concentration with respect to a distribution $\mathscr{D}$ over $\mathscr{X}$, if there exists a real number $\mu \in \mathbb{R}$ and a value $\delta \in 2^{-O(n)}$ such that

$$\Pr_{x,y\sim\mathscr{D}}[|\kappa(x,y) - \mu| \geq \delta] \in 2^{-\Omega(n)}. \tag{5.8}$$

In this case, all the entries of the Gram matrix are exponentially close to μ with exponentially high probability, making the optimization of the vector $\boldsymbol{a}$ an information-theoretically hard task.

5.1.0.1 Assumption on the training data distribution

Assume that each point in the training set is sampled from a distribution $\mathscr{D} : \mathscr{X} \to [0,1]$ and denote by ν' the corresponding induced distribution over quantum channels. Then we make the following assumptions over ν':

1. each layer is invariant under post-processing by a layer of single-qubit Clifford gates;
2. moreover, the circuit is ended by a layer of single-qubit Clifford gates sampled uniformly at random.

The second and third assumptions will play a pivotal role in the proof of absence of exponential concentration for the projected quantum kernels. We also remark that these assumptions could be further relaxed, since our computation only involve (up to) fourth moments.

5.1.1 A technical lemma

Prior to delving into the analysis of quantum kernels, we give the following technical tool, which will be employed in the following.

Table 5.1: **Fidelity quantum kernel**

Noise model	**Exponential concentration**
Noiseless [Our work],[TWH22]	Yes
Unital noise at linear depth [TWH22]	Yes
Unital and non-unital noise at any depth depth [Our work]	Yes

Table 5.1 resumes the exponential concentration results for the fidelity quantum kernels, which arise both for the noiseless case and for the noisy case. Our work extends the analysis to non-unital sources of noise, and demonstrates that exponential concentration arises at any depth.

Table 5.2: **Projected quantum kernel**

Noise model	**Exponential concentration**
Noiseless [TWH22]	Yes
Unital noise at linear depth [TWH22]	Yes
Non-unital noise at any depth [Our work]	No

Table 5.2 illustrates that if a state is prepared by a non-unital noisy-random quantum circuit, the expectation value of projected quantum kernels does not exhibits exponential concentration at any depth around a fixed value. This stands in stark contrast to the noiseless and unital noise regimes.

Lemma 5.1 (Adapted from ([TWH22], Theorem 3))**.** *The following inequality holds,*

$$|1-\kappa^{PQ}(x,y)| \leq \gamma \sum_{k=1}^{n} \left(2\left\|\rho_k(x)-\frac{I}{2}\right\|_2^2 + 2\left\|\rho_k(y)-\frac{I}{2}\right\|_2^2\right). \tag{5.9}$$

Proof. We can rearrange the projected quantum kernel as follows:

$$|1-\kappa^{PQ}(x,y)| = \left|1-\exp\left(-\gamma\sum_{k=1}^{n}\|\rho_k(x)-\rho_k(y)\|_2^2\right)\right| \tag{5.10}$$

$$\leq \gamma \sum_{k=1}^{n} \|\rho_k(x)-\rho_k(y)\|_2^2 \tag{5.11}$$

$$\leq \gamma \sum_{k=1}^{n} \left(\left\|\rho_k(x)-\frac{I}{2}\right\|_2 + \left\|\rho_k(y)-\frac{I}{2}\right\|_2\right)^2 \tag{5.12}$$

$$\leq \gamma \sum_{k=1}^{n} \left(2\left\|\rho_k(x)-\frac{I}{2}\right\|_2^2 + 2\left\|\rho_k(y)-\frac{I}{2}\right\|_2^2\right), \tag{5.13}$$

where the first inequality is due to the standard inequality $1-e^{-t} \leq t$, the second inequality is due to the triangle inequality, the third inequality is due to the fact that $(s+t)^2 \leq 2s^2+2t^2$. ■

5.2 Ensemble-induced concentration

We now provide several examples of exponential concentration arising due to properties of the ensemble $\{U_x\}_{x\in\mathcal{X}}$. For all $x \in \mathcal{X}$, let $\rho(x) = U_x|0^n\rangle\langle 0^n|U_x^\dagger$. In the related work from [TWH22], the

authors showed that exponential concentration of quantum kernels can be caused from the closeness to a global 2-design. Moreover, the fidelity quantum kernel incurs in exponential concentration when the encoding circuit is the product of local unitaries, and the projected quantum kernels is exponentially concentrated when the state $\rho(x)$ is highly entangled. We first notice that the scrambling properties of random quantum circuits can be related the concentration of the projected quantum kernel.

Proposition 5.1 (Scrambling-induced concentration)**.** *If $\{U_x\}_{x\in\mathcal{X}}$ is an $(\varepsilon,1)$-approximate scrambler, then*

$$|1-\kappa^{PQ}(x,y)| \leq 4\gamma n\varepsilon. \tag{5.14}$$

Proof. The result readily follows by combining Lemma 5.1 with the definition of approximate scrambler (Definition 3.5). ■

Theorem 5.1 (Concentration induced by Pauli invariance)**.** *Let $x \in \mathcal{X}$ such that the ensemble $\{U_x\}_{x\in\mathcal{X}}$ is invariant under right-hand multiplication of Pauli in $\{I,X\}^{\otimes n}$. Then the fidelity quantum kernel is exponentially concentrated:*

$$\mathbb{E}_{x,y}\kappa^{FQ}(x,y) \leq \frac{1}{2^n}. \tag{5.15}$$

Proof. The fidelity quantum kernels can be rearranged as follows

$$\kappa^{FQ}(x,y) = \mathrm{Tr}[\rho(x)\rho(y)] = \mathrm{Tr}[U_x|0^n\rangle\langle 0^n|U_x^\dagger U_y|0^n\rangle\langle 0^n|U_y^\dagger] \tag{5.16}$$

$$= |\langle 0^n|U_x^\dagger U_y|0^n\rangle|^2. \tag{5.17}$$

By Pauli invariance of the ensemble $\{U_x\}_{x\in\mathcal{X}}$, we have,

$$\forall \ell \in \{0,1\}^n : \mathbb{E}_{x,y}|\langle 0^n|U_x^\dagger U_y|0^n\rangle|^2 = \mathbb{E}_{x,y}|\langle \ell|U_x^\dagger U_y|0^n\rangle|^2 \tag{5.18}$$

Conservation of probabilities implies

$$2^n\mathbb{E}_{x,y}|\langle 0^n|U_x^\dagger U_y|0^n\rangle|^2 = \sum_{\ell\in\{0,1\}^n}|\langle \ell|U_x^\dagger U_y|0^n\rangle|^2 \tag{5.19}$$

$$\leq \sum_{\ell\in\{0,1\}^n}|\langle \ell|U_x^\dagger U_y|0^n\rangle| = 1 \tag{5.20}$$

Thus rearranging gives the desired result:

$$\mathbb{E}_{x,y}\kappa^{FQ}(x,y) = \mathbb{E}_{x,y}|\langle 0^n|U_x^\dagger U_y|0^n\rangle|^2 \leq \frac{1}{2^n}. \tag{5.21}$$

■

This result demonstrate that exponential concentration may arise even for very simple encodings. For instance, if $y=(\theta_1,\theta_2,\ldots,\theta_n)$ is sampled uniformly at random from $[0,2\pi)^n$, it's easy to see that the following encoding satisfies the hypothesis Theorem 5.1:

$$\rho(y) = \bigotimes_{k\in[n]} R_X(\theta_k)|0\rangle$$

An analogous result for a similar product encoding was also given in ([TWH22], Proposition 1). However, our approach provides further insight by connecting the exponential concentration with Pauli invariance. Moreover, we also argue that the assumption on Pauli invariance can be relaxed to the following condition,

$$\forall \ell \in \{0,1\}^n : \mathbb{E}_{x,y} |\langle 0^n | U_x^\dagger U_y | 0^n \rangle|^2 \leq \beta \cdot \mathbb{E}_{x,y} |\langle \ell | U_x^\dagger U_y | 0^n \rangle|^2, \tag{5.22}$$

for $\beta/2^n = 2^{\Omega(-n)}$.

In addition, we show that ensembles with low collision probability are also more prone to exponential concentration. This result is consistent with the literature on barren plateaus, and particularly with [Nap22], which drew a connection between anticoncentration and barren plateaus for global cost function. We recall that an ensemble anticoncentrates if its scaled collision probability is at most constant, as discussed in details in Section 3.2.3.

Theorem 5.2 (Concentration induced by low collision probability)**.** *Let x, y two random points sampled independently and uniformly at random from $\mathscr{X}$, and denote by ν the distribution of the ensemble $\left\{U_x^\dagger U_y\right\}_{x,y\in\mathscr{X}}$. Let $\mathscr{Z}(\nu)$ be the scaled collision probability of the ensemble ν, i.e.*

$$\mathscr{Z}(\nu) := 2^n \cdot \mathbb{E}_{\rho\sim\nu} \left[\sum_{i\in\{0,1\}^n} \mathrm{Tr}[\rho |i\rangle\langle i|]^2 \right] - 1. \tag{5.23}$$

Then the fidelity quantum kernel satisfies the following inequality,

$$\mathbb{E}_{x,y} \kappa^{FQ}(x,y) \leq \frac{\mathscr{Z}(\nu)+1}{2^n}. \tag{5.24}$$

Proof. The fidelity quantum kernels can be rearranged as follows

$$\kappa^{FQ}(x,y) = \mathrm{Tr}[\rho(x)\rho(y)] = \mathrm{Tr}[U_x|0^n\rangle\langle 0^n|U_x^\dagger U_y|0^n\rangle\langle 0^n|U_y^\dagger] \tag{5.25}$$

$$= |\langle 0^n|U_x^\dagger U_y|0^n\rangle|^2. \tag{5.26}$$

By definition of scaled collision probability we have:

$$\mathscr{Z}(\nu) = 2^n \mathbb{E}_{x,y} \sum_{\ell\in\{0,1\}^n} |\langle \ell|U_x^\dagger U_y|0^n\rangle|^2 - 1 \geq 2^n \mathbb{E}_{x,y} |\langle 0|U_x^\dagger U_y|0^n\rangle|^2 - 1 \tag{5.27}$$

And rearranging yields

$$\mathbb{E}_{x,y} \kappa^{FQ}(x,y) \leq \frac{\mathscr{Z}(\nu)+1}{2^n} \tag{5.28}$$

■

Several upper bounds on the value of the collision probability for several families of random circuits can be found in the previous literature, such as [DHJB22]. In particular, they showed that random circuits based on the 1D or the complete-graph architecture anticoncentrate at logarithmic depth, and conjectured the same to hold also for all regularly connected architectures. Moreover, for all these family of circuits, the RHS of Eq. 5.24, i.e. the collision probability, is exponentially small even at constant depth.

5.3 Noise-induced concentration

As previously noted in [TWH22], both fidelity and projected quantum kernels can suffer from noise-induced exponential concentration under the action of unital noise. Here, we improve their bound in a twofold way for the case of the fidelity quantum kernel:

- first, our upper bound on fidelity quantum kernels converges to an exponentially small value *double* exponentially fast in the number of layer, whereas the prior bound predicts an exponentially fast convergence rate;
- second, we show that this convergence is attained even under the action of non-unital noise, provided that the unital component of the noise "dominates" the non-unital one.

First of all, we recall the definition of the following noise channel, obtained by composing the depolarizing channel of noise rate p and the amplitude damping channel of noise rate q.

$$\mathcal{N}_{p,q}^{(\mathrm{dep,amp})} := \mathcal{N}_p^{(\mathrm{dep})} \circ \mathcal{N}_q^{(\mathrm{amp})} \tag{5.29}$$

$$\mathcal{N}_{q,p}^{(\mathrm{amp,dep})} := \mathcal{N}_q^{(\mathrm{amp})} \circ \mathcal{N}_p^{(\mathrm{dep})} \tag{5.30}$$

Our result follows as a simple consequence of Corollary 4.1.

Proposition 5.2 (Noise-induced concentration)**.** *Let Φ_x, Φ_y two noisy circuits interspersed by m layers of local noise, either of the form $\mathcal{N}_{p,q}^{(\mathrm{dep,amp}),\otimes n}$ or $\mathcal{N}_{q,p}^{(\mathrm{amp,dep}),\otimes n}$. Denote by $\rho(x) = \Phi_x(\rho_0)$ and $\rho(y) = \Phi_y(\rho_0)$ the output states of the noisy circuits. Then the fidelity quantum kernel $\kappa^{FQ}(x,y)$ satisfies the following upper bound.*

$$\kappa^{FQ}(x,y) \leq 2^{n(\delta_m - 1)}, \tag{5.31}$$

$$|1 - \kappa^{PQ}(x,y)| \leq 2\gamma n^2 \delta_m \tag{5.32}$$

where $\delta_m = (1-p)^{2m} + q\frac{1-(1-p)^{2m}}{2p-p^2}$.

Proof. The Cauchy-Schwarz inequality implies that the fidelity quantum kernel can be upper bounded by the purities of the output states:

$$\mathrm{Tr}[\rho(x)\rho(y)] \leq \sqrt{\mathrm{Tr}\left[\rho(x)^2\right]\mathrm{Tr}\left[\rho(y)^2\right]}. \tag{5.33}$$

As for the projected quantum kernel, we have:

$$\left\| \rho_k(x) - \frac{I}{2} \right\|_2^2 = \mathrm{Tr}[\rho_k(x)^2] - 2\frac{\mathrm{Tr}[\rho_k]}{2} + \frac{\mathrm{Tr}[I]}{2^2} \tag{5.34}$$

$$= \mathrm{Tr}[\rho_k(x)^2] - \frac{1}{2} \leq \frac{1}{2}\left(2^{D_2(\rho\|I/2^n)} - 1\right) \leq \frac{D_2(\rho\|I/2^n)}{2}. \tag{5.35}$$

$$|1 - \kappa^{PQ}(x,y)| \leq 2\gamma n D_2(\rho\|I/2^n). \tag{5.36}$$

Thus the desired results follows by invoking Corollary 4.1. ■

We emphasize that for $q = 0$, our bound predicts that the kernel $\kappa^{FQ}(x, y)$ is at most $2^{-n(2p-p^2)} = 2^{-\Omega(n)}$, even after a single layer of noise, whereas ([TWH22], Theorem 3) only predicts that $|\kappa^{FQ}(x, y) - 1/2^n| \leq (1-p)^2 = \Theta(1)$. Moreover, for constant $q > 0$, $\kappa^{FQ}(x, y) = 2^{-\Omega(n)}$ at depth m as soon as $\delta_m < 1$. Crucially, δ_m converges exponentially fast in m to its limiting value $q/(2p - p^2)$, which makes the convergence of the upper bound in Proposition 5.2 double exponentially fast.

On the other, the upper bound for the projected quantum kernel recovers the one in [TWH22] for $q = 0$, and moreover predicts exponential concentration at linear depth if $q = 2^{\Omega(-n)}$. We emphasize that, for larger values of q, the upper bound does not predict exponential concentration. This suggests that projected kernels are less prone to noise-induced concentration, especially in the non-unital regime. In the next section, we will support this intuition by upper bounding $\kappa^{PQ}(x, y)$, which is equivalent to lower bounding $|1 - \kappa^{PQ}(x, y)|$.

5.4 Absence of exponential concentration for the projected quantum kernel under non-unital noise

We will now show that projected kernels behave in a fundamentally different way under the action of non-unital noise. In this section, we make use of the notation $\|\boldsymbol{t}\|_2^2 := t_{I,X}^2 + t_{I,Y}^2 + t_{I,Z}^2$.

Theorem 5.3 (Variance)**.** *We have,*

$$\mathrm{Var}_{x,x'} \sum_{k=1}^{n} \left[\|\rho_k(x) - \rho_k(x')\|_2^2 \right] \geq \Omega(n \|\mathbf{t}\|_2^4). \tag{5.37}$$

Proof. For the scope of this proof, we only need to consider the last two layers of unitaries and the last layer of local noise. In particular, we will use the fact that the last layer is a tensor product of random single-qubit Cliffords, and the second-to-last layer is invariant under post-processing by tensor products of random single-qubit Cliffords.

Thus, we can re-express the reduced states $\rho_k(x)$ and $\rho_k(x')$ as follows

$$\rho_k(x) = V_x \mathcal{N} \left(\widetilde{V}_x \hat{\rho}(x) \widetilde{V}_x^\dagger \right) V_x^\dagger, \tag{5.38}$$

$$\rho_k(x') = V_{x'} \mathcal{N} \left(\widetilde{V}_{x'} \hat{\rho}(x') \widetilde{V}_{x'}^\dagger \right) V_{x'}^\dagger, \tag{5.39}$$

where $V_x, V_{x'}, \widetilde{V}_x$ and $\widetilde{V}_{x'}$ are single-qubit Cliffords. We denote by $\hat{\rho}(x)$ and $\hat{\rho}(x')$ the states obtained after the action of the first $L-1$ layers of the noisy circuits and after tracing out all the qubits except from the k-th.

Throughout this proof, we will consider the conditional expectation with respect to the following event:

$$A = \{V_x = V_{x'}\}, \tag{5.40}$$

that is, we condition upon the last gate acting on the k-th qubit being the same for both the classical inputs x and x'. Since V_x and $V_{x'}$ are single-qubit Cliffords sampled uniformly at random from Cl(1), the event A happens with constant probability: $\Pr[A] = |\mathrm{Cl}(1)|^{-1} = \Theta(1)$.

First, we notice that the expected purities of $\rho_k(x)$ and $\rho_k(x')$ do not change if we condition on the event A:

$$\mathbb{E}_x \operatorname{Tr}[\rho_k(x)^2] = \mathbb{E}_{x'} \operatorname{Tr}[\rho_k(x')^2] = \mathbb{E}_x\{\operatorname{Tr}[\rho_k(x)^2]|A\} = \mathbb{E}_{x'}\{\operatorname{Tr}[\rho_k(x')^2]|A\}. \tag{5.41}$$

Moreover, since the Clifford group forms a 1-design, the expected overlap takes the value

$$\mathbb{E}_{x,x'} \operatorname{Tr}[\rho_k(x)\rho_k(x')] = \operatorname{Tr}\left[\left(\frac{I}{2}\right)^2\right] = \frac{1}{2}. \tag{5.42}$$

Similarly, conditioning on A we obtain

$$\mathbb{E}_{x,x'}\{\operatorname{Tr}[\rho_k(x)\rho_k(x')]|A\} = \tag{5.43}$$

$$\mathbb{E}_{x,x'} \operatorname{Tr}[\mathcal{N}\left(\widetilde{V}_x\hat{\rho}(x)\widetilde{V}_x^\dagger\right)\mathcal{N}\left(\widetilde{V}_{x'}\hat{\rho}(x')\widetilde{V}_{x'}^\dagger\right)] \tag{5.44}$$

$$= \operatorname{Tr}\left[\mathcal{N}\left(\frac{I}{2}\right)^2\right] = \frac{1+\|\mathbf{t}\|_2^2}{2}. \tag{5.45}$$

Therefore, the difference between $\mathbb{E}_{x,x'}\left\{\|\rho_k(x)-\rho_k(x')\|_2^2|A\right\}$ and $\mathbb{E}_{x,x'}\|\rho_k(x)-\rho_k(x')\|_2^2$ can be expressed as

$$\mathbb{E}_{x,x'}\{|\rho_k(x)-\rho_k(x')\|_2^2|A\} - \mathbb{E}_{x,x'}\|\rho_k(x)-\rho_k(x')\|_2^2 \tag{5.46}$$

$$= 2\mathbb{E}_{x,x'}\{\operatorname{Tr}[\rho_k(x)\rho_k(x')]|A\} - 2\mathbb{E}_{x,x'} \operatorname{Tr}[\rho_k(x)\rho_k(x')] \tag{5.47}$$

$$= 2\operatorname{Tr}\left[\mathcal{N}\left(\frac{I}{2}\right)^2\right] - 2\operatorname{Tr}\left[\left(\frac{I}{2}\right)^2\right] = \|\mathbf{t}\|_2^2 \tag{5.48}$$

This immediately translates into a lower bound on the variance of $\|\rho_k(x)-\rho_k(x')\|_2^2$ with respect to the random gates $V_x, V_{x'}, \widetilde{V}_x, \widetilde{V}_{x'}$. Let $\mu := \mathbb{E}_{x,x'}\left[\|\rho_k(x)-\rho_k(x')\|_2^2\right]$. We have

$$\operatorname{Var}_{V_x,V_{x'},\widetilde{V}_x,\widetilde{V}_{x'}}\left[\|\rho_k(x)-\rho_k(x')\|_2^2\right] \tag{5.49}$$

$$= \mathbb{E}_{V_x,V_{x'},\widetilde{V}_x,\widetilde{V}_{x'}}\left[(\mu - \|\rho_k(x)-\rho_k(x')\|_2^2)^2\right] \tag{5.50}$$

$$= \Pr[A]\left(\mu - \mathbb{E}[\|\rho_k(x)-\rho_k(x')\|_2^2|A]\right)^2 + \Pr[\overline{A}]\left(\mu - \mathbb{E}[\|\rho_k(x)-\rho_k(x')\|_2^2|\overline{A}]\right)^2 \tag{5.51}$$

$$\geq \Pr[A]\left(\mu - \mathbb{E}[\|\rho_k(x)-\rho_k(x')\|_2^2|A]\right)^2 = \Pr[A]\|\mathbf{t}\|_2^4 \geq \Omega(\|\mathbf{t}\|_2^4). \tag{5.52}$$

We can easily lower bound the variance of the sum of all the terms, i.e., $\sum_{k=1}^n\left[\|\rho_k(x)-\rho_k(x')\|_2^2\right]$,

$$\operatorname{Var}_{x,x'}\sum_{k=1}^n\left[\|\rho_k(x)-\rho_k(x')\|_2^2\right] \geq \operatorname{Var}_{V_x,V_{x'},\widetilde{V}_x,\widetilde{V}_{x'}}\sum_{k=1}^n\left[\|\rho_k(x)-\rho_k(x')\|_2^2\right] \tag{5.53}$$

$$\geq n\cdot\min_{k\in[n]}\operatorname{Var}_{V_x,V_{x'},\widetilde{V}_x,\widetilde{V}_{x'}}\|\rho_k(x)-\rho_k(x')\|_2^2 \geq \Omega(n\|\mathbf{t}\|_2^4) \tag{5.54}$$

■

The lower bound computed above can be transferred to the projected quantum kernels by a McLaurin's expansion with a first-order approximation. Whenever the term $\gamma\sum_{k=1}^n\|\rho(x)_k-\rho(x')_k\|_2^2 \ll 1$, we have

$$\kappa^{PQ}(x,x') \approx 1-\gamma\sum_{k=1}^n\|\rho(x)_k-\rho(x')_k\|_2^2, \tag{5.55}$$

therefore,

$$\operatorname{Var}_{x,x'}\kappa^{PQ}(x,x') \gtrsim \Omega\left(n\gamma^2\|\mathbf{t}\|_2^4\right). \tag{5.56}$$

5.5 The "effective depth" noisy circuit

Theorem 5.3 shows that projected quantum kernels do not exhibit exponential concentration at any depth, provided that the circuit layers are interspersed by local noise with a sufficiently strong non-unital component. However, it should be noted that this result does not bring any guarantee on the performance of the quantum kernel methods. As exponential concentration implies that estimating the entries of the Gram matrix requires exponentially many samples, the absence of such concentration is merely a necessary condition for trainability. In particular, we emphasize that the proof of Theorem 5.3 involves only the last two layers of the encoding circuit. This shows that, despite the presence of noise, the Gram matrix is sensitive to the classical information encoded in the final layers of the circuits. However, the information encoded in the early layers of the circuit could still be irremediably corrupted by the action of the noise.

In the following we will provide some insights in favor of the existence of an "effective depth" circuit, i.e., that the output is highly influenced by the last m layers of the circuit, where m is a function of the noise strength, while the initial ones have a marginal impact. In particular, no matter the depth of circuit, the "effective depth" will be bounded by $\mathcal{O}(\log n)$ if the noise is constant.

5.5.1 A Wasserstein distance approach

While we will not provide a general proof of the "effective depth" picture, we will prove it for certain ranges of the noise strength, i.e., if the parameters p, q of the depolarizing and amplitude damping channels exceed some constant threshold. To this end, we will upper bound the trace distance of the outputs of two arbitrary input states, by resorting to the contraction coefficients the quantum Wasserstein distance of order 1, introduced in Ref. [DPMTL21a]. We remark that similar kinds of *reverse threshold theorems* have already appeared in the literature. For instance, let ρ, σ the outputs of an m-depth circuit obtained from two arbitrary input states. Assuming that the circuit is interspersed with local depolarizing noise with constant noise strength $p > 2/3$, ([HRF23], Proposition IV.8) showed that $\|\rho - \sigma\|_1 \leq 2^{\Omega(-m)}$, complementing previous bounds of [Raz03, KRUDW08]. Moreover, for the global depolarizing noise with p an arbitrarily small constant, the result of [Rag03] implies that $\|\rho - \sigma\|_1 \leq 2^{\Omega(-m)}$. For non-unital noise, if the noise strength is an arbitrarily small constant, [FMHS22] showed that the trace distance is exponentially small at exponential depth, with implications for the space overhead of quantum error correction.

Our bounds extends the results of [HRF23] from the local depolarizing noise to a broad class of local non-unital channels, assuming that the local noise is preceded by a single-qubit gate drawn from a 2-design. Moreover, we note that an analogous result holds for the amplitude damping noise, even in absence of randomness. We emphasize that both these bounds require the noise strength to be above a certain threshold. We start by recalling that the Wasserstein distance of order 1 and the trace distance are within a factor n. This property is particularly suitable for our goals, as we aim at proving that the outputs of a noisy circuit are exponentially close for any arbitrary pair of initial

states, and thus the factor n is of little importance in our setting. As in Ref. [HRF23], our argument is based on the contraction coefficient $\|\cdot\|_{W_1\to W_1}$. Importantly, as showed in Ref. [DPMTL21a], if Φ is a layer of of k-qubit gates, the contraction coefficient of Φ can be bounded by light-cone argument as follows

$$\|\Phi\|_{W_1\to W_1} \leq \begin{cases} 1 & \text{if } k=1, \\ \frac{3}{2}k & \text{if } k>1 \ \text{([DPMTL21a], Proposition 13).} \end{cases} \tag{5.57}$$

And thus a layer of two qubit gates has contraction coefficient at most 3. Moreover, $\|\mathcal{N}_p^{(\mathrm{dep})\otimes n}\|_{W_1\to W_1} = 1-p$, which readily implies ([HRF23], Proposition IV.8). If $\mathcal{N}$ is a single-qubit channel, the contraction coefficient of the tensor power channel $\mathcal{N}^{\otimes n}$ can be upper bounded by the diamond distance between $\mathcal{N}$ and a suitable 1-qubit channel $\mathcal{E}$, as detailed in Proposition 3.1. This yields the following result.

Proposition 5.3. *Let c be a positive constant and let $\mathcal{C}$ a noisy circuit consisting in m layers of 2-qubit gates interspersed with local noise, either of the form $\mathcal{N}_{p,q}^{(\mathrm{dep,amp}),\otimes n}$ or $\mathcal{N}_{q,p}^{(\mathrm{amp,dep}),\otimes n}$. Then, if*

$$\min\{1-p, 2(1-p)(1-q)\} \leq \frac{c}{3}, \tag{5.58}$$

we have

$$W_1(\mathcal{C}(\rho),\mathcal{C}(\sigma)) \leq \|\mathcal{C}(\rho)-\mathcal{C}(\sigma)\|_1 \leq 2c^m W_1(\rho,\sigma) \leq nc^m\|\rho-\sigma\|_1. \tag{5.59}$$

for all $\rho,\sigma\in\mathcal{S}_n$. This implies that W_1 distance decays exponentially fast in m if $c<1$.

Proof. If $q=0$, the result coincides with ([HRF23], Proposition IV.8). We note that the result in Ref. [HRF23] still holds if $q\neq 0$, as local channels do not increase the W_1 distance. Moreover, as showed in (Proposition 12, [DPMTL21a]), the contraction coefficient of the depolarizing noise is $(1-p)$.

It remains to upper bound the contraction coefficient of the amplitude damping channel. By (Proposition 11, [DPMTL21a]), it suffices to upper bound $\|\mathcal{N}_q^{(\mathrm{amp})}-\mathcal{N}_1^{(\mathrm{amp})}\|_\diamond$, where $\mathcal{N}_1^{(\mathrm{amp})}$ is the channel sending all states to $|0\rangle\langle 0|$. By [PP21], this is at most $2(1-q)$, and hence the contraction coefficient of the amplitude damping channel is at most $\min\{1,2(1-q)\}$. Finally, we recall that the contraction coefficient of the layer of unitaries is at most 3 from ([DPMTL21a], Proposition 13). Multiplying all the contraction coefficients and iterating over m layers yields the desired result:

$$\|\mathcal{C}(\rho)-\mathcal{C}(\sigma)\|_1 \leq 2W_1(\mathcal{C}(\rho),\mathcal{C}(\sigma)) \leq 2c^m W_1(\rho,\sigma) \leq nc^m\|\rho-\sigma\|_1. \tag{5.60}$$

■

For a general noise channel $\mathcal{N}$, we can provide a similar result by making further assumptions on the circuit randomness.

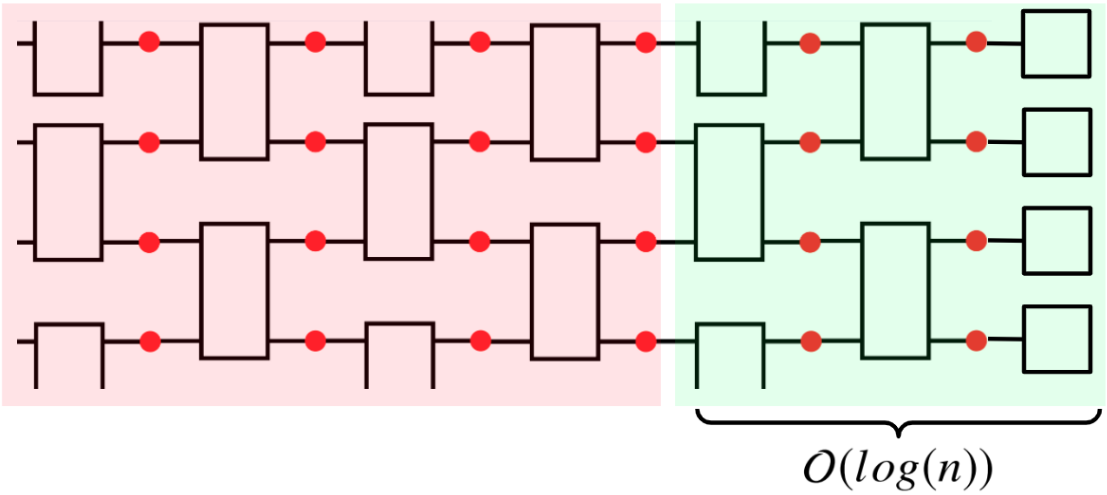


Figure 5.2: Here we provide a pictorial representation of the "effective depth" noisy circuit. We conjecture that the final state $\rho(x) = \Phi_x(\rho_0)$ bears little information about the initial portion of the circuit. In the context of cost function-based variational machine learning, this also implies that the parameters encoded in this initial portion are not trainable.

Proposition 5.4. *Let c be a positive constant and let $\mathscr{C}$ a noisy circuit consisting in m layers of 2-qubit gates interspersed with local noise, modeled by the channel $\mathcal{N}^{\otimes n}$. Then, if*

$$\sqrt{\frac{4}{3}\sum_{P,Q\in\{X,Y,Z\}} t_{P,Q}^2} \leq \frac{c}{3}, \tag{5.61}$$

where $\{t_{P,Q}\}_{P,Q\in\mathscr{P}_1}$ are the entries of the Pauli Transfer Matrix of the channel $\mathcal{N}$. Let ν be the distribution over the circuit gates, and assume that each 2-qubit gate is sampled from a local 2-design.

$$\mathbb{E}_\nu W_1(\mathscr{C}(\rho),\mathscr{C}(\sigma)) \leq \mathbb{E}_\nu \|\mathscr{C}(\rho)-\mathscr{C}(\sigma)\|_1 \leq 2c^m W_1(\rho,\sigma) \leq nc^m \|\rho-\sigma\|_1. \tag{5.62}$$

for all $\rho,\sigma \in \mathscr{S}_n$. This implies that expected W_1 distance decays exponentially fast in m if $c<1$.

Conjecture 5.1. *Let $\mathscr{C}$ be a noisy circuit consisting in m layers of 2-qubit gates interspersed with local noise, either of the form $\mathcal{N}_{p,q}^{(\mathrm{dep,amp}),\otimes n}$ or $\mathcal{N}_{q,p}^{(\mathrm{amp,dep}),\otimes n}$. Moreover, assume that each 2-qubit gate is sampled independently from a local 2-design. Then if $p,q=\Omega(1)$, for all $\rho,\sigma\in\mathscr{S}_n$, we have*

$$\mathbb{E}\|\mathscr{C}(\rho)-\mathscr{C}(\sigma)\|_{\mathrm{tr}} \in 2^{\Omega(-n)} \tag{5.63}$$

5.5.2 Kernel-based algorithms in the high noise regime

We will now draw an inference from Proposition 5.3 by applying it to the realm of projected quantum kernels. This analysis reveals that the data stored in the initial layers of the circuit diminishes significantly in influence, akin to being "forgotten", in the resulting Gram matrix, provided that the noise rates exceeds some fixed thresholds.

Theorem 5.4. *Let $x = (x_1, x_2, \dots, x_d), y = (y_1, y_2, \dots, y_d) \in \mathscr{X}$ two input vector such that $\forall i \neq j : x_i = y_i$, i.e. x and y coincides up to the j-th entry. Let c be a positive constant and let Φ_x, Φ_y a noisy circuit parametrized by x and y, consisting in m layers of 2-qubit gates interspersed with local noise, either of the form $\mathscr{N}_{p,q}^{(\mathrm{dep,amp}),\otimes n}$ or $\mathscr{N}_{q,p}^{(\mathrm{amp,dep}),\otimes n}$. Assume that the j-th entry of the input vector is encoded in the ℓ-th layer. Then, if*

$$\min\{1-p, 2(1-p)(1-q)\} \leq \frac{c}{3}, \tag{5.64}$$

we have

$$|1-\kappa^{PQ}| \leq \gamma n^3 c^{2(m-\ell)}. \tag{5.65}$$

In particular, if $m \geq t \cdot n + \ell$ for a sufficiently large constant t and $\gamma \in O(\mathrm{poly}(n))$, we have the following upper bound on the variance

$$|1-\kappa^{PQ}| \in 2^{-\Omega(n)}. \tag{5.66}$$

Proof. We let $\Phi_x = \Phi_x^{(A)} \circ \Phi_x^{(B)}$, where $\Phi_x^{(A)}$ represents the evolution of the last $m-\ell$ layers and $\Phi_x^{(B)}$ represents the evolution of the first ℓ layers, and analogously we let $\Phi_y = \Phi_y^{(A)} \circ \Phi_y^{(B)}$. We denote $\Phi^{(A)} := \Phi_x^{(A)} = \Phi_y^{(A)}$. We have,

$$\rho(x) = \Phi^{(A)} \circ \Phi_x^{(B)}(\rho_0), \tag{5.67}$$

$$\rho(y) = \Phi^{(A)} \circ \Phi_y^{(B)}(\rho_0), \tag{5.68}$$

As $\Phi_x^{(A)}$ has depth $m-\ell$, by Proposition 5.3 we obtain

$$\|\rho_k(x) - \rho_k(y)\|_{\mathrm{tr}} \leq \|\rho(x) - \rho(y)\|_{\mathrm{tr}} \leq n c^{m-\ell}. \tag{5.69}$$

Therefore,

$$|1-\kappa^{PQ}(x,y)| = \left|1 - \exp\left(-\gamma \sum_{k=1}^{n} \|\rho_k(x) - \rho_k(y)\|_2^2\right)\right| \tag{5.70}$$

$$\leq \gamma \sum_{k=1}^{n} \|\rho_k(x) - \rho_k(y)\|_2^2 \leq \gamma \sum_{k=1}^{n} \|\rho_k(x) - \rho_k(y)\|_2^2 \tag{5.71}$$

$$\leq \gamma n^3 c^{2(m-\ell)}. \tag{5.72}$$

■

5.5.3 Cost-based algorithms in the high noise regime

In a similar manner, we can tailor Proposition 5.3 to the specific scenario of cost-based algorithms. This specialized analysis yields a result akin to the concept of "barren plateaus", particularly in relation to the trainable parameters situated within the early layers of a noisy circuit. It demonstrates how these parameters have limited influence on the overall outcome.

Theorem 5.5. *Let c be a positive constant and let $\mathscr{C}$ a noisy circuit consisting in m layers of 2-qubit gates interspersed with local noise, either of the form $\mathcal{N}_{p,q}^{(\text{dep,amp}),\otimes n}$ or $\mathcal{N}_{q,p}^{(\text{amp,dep}),\otimes n}$. Assume that the gates composing $\mathscr{C}$ are parametrized by a vector $(\boldsymbol{\theta})$. Denote the output state by $\rho(\boldsymbol{\theta}) := \mathscr{C}(|0^n\rangle\langle 0^n|)$. Let H be a an observable with $\|H\| \leq 1$, and let $C(\boldsymbol{\theta}) := \text{Tr}(H\rho(\boldsymbol{\theta}))$ be the associated cost function. Let $\partial_\mu C(\boldsymbol{\theta})$ be the gradient of the cost function with respect to a parameter μ encoded in the j-th layer. Then, if*

$$\min\{1-p, 2(1-p)(1-q)\} \leq \frac{c}{3}, \tag{5.73}$$

we have

$$|\partial_\mu C(\boldsymbol{\theta})| \leq nc^{m-j}. \tag{5.74}$$

In particular, if $m \geq t \cdot n + j$ for a sufficiently large constant t, we have the following upper bound on the variance

$$\text{Var}[\partial_\mu C(\boldsymbol{\theta})] \leq O(2^{-n}). \tag{5.75}$$

Proof. Recall that Proposition 5.3 implies for an m-depth circuit $\mathscr{C}$ and for two arbitrary states ρ, σ,

$$\frac{1}{2}\|\mathscr{C}(\rho) - \mathscr{C}(\sigma)\|_1 \leq nc^m. \tag{5.76}$$

More generally, we can write the noisy circuit as $\mathscr{C} = \Phi_A \circ \Phi_B$, where Φ_A represents the evolution of the last $m-j$ layers and Φ_B represents the evolution of the first j layers. Then for a traceless self-adjoint linear operator X we can write,

$$\|\Phi_A(X)\|_1 \leq 2\|\Phi_A(X)\|_{W_1} \leq 2c^{m-j}\|X\|_{W_1} \leq nc^{m-j}\|X\|_1. \tag{5.77}$$

where Φ_A is the channel corresponding to the last $L-j$ layers of the noisy circuit and Φ_B is the channel corresponding to the first j layers. Recall that the gradient with respect to a parameter μ encoded in the j-th layer is proportional the expectation of a suitable observable as

$$\begin{aligned} |\partial_\mu C(\boldsymbol{\theta})| &= \left|\text{Tr}\left[\Phi_B(\rho_0)\left[H_\mu, \Phi_A^\dagger(H)\right]\right]\right| \\ &= \left|\text{Tr}\left[\Phi_A^\dagger(H)\left[\Phi_B(\rho_0), H_\mu\right]\right]\right| = \left|\text{Tr}\left[H\Phi_A\left(\left[\Phi_B(\rho_0), H_\mu\right]\right)\right]\right| \\ &\leq \left|\|H\|_\infty \left\|\Phi_A\left(i\left[\Phi_B(\rho_0), H_\mu\right]\right)\right\|_1\right| \leq nc^{m-j}\|H\|_\infty \left\|\left[\Phi_B(\rho_0), H_\mu\right]\right\|_1, \end{aligned} \tag{5.78}$$

where the first inequality is Holder's inequality , and the second one follows from the definition of contraction coefficient. We can also bound the 1-norm of the commutator as

$$\begin{aligned} \left\|\left[\Phi_B(\rho_0), H_\mu\right]\right\|_1 &\leq 2\max\{\|\Phi_B(\rho_0)H_\mu\|_1, \|H_\mu\Phi_B(\rho_0)\|_1\} = 2\max_{\|R\|_\infty=1}\left|\text{Tr}(\Phi_B(\rho_0)H_\mu R)\right| \\ &\leq 2\max_{\|R'\|_\infty=1}\left|\text{Tr}(\Phi_B(\rho_0)R')\right| = 2\|\Phi_B(\rho_0)\|_1 \leq 4. \end{aligned} \tag{5.79}$$

Thus, if $L \geq t \cdot n + j$ for a sufficiently large constant t, then $|\partial_\mu C(\boldsymbol{\theta})| \leq O(2^{-n})$, i.e., the gradient $\partial_\mu C(\boldsymbol{\theta})$ is exponentially concentrated around 0 for each value of $\boldsymbol{\theta}$. We can also transfer this bound to the variance by Popoviciu's inequality,

$$\text{Var}[\partial_\mu C] \leq O(2^{-n}). \tag{5.80}$$

■

CHAPTER 6

LEARNING UNITARIES WITH QUANTUM STATISTICAL QUERIES

In the previous section, we discussed how noise can degrade the performance of variational quantum algorithms. In particular, we modeled noise as a series of local channels interspersed in-between the quantum gates, therefore perturbing the ideal unitary evolution of the input state. In this section, we study a distinct but related model, meant to capture the capabilities of a learner with modest quantum resources. Particularly, in the *quantum statistical query* (QSQ) model, we consider a learner without quantum memory that can only access noisy estimates of the expected values of chosen observables on an unknown initial state. We will show that several algorithms for learning unitaries from oracle access can be fruitfully rephrased in this model. Our methods hinge on a novel technique for estimating the Fourier mass of a unitary on a subset of Pauli strings with a single quantum statistical query, generalizing a previous result for uniform quantum examples. Exploiting this insight, we show that the quantum Goldreich-Levin algorithm can be implemented with quantum statistical queries, whereas the prior version of the algorithm involves oracle access to the unitary and its inverse. Furthermore, we demonstrate that $\mathcal{O}(\log n)$-juntas and quantum Boolean functions with constant total influence are efficiently learnable in our model, and constant-depth circuits are learnable sample-efficiently with quantum statistical queries. On the other hand, all previous algorithms for these tasks require direct access to the Choi-Jamiolkowski state or oracle access to the unitary. Additionally, our upper bounds imply the efficient learning of those classes of unitaries with respect to locally scrambled ensembles. We also demonstrate that,

despite these positive results, quantum statistical queries lead to an exponentially larger sample complexity for certain tasks, compared to separable measurements to the Choi-Jamiolkowski state. In particular, we show an exponential lower bound for learning a class of phase-oracle unitaries and a double exponential lower bound for testing the unitarity of channels, adapting to our setting previous arguments for quantum states. Finally, we propose a new definition of average-case surrogate models, showing a potential application of our results to hybrid quantum machine learning.

6.1 Motivation and context

Learning the dynamic properties of quantum systems is a fundamental problem at the intersection of machine learning (ML) and quantum physics. In the most general case, this task can be achieved under the broad framework of quantum process tomography (QPT) [CN97]. However, QPT can be extremely resource-intensive, as learning the entire classical description of a unitary transformation requires exponentially many queries [GJ14] in the worst case. This complexity can be significantly reduced if the unitary is not completely arbitrary, but instead it belongs to a specific class. For instance, this approach has been fruitfully adopted for quantum Boolean functions [MO10], quantum juntas [CNY23, BY23] and quantum circuits with bounded covering numbers [FQR22]. On the other hand, the complexity of quantum process tomography could be drastically reduced if we restrict our attention only on local properties of the output state, as recently demonstrated in [HCP22]. Another scenario of interest is the one of property testing, where the learner is not asked to retrieve the classical description of the target process, but solely to *test* whether is satisfies some specific property [MdW13]. A further figure of merit in quantum process learning is the type of resources that the learner is allowed to use. For the special case of unitary transformations, the learner is usually given oracle access to the target unitary U and its inverse $U^\dagger$, or, alternatively, to the corresponding Choi-Jamiolkowski state. In this chapter we consider this latter approach and we ask the following question:

Which classes of unitaries are efficiently learnable with noisy separable binary measurements of the Choi-Jamiolkowski state?

This question is motivated by near-term implementations of quantum algorithms, which involve several sources of noise and severely limited entangling capacity [Pre18a]. To this end, we adopt the model of *quantum statistical queries* (QSQs), previously introduced in [AGY20] as an extension of the (classical) statistical query model [Kea98a]. In the QSQ model, we consider a learner without quantum memory that can only access noisy estimates of the expected values of chosen observables on an unknown initial state. As noted in [AHS23], this is essentially equivalent to performing noisy separable binary measurements. Thus our leading question can be rephrased as follows.

Question 2. *Can we employ quantum statistical queries to learn quantum dynamics?*

Interestingly, several concept classes such as parities, juntas function, and DNF formulae are efficiently learnable in the QSQ model, whereas the classical statistical query model necessitates an exponentially larger number of samples. Despite these positive results, resorting to quantum statistical queries can be considerably limiting for some tasks. In particular, the authors of [AHS23] have established an exponential gap between QSQ learning and learning with quantum examples in the presence of classification noise. Quantum statistical queries have also found practical applications in classical verification of quantum learning, as detailed in [CHI+23]. Furthermore, they have been employed in the analysis of quantum error mitigation models [QFK+22, AHS23] and quantum neural networks [DHL+21a]. Alternative variations of quantum statistical queries have also been explored in [HIN+23, GL22, NIS+23]. Moreover, the connection between quantum statistical queries and quantum differential privacy was investigated in [AGY20], and an equivalence between quantum statistical query learning and quantum local differential privacy [AK22a] .

Our contributions. In this chapter we demonstrate that several classes of unitaries are efficiently learnable with quantum statistical queries with respect to their Choi state. In particular, we show our result for a natural distance over unitaries induced by the Choi-Jamiolkowski isomorphism and previously adopted in [MdW13, BY23]. It is crucial to note that this distance choice enables the prediction of a target unitary's action on a randomly sampled input state from a locally scrambled ensemble [CHE+23].

To provide a more accessible overview of our upper bounds, we will offer an informal description. Unless explicitly stated otherwise, the tolerance of a quantum statistical query is, at least, polynomially small.

- Constant depth circuits are learnable with polynomially many quantum statistical queries (Theorem 6.2).
- Quantum $\mathcal{O}(\log n)$-juntas are efficiently learnable with polynomially many quantum statistical queries (Theorem 6.3).
- Quantum Boolean functions with constant total influence are efficiently learnable with polynomially many quantum statistical queries (Theorem 6.5). In order to prove this result, we show that the quantum Goldreich-Levin algorithm can be implemented with quantum statistical queries (Theorem 6.4).

While these positive results show that a wide class of unitaries can be efficiently learned in our model, we also argue that resorting to quantum statistical queries leads to an exponentially larger sample complexity for certain tasks. In particular, we give the following lower bounds.

- There is a class of phase oracle unitaries that requires exponentially many quantum statistical queries with polynomially small tolerance to be learnt below distance 0.005 with high probability (Theorem 6.6);

- Estimating the unitarity of a quantum channel with error smaller than 0.24 and polynomially small tolerance requires double-exponentially many quantum statistical queries (Theorem 6.7).

Moreover, prior results imply that both tasks can be efficiently performed with polynomially many copies of the associated Choi-Jamiolkowski state [MdW13, ABDY22]. In Section 6.3.3.1, we complement our theoretical findings with a numerical simulation the quantum Goldreich-Levin algorithm implemented with quantum statistical queries. Finally, in Section 6.5 we suggest a potential application of our results to hybrid quantum machine learning. Prior work [SEM23, JGM$^+$23] showed that certain quantum learning models can be replaced by classical surrogates during the prediction phase. We argue that the learning algorithms provided in the present chapter can also serve to this scope. To this end, we extend the definition of classical surrogates from the worst-case to the average-case.

Related work. Our results generalize prior work in two ways. On one hand, we show that several classes of unitaries are learnable in the QSQ model, while all previous results involved the access to stronger oracles. The adoption of a weaker oracle is particularly advantageous for near-term implementation, since the definition of QSQs accounts for the measurement noise. On the other hand, we demonstrate that prior QSQ algorithms for learning classical Boolean functions can be generalized to unitary learning. In particular, the authors of [CNY23] demonstrated that k-junta unitaries can be effectively learned using $\mathcal{O}(4^k)$ copies of the Choi state. Furthermore, the quantum Goldreich-Levin algorithm, as initially proposed in [MO10], relies on oracle access to both the target unitary and its inverse. This quantum algorithm builds upon the foundations of the classical Goldreich-Levin algorithm, first introduced in [GMW87].

Furthermore, an algorithm for learning classical k-junta functions with $\mathcal{O}(2^k)$ uniform quantum examples was provided in [AS07], and the authors of [AGY20] demonstrated that several classes of quantum Boolean functions are learnable with quantum statistical queries with respect to uniform quantum examples. In particular, they showed that classical k-junta functions are learnable with $\mathcal{O}(2^k + n)$ quantum statistical queries, and moreover that the (classical) Goldreich-Levin algorithm can be implemented in the QSQ model. In a subsequent work [AHS23], it was demonstrated that the output of constant-depth circuits is learnable with poly(n) quantum statistical queries and provided several hardness results for the QSQ model. Specifically, the authors showed an exponential lower bound for learning a class of classical Boolean functions, and a double exponential lower bound for testing the purity of a target state.

Moreover, a simultaneous work [WD23] devised a general QSQ oracle for learning quantum processes where a learner can select both the input state and the measurement, showing that the algorithm for learning arbitrary quantum processes from [HCP22] can be implemented in their model.

6.1.1 Fourier analysis on the unitary group

Let $U \in \mathscr{U}_n$ a unitary and consider the Pauli expansion $U = \sum_{P\in\mathscr{P}_n} \widehat{U}_P P$. We observe that the corresponding Choi state $|v(U)\rangle$ admits an analogous expansion with the same coefficients:

$$|v(U)\rangle = \left(I_n \otimes \sum_{P\in\mathscr{P}_n} \widehat{U}_P P\right)\left(\frac{1}{\sqrt{2^n}} \sum_{i\in\{0,1\}^n} |i,i\rangle\right) = \sum_{P\in\mathscr{P}_n} \widehat{U}_P |v(P)\rangle\,. \tag{6.1}$$

We now recall the notion of influence of qubits on linear operators, introduced in [MO10] in the context of Hermitian operators and further developed in [CNY23, RWZ22]. The related influence of variables is widely used in the analysis of Boolean functions [O'D21]. We define the quantum analogue of the bit-flip map as superoperator on $\mathscr{L}_n$:

$$d_j := I^{\otimes(j-1)} \otimes \left(I - \frac{1}{2}\mathrm{Tr}\right) \otimes I^{\otimes(n-j)}. \tag{6.2}$$

Then for $P = \bigotimes_{i=1}^n P_i \in \mathscr{P}_n$, we have

$$d_j P = \begin{cases} P & \text{if } P_j \neq I, \\ 0 & \text{if } P_j = I. \end{cases} \tag{6.3}$$

For a linear operator $A \in \mathscr{L}_n$, $A = \sum_{P\in\mathscr{P}_n} \widehat{A}_P P$, we have

$$d_j A = \sum_{P:P_j\neq I} \widehat{A}_P P. \tag{6.4}$$

For $p \geq 1$, we denote by $\mathrm{Inf}_j^p(A) := \|d_j A\|_p^p$ the L^p-influence of j on the operator A. For $S \in [n]$, we denote by $\mathrm{Inf}^p(A) := \sum_{j=1}^n \mathrm{Inf}_j^p(A)$ the associated total L^p-influence. We will often omit the index p when $p = 2$. Following [CNY23], we also define the influence of a subset of qubits $S \in [n]$ as

$$\mathrm{Inf}_S(A) = \sum_{\substack{P\in\mathscr{P}_n: \\ \mathrm{supp}(P)\cap S\neq\emptyset}} |\widehat{A}_P|^2. \tag{6.5}$$

We observe that $\mathrm{Inf}_j(A) = \mathrm{Inf}_{\{j\}}(A) = \sum_{P\in\mathscr{P}_n:P_j\neq I} |\widehat{A}_P|^2$, as expected. Intuitively, the influence of a unitary U on a subset of qubits is a quantitative measure of the action of U on such subset.

6.2 The model

We first give the definition of the QSQ oracle. For a state $\rho \in \mathscr{S}_n$, the QStat_ρ oracle receives as input an observable $O \in \mathscr{L}_n$, $\|O\| \leq 1$ and a tolerance parameter $\tau \geq 0$, and returns a τ-estimate of $\mathrm{Tr}[O\rho]$, i.e.

$$\mathsf{QStat}_\rho : (O,\tau) \mapsto \mathrm{Tr}[\rho O] \pm \tau. \tag{6.6}$$

A typical choice of the target state is the uniform quantum example $|\psi_f\rangle := \sum_{x\in\{0,1\}^n} \frac{1}{\sqrt{2^n}} |x, f(x)\rangle$, for a suitable Boolean function $f : \{0,1\}^n \to \{0,1\}$, which was first introduced in [BJ95] and widely

employed in previous works on quantum statistical query learning [AGY20, AHS23]. In this case, we will shorten the notation to $\mathsf{QStat}_f = \mathsf{QStat}_{|\psi_f\rangle\langle\psi_f|}$. To adapt their framework to our goal of learning unitaries, we need to devise an alternative input state. A natural choice is the Choi-Jamiolkowki state, which found many applications in prior work about unitary learning [CNY23], and more broadly process learning [Car22], motivating its adoption in the context of quantum statistical query. For brevity, we will write QStat_U instead of $\mathsf{QStat}_{|v(U)\rangle\langle v(U)|}$. We now detail the mutual relationship between the oracle QStat_U and the previous oracles defined in terms of quantum examples. To this end, we consider two unitaries implementing f, notably the bit-flip oracle U_f and the phase oracle V_f. We have,

$$\forall x \in \{0,1\}^n, y \in \{0,1\} : U_f |x,y\rangle = |x, y \oplus f(x)\rangle\,, \tag{6.7}$$

$$\forall x \in \{0,1\}^n : V_f |x\rangle = (-1)^{f(x)} |x\rangle \tag{6.8}$$

In particular we note that $|\psi_f\rangle = \frac{1}{\sqrt{2^n}} U_f \sum_{x\in\{0,1\}^n} |x,0\rangle$. We show that QStat_f can be simulated by QStat_{U_f} and conversely QStat_{V_f} can be simulated by QStat_f. The first result shows that our framework generalizes the previous one based on quantum examples, while the second one allows us to transfer lower bounds from classical Boolean functions to unitaries, as formalized in Theorem 6.6.

Lemma 6.1 (Relations between QSQ oracles)**.** *Let $f : \{0,1\}^n \to \{0,1\}$ a Boolean function and consider the bit-flip oracle U_f and the phase oracle V_f. Then for every observable $A \in \mathscr{L}_{n+1}$, there exists an observable $A' \in \mathscr{L}_{2n+2}$ such that*

$$\langle\psi_f|A|\psi_f\rangle = \langle v(U_f)|A'|v(U_f)\rangle\,. \tag{6.9}$$

and, similarly, for every observable $B \in \mathscr{L}_{2n}$, there exists an observable $B' \in \mathscr{L}_{n+1}$ such that

$$\langle v(V_f)|B|v(V_f)\rangle = \langle\psi_f|B'|\psi_f\rangle\,. \tag{6.10}$$

Proof. The first result follows by selecting $A' = I_n \otimes |0\rangle\langle 0| \otimes A$. As for the second result, we can write the following expansion $B = \sum_{P,Q\in\mathscr{P}_n} c_{P,Q} |v(P)\rangle\langle v(Q)|$. From ([MO10], Proposition 9), we know that $|v(V_f)\rangle = \sum_{x\in\{0,1\}^n} \widehat{f(x)}\,|v(Z^x)\rangle$, where we denoted $Z^x := \bigotimes_{i\in[n]} Z^{x_i}$, with $Z^0 = I$ and $Z^1 = Z$. Hence

$$\langle v(U_f)|B|v(U_f)\rangle = \sum_{x\in\{0,1\}^n} c^2_{Z^x,Z^x} \widehat{f(x)}^2. \tag{6.11}$$

Now, consider the observable $T = \sum_{x\in\{0,1\}^n} c_{Z^x,Z^x} |x\rangle\langle x| \in \mathscr{L}_n$ and define

$$B' = H^{\otimes(n+1)}(I_n \otimes |1\rangle\langle 1|) \cdot T \cdot (I_n \otimes |1\rangle\langle 1|) H^{\otimes(n+1)}, \tag{6.12}$$

which is equivalent to perform the Fourier transform on $|\psi_f\rangle$, post-selecting on the last qubit being 1 and finally applying T on n qubits. The Fourier transform and the projection on $|1\rangle\langle 1|$ give rise to

$$|\widehat{\psi}_f\rangle = \sum_{x\in\{0,1\}^n} \widehat{f(x)}\,|x\rangle\,. \tag{6.13}$$

Then the desired result follows by noting that

$$\langle \psi_f | B' | \psi_f \rangle = \langle \widehat{\psi}_f | T | \widehat{\psi}_f \rangle = \sum_{x \in \{0,1\}^n} c^2_{Z^x, Z^x} \widehat{f(x)}^2. \tag{6.14}$$

■

We argue that this choice of the oracle is particularly suitable for learning the unitary evolution of states sampled from locally scrambled ensembles. This comes as a direct consequence of Lemmas 3.1 and 3.5, that together imply the following proposition.

Lemma 6.2. *For quantum unitaries $U, V \in \mathscr{U}_n$ and $\nu \in \mathscr{S}_{LS}$ a locally scrambled ensemble of states, it holds that*

$$\frac{1}{2} D(U,V)^2 \leq \mathscr{R}_\nu(U,V) \leq D(U,V)^2, \tag{6.15}$$

where $D(U,V)^2 = 1 - |\langle \nu(U) | \nu(V) \rangle|^2$.

We also introduce the following notion of learnability of classes of unitaries with quantum statistical queries.

Definition 6.1 (Unitary learning with QSQs)**.** Let $\varepsilon \in [0,1]$, $\mathscr{C} \subseteq \mathscr{U}_n$ a class of unitaries and ν an ensemble of n-qubit states. We say that $\mathscr{C}$ is efficiently ε-learnable with quantum statistical queries with respect to ν if, for all $U \in \mathscr{C}$, there exists an algorithm $\mathscr{A}$ that runs in time $\mathrm{poly}(n)$, performs $\mathrm{poly}(n)$ queries to the oracle QStat_U with tolerance at least $1/\mathrm{poly}(n)$ and outputs a unitary $V \in \mathscr{U}_n$ such that

$$\mathscr{R}_\nu(U,V) \leq \varepsilon. \tag{6.16}$$

We emphasize that all the algorithms proposed in this chapter are *proper* learners, in the sense that they output a unitary $V \in \mathscr{C}$. Moreover, they are classical randomized algorithms, as they use no other quantum resource apart from the query access to QStat_U. The QSQ model is considerably more restrictive than the *oracle access* model, where a learner has the freedom to implement the unitary U and its inverse $U^\dagger$ on an arbitrary input state. Then, every algorithm implementable with QSQs can be also implemented with oracle access, but the converse it is not true in general. In particular, we demonstrate in Theorem 6.6 that there is a class of unitaries that is efficiently learnable with direct access to the Choi state, but requires exponentially many quantum statistical queries.

6.3 Learning classes of unitaries with quantum statistical queries

Our results are based on the following technical lemma, which extends ([AGY20], Lemma 4.1) to unitary operators. In particular, this lemma allows us to estimate the influence of subset of qubits defined in Eq. 6.5.

Lemma 6.3 (Learning the influence of a subset with a single QSQ)**.** *Let $A \in \mathcal{U}_n$ be a unitary operator and QStat_A be the quantum statistical query oracle associated to the Choi state $|v(A)\rangle$. There is a procedure that on input a subset of Pauli strings $T \subseteq \mathscr{P}_n$, outputs τ-estimate of $\sum_{P\in T} |\widehat{A}_P|^2$ using one query to QStat_A with tolerance τ.*

Proof. Let $M = \sum_{P\in T} |v(P)\rangle\langle v(P)|$. We note that

$$\langle v(A)|M|v(A)\rangle = \left(\sum_{P\in\mathscr{P}_n} \widehat{A}_P^* \langle v(P)|\right)\left(\sum_{Q\in T} |v(Q)\rangle\langle v(Q)| \sum_{P\in\mathscr{P}_n} \widehat{A}_P |v(P)\rangle\right) \tag{6.17}$$

$$= \left(\sum_{P\in\mathscr{P}_n} \widehat{A}_P^* \langle v(P)|\right)\left(\sum_{Q\in T} \widehat{A}_Q |v(Q)\rangle\right) = \sum_{P\in T} |\widehat{A}_P|^2. \tag{6.18}$$

Thus a single query to Qstat_A with input (M,τ) yields the desired outcome. ■

Remark 6.1 (Computational efficiency). We observe that the circuit implementing the measurement $M = \sum_{P\in T} |v(P)\rangle\langle v(P)|$ can have exponential depth in the worst case. However, in some cases, even if the set T has exponential size, we can implement M with a poly(n) circuit. For instance, the influence of the j-th qubit $\mathrm{Inf}_j(A)$ can be expressed as

$$\mathrm{Inf}_j(A) = \sum_{\substack{P\in\mathscr{P}_n:\\ P_j\neq I}} |\widehat{A}_P|^2 = 1 - \sum_{\substack{P\in\mathscr{P}_n:\\ P_j = I}} |\widehat{A}_P|^2. \tag{6.19}$$

Thus it suffices to estimate the expected value of $|v(I)\rangle_j\langle v(I)|_j \otimes I_{n-1}$. More generally, we can consider the indicator string $\mathbb{S} = (x_1, x_2, \ldots, x_k, *, * \ldots, *)$ to denote the set of n-bit strings whose first k elements are $x_1, x_2, \ldots, x_k$, i.e. $\mathbb{S} = \{(t_1, t_2, \ldots, t_n) \in \{0,1,2,3\}^n | \ \forall i \in [k] : x_i = t_i\}$. Then we have,

$$\sum_{P\in\mathbb{S}} |v(P)\rangle\langle v(P)| = |v(\sigma_{x_1}\otimes\sigma_{x_2}\otimes\cdots\otimes\sigma_{x_k})\rangle\langle v(\sigma_{x_1}\otimes\sigma_{x_2}\otimes\cdots\otimes\sigma_{x_k})| \otimes I_{n-k}, \tag{6.20}$$

which again can be implemented by a poly(n) circuit.

We will also need a further technical tool, which is an implementation of state tomography with quantum statistical queries, also previously exploited in [AHS23] for learning the output of shallow circuits. Here we propose a refined argument for the special case of pure states. Since the complexity is exponential in the number of qubits, this primitive can be used to efficiently estimate the reduced states of subsets of logarithmic size.

Lemma 6.4 (State tomography)**.** *Let $\rho \in \mathscr{S}_n$. There exists an algorithm that performs 4^n queries to the oracle QStat_ρ with tolerance at least $\varepsilon \cdot 4^{-n}$ and returns a state $\widehat{\rho}$ such that*

$$\|\rho - \widehat{\rho}\|_2 \leq \varepsilon. \tag{6.21}$$

Moreover, if $\rho = |\psi\rangle\langle\psi|$ is a pure state, there exists an algorithm that performs 4^n queries to the oracle QStat_ρ with tolerance at least $\varepsilon \cdot 2^{-n/2}$ and returns a pure state $|\widehat{\psi}\rangle$ such that

$$\|\rho - |\widehat{\psi}\rangle\langle\widehat{\psi}|\|_{\mathrm{tr}} \leq \varepsilon. \tag{6.22}$$

Proof. We perform a state tomography by querying all $4^n - 1$ non-identity Pauli strings with tolerance $\tau = \varepsilon \cdot 4^{-n}$. For all $P \in \mathscr{P}_n$, denote the obtained outcome by

$$o_P = \mathrm{Tr}[P\rho] \pm \tau$$

and set $x_P = \min\{o_P, 1\}$. Denote the estimated state by

$$\widehat{\rho} := \frac{1}{2^n}\left(I + \sum_{P \in \mathscr{P}_n \setminus I} x_P P\right). \tag{6.23}$$

This allows to upper bound the distance between the partial state ρ and its estimate $\widehat{\rho}$.

$$\|\rho - \widehat{\rho}\|_2^2 = \mathrm{Tr}\left[(\rho - \widehat{\rho})^2\right] = \frac{1}{4^n}\mathrm{Tr}\left[\left(\sum_{P \in \mathscr{P}_n \setminus I} (\mathrm{Tr}[P\rho] - x_P)P\right)^2\right] \tag{6.24}$$

$$= \frac{1}{2^n}\sum_{P \in \mathscr{P}_n \setminus I} (\mathrm{Tr}[P\rho] - x_P)^2 \leq 2^n \tau^2, \tag{6.25}$$

where we used the inequality $(x+y)^2 \leq 2(x^2+y^2)$. Then picking $\tau = \varepsilon/\sqrt{2^n}$ gives the desired result. We now delve into the case where the input state is pure. Thanks to ([CCC19], Theorem 1), and since $\rho = |\psi\rangle\langle\psi|$ has rank 1, we obtain the following bound for the 1-distance:

$$\|\rho - \widehat{\rho}\|_1 \leq \sqrt{\frac{2^n}{2^n+1}}\,\|\rho - \widehat{\rho}\|_2 \leq \varepsilon. \tag{6.26}$$

We now consider the dominant eigenstate of $\widehat{\rho}$, denoted by $|\widehat{\psi}\rangle$, which can be computed in poly(2^n) time. By ([MGN20], Proposition 2) we know that $|\widehat{\psi}\rangle\langle\widehat{\psi}|$ is the unique closest pure state to $\widehat{\rho}$. Since ρ is also a pure state, this immediately implies

$$\left\||\widehat{\psi}\rangle\langle\widehat{\psi}| - \rho\right\|_{\mathrm{tr}} \leq \left\||\widehat{\psi}\rangle\langle\widehat{\psi}| - \widehat{\rho}\right\|_{\mathrm{tr}} + \|\rho - \widehat{\rho}\|_{\mathrm{tr}} \tag{6.27}$$

$$\leq 2\|\rho - \widehat{\rho}\|_{\mathrm{tr}} \leq \varepsilon, \tag{6.28}$$

■

6.3.1 Appetizer: learning constant-depth circuits

As a first application of the tools introduced before, we show that very shallow circuits are learnable sample-efficiently with QSQs according to a locally scrambled distribution. We will rely on the following recent result of [YW23], which essentially shows that "learning marginal suffices", i.e. learning the k-reduced density matrices of a state produced by a shallow circuit allows to perform a state tomography.

Theorem 6.1 (Adapted from [YW23], Theorem 4.3)**.** *Let $\psi = |\psi\rangle\langle\psi|$ a state produced by a circuit of depth at most D. For any state ρ, one of the following conditions must be satisfied: either $\|\rho - \psi\|_{\mathrm{tr}} < \varepsilon$; or $\|\rho_s - \psi_s\|_{\mathrm{tr}} > \varepsilon^2/n$ for some $s \subseteq \{0,1,\ldots,n-1\}$ with $|s| = 2^D$.*

An application of this result was also given in [AHS23], where the authors showed that the class of n-qubit trivial states is learnable with $\mathrm{poly}(n)$ quantum statistical queries. We now extend their result from states to unitaries.

Theorem 6.2 (Learning constant-depth circuits via QSQs)**.** *Let $\mathscr{C}$ the class of $\mathscr{O}(1)$-depth circuits. Then for all $U \in \mathscr{C}$, there exists an algorithm that makes $\mathrm{poly}(n)$ queries to QStat_U with tolerance at least $\frac{\varepsilon^2}{4n} \cdot 2^{-D/2}$ and returns a unitary $W \in \mathscr{U}_n$ such that*

$$D(U,W) \leq \varepsilon. \tag{6.29}$$

Proof. Let D be the depth of the circuit. First, we consider the Choi state $|v(U)\rangle = I \otimes U |\Omega\rangle$ and recall that $|\Omega\rangle$ can be produced with a circuit of depth 2 over $2n$ qubits. Then we have $|v(U)\rangle = V|0^{2n}\rangle$ for a suitable unitary $V \in \mathscr{U}_{2n}$ implemented by a circuit of depth $D+2$. Let $k = 2^{D+2}$. Then it suffices to learn all the k-local reduced density matrices of the states $|v(U)\rangle$. There are $\binom{2n}{k} = \mathscr{O}\left(n^{2^D}\right)$ of them and each of them is learnable in trace distance with accuracy $\frac{\varepsilon^2}{2n}$ by performing 4^{D+2} quantum statistical queries with tolerance $\frac{\varepsilon^2}{4n} \cdot 2^{-D/2}$ by means of Lemma 6.4. We can thus determine thanks to Theorem 6.1 a state $|v(W)\rangle$ such that $\| |v(W)\rangle\langle v(W)| - |v(U)\rangle\langle v(U)| \|_{\mathrm{tr}} \leq \varepsilon$. This immediately implies Eq. 6.29 by Lemma 6.2. ■

6.3.2 Learning quantum juntas

A unitary $U \in \mathscr{U}_n$ is a quantum k-junta if there exists $S \subseteq [n]$ with $|S| = k$ such that

$$U = V_S \otimes I_{\overline{S}}$$

for some $V_S \in \mathscr{U}_k$. For a Pauli string $P = \bigotimes_{i\in[n]} P_i \in \mathscr{P}_n$, we denote the reduced string as $P_S = \bigotimes_{i\in S} P_i \in \mathscr{P}_k$. We now consider the Pauli expansions $U = \sum_{P\in\mathscr{P}_n} \widehat{U}_P P$ and $V_S = \sum_{P_S\in\mathscr{P}_k} \widehat{V}_{P_S} P_S$. Their coefficients satisfy the following relation.

$$\widehat{U}_P = \frac{1}{2^n}\mathrm{Tr}[UP] = \frac{1}{2^n}\mathrm{Tr}[V_S P_S]\,\mathrm{Tr}[P_{\overline{S}} I_{\overline{S}}] = \begin{cases} \widehat{V}_{P_S} & \text{if } \mathrm{supp}(P) \in S, \\ 0 & \text{else.} \end{cases}$$

As for the Choi state, we have

$$|v(U)\rangle = \sum_{P\in\mathscr{P}_n} \widehat{U}_P |v(P)\rangle = \sum_{\mathrm{supp}(P)\in S} \widehat{V}_{P_S} |v(P_S \otimes I_{\overline{S}})\rangle = |v(V_S)\rangle\,|v(I_{\overline{S}})\rangle\,.$$

We will now show that quantum k-juntas are efficiently learnable in our model. Our proof combines the techniques used in [CNY23] for learning quantum k-juntas from oracle access and the ones used in [AGY20] for learning (classical) k-juntas with quantum statistical queries. Note that the algorithm given in ([CNY23], Theorem 28) has query complexity independent of n. Crucially, their algorithm involves a Pauli sampling as a subroutine to estimate the support of the Pauli strings with non-zero Fourier coefficients. We replaced this procedure by estimating the influences of each qubit by means of Lemma 6.3, introducing an additional factor n in the query complexity.

Algorithm 1 Learning quantum k-juntas with statistical queries

for $i = 1$ to n **do**

Estimate $\mathrm{Inf}_i^2(U)$ with a quantum statistical query with accuracy $\varepsilon^2/(20k)$ and store the result in the variable α_i.

end for

Define the subset $T = \left\{i \in [n] : \alpha_i \geq \varepsilon^2/(16k)\right\}$ and consider the set T_2, which includes the qubits in T and the associated qubits in the dual space.

for $P \in \mathscr{P}_{|T_2|}$ **do**

Produce an estimate o_p of

$$\mathrm{Tr}[P \cdot |v(I^{\otimes(n-\ell)})\rangle\langle v(I^{\otimes(n-\ell)})| \cdot (|v(U)\rangle\langle v(U)|) \cdot |v(I^{\otimes(n-\ell)})\rangle\langle v(I^{\otimes(n-\ell)})|]$$

with a quantum statistical query with tolerance $2^{-\ell}\varepsilon/3$.

Set $x_P = \min\{o_P, 1\}$.

end for

Reconstruct the density matrix $\widehat{\rho}_T = \frac{1}{2^{2\ell}}\left(I^{\otimes 2\ell} + \sum_{P \in \mathscr{P}_{2\ell} \setminus I^{\otimes 2\ell}} x_P P\right)$ and compute its dominant eigenstate $|\widehat{\psi}_T\rangle$.

Compute W such that $|v(W)\rangle := |\widehat{\psi}_T\rangle$

return $W \otimes I^{\otimes(n-\ell)}$.

Theorem 6.3 (Learning quantum k-juntas via QSQs)**.** *Let U be a quantum k-junta. There is a* $\mathrm{poly}(n, 2^k, \varepsilon)$*-time algorithm that accesses the state $|v(U)\rangle$ via QStat_U queries with tolerance* $\mathrm{poly}(2^{-k}, \varepsilon)$ *and outputs a unitary $\widetilde{U}$ such that*

$$\mathrm{D}(U, \widetilde{U}) \leq \varepsilon. \tag{6.30}$$

Proof. Throughout this proof, we will use the following notation to deal with the reduced Choi state with respect to a given subset of the qubits. Recall that the Choi state is a state over a set of $2n$ qubits, which we label as $\{i_1, i_2, \ldots, i_n, i'_1, i'_2, \ldots, i'_n\}$. For $S = \{i_{j+1}, i_{j+2}, \ldots\} \subseteq \{i_1, i_2, \ldots, i_n\}$ we will denote $S_2 := \{i_{j+1}, i_{j+2}, \ldots\} \cup \{i'_{j+1}, i'_{j+2}, \ldots\}$. Clearly, $|S_2| = 2|S|$.

Our algorithm consists in two separate steps: first we perform n QStat_U queries with tolerance $\Theta(\varepsilon^2/k)$ to learn a subset $T \subseteq [n]$ containing all the variables i for which $\mathrm{Inf}_i^2(U) \geq \varepsilon^2/(16k)$. Next we will define a reduced state on the subset T_2 and we will learn it by performing a state tomography with $4^{2|T|} - 1$ QStat_U queries with tolerance $\Omega(\varepsilon 4^{-2k})$.

Let U be a quantum k-junta over the subset $Q \subseteq [n]$. Then, it is not hard to see that $\mathrm{Inf}_i^2(U) = 0$ if $i \notin Q$. For each $j \in [n]$, we use Lemma 6.3 to estimate $\mathrm{Inf}_j^2(U) \pm \varepsilon^2/(20k)$ via a single Qstat_U query. Suppose the outcomes of these queries are $\alpha_1, \ldots \alpha_n$, and let

$$T = \left\{i \in [n] : \alpha_i \geq \varepsilon^2/(16k)\right\}.$$

We observe that $T \subseteq Q$, as $\mathrm{Inf}_i^2(U) = 0$ implies that $\alpha_i \leq \varepsilon^2/(20k)$. On the other hand, for every $i \in Q \setminus T$, we have that $\mathrm{Inf}_i^2(U) < \varepsilon^2/(8k)$. Assume by contradiction that $i \notin T$ and $\mathrm{Inf}_i^2(U) \geq \varepsilon^2/(4k)$. Then we have:

$$\alpha_i \geq \mathrm{Inf}_i^2(U) - \frac{\varepsilon}{20k} > \frac{\varepsilon^2}{16k},$$

contradicting the fact that $i \notin T$. As a consequence,

$$\sum_{i\in\overline{T}} \mathrm{Inf}_i^2(U) = \sum_{i\in Q\setminus T} \mathrm{Inf}_i^2(U) \leq k\cdot\frac{\varepsilon^2}{8k} = \frac{\varepsilon^2}{8}, \tag{6.31}$$

where the inequality follows from $|Q| \leq k$.

We now describe the second phase of the learning algorithm. Let $|T| = \ell$ and consider the identity operator $I^{\otimes(n-\ell)}$ acting on the subset $\overline{T}$. Let ρ be the state obtained by measuring $|v(U)\rangle$ according to the projectors $\left(|v(I^{\otimes(n-\ell)})\rangle\langle v(I^{\otimes(n-\ell)})|, I^{\otimes(n-\ell)} - |v(I^{\otimes(n-\ell)})\rangle\langle v(I^{\otimes(n-\ell)})|\right)$, and then conditioning on the first outcome,

$$|\psi\rangle := \frac{\left(I^{\otimes(\ell)}\otimes|v(I^{\otimes(n-\ell)})\rangle\langle v(I^{\otimes(n-\ell)})|\right)|v(U)\rangle}{\left|\left(\mathrm{Tr}_{T_2}\langle v(U)|\right)|v(I^{\otimes(n-\ell)})\rangle\right|} := |v(V\otimes I^{\otimes(n-\ell)})\rangle,$$

where in the last line we introduced the ℓ-qubit unitary V such that $|\psi\rangle$ is the state isomorphic to $V\otimes I^{\otimes(n-\ell)}$. We make the following claim on the distance between U and $V\otimes I^{\otimes(n-\ell)}$, which we will prove in the following.

Claim 6.1. $D(U, V\otimes I^{\otimes(n-\ell)}) \leq \varepsilon/2$.

Denote $\rho := |\psi\rangle\langle\psi|$. We will learn $\rho_{T_2} = \mathrm{Tr}_{\overline{T_2}}[\rho]$ by performing a state tomography via QStat queries on a reduced state of 2ℓ qubits. To this end, we query all $4^{2\ell}-1$ non-identity Pauli strings with support on T with tolerance $\tau = \varepsilon 2^{-2\ell-1}$. For all $P\in\mathscr{P}_{2\ell} = \{I,X,Y,Z\}^{\otimes 2\ell}$, denote the obtained outcome by

$$o_P = \mathrm{Tr}[P\cdot|v(I^{\otimes(n-\ell)})\rangle\langle v(I^{\otimes(n-\ell)})|\cdot(|v(U)\rangle\langle v(U)|)\cdot|v(I^{\otimes(n-\ell)})\rangle\langle v(I^{\otimes(n-\ell)})|] \pm \tau$$

and set $x_P = \min\{o_P, 1\}$. Denote the estimated 2ℓ-qubit state by

$$\widehat{\rho}_T = \frac{1}{2^{2\ell}}\left(I^{\otimes 2\ell} + \sum_{P\in\mathscr{P}_{2\ell}\setminus I^{\otimes 2\ell}} x_P P\right).$$

Let $|\widehat{\psi}_T\rangle$ be the dominant eigenstate of $\widehat{\rho}_T$ and let W be the unitary encoded by the state $|\widehat{\psi}_T\rangle$, i.e. let $|v(W)\rangle := |\widehat{\psi}_T\rangle$. We make a further claim and we delay its proof to the end.

Claim 6.2. $D(V, W) \leq \varepsilon/2$.

Then the theorem follows by combining Claims 6.1 and 6.2 with the triangle inequality and letting $\widetilde{U} = W\otimes I^{\otimes(n-\ell)}$. ■

We present the proofs of Claims 6.1 and 6.2 below.

Proof of Claim 6.1 Recall that $U = U_Q\otimes I_{\overline{Q}}$ is a k-junta which acts non trivially only on the set Q and that $T\subseteq Q$ is the set of qubits with non-negligible influence learnt by the algorithm. It is sufficient to

show that $\mathrm{dist}(U_Q, V) \leq \varepsilon/2$. First, we observe that $|v(U)\rangle = |v(U_Q)\rangle \otimes |v(I^{\otimes(n-k)})\rangle$. We will need the following decomposition of $|v(U_Q)\rangle$:

$$|v(U_Q)\rangle = \sum_{P_Q \in \mathscr{P}_k} \widehat{U}_{P_Q} |v(P_Q)\rangle \qquad = \sum_{\substack{P_Q \in \mathscr{P}_k \\ \mathrm{supp}(P_Q)\cap \overline{T} = \emptyset}} \widehat{U}_{P_Q} |v(P_Q)\rangle + \sum_{\substack{P_Q \in \mathscr{P}_k: \\ \mathrm{supp}(P_Q)\cap \overline{T} \neq \emptyset}} \widehat{U}_{P_Q} |v(P_Q)\rangle\,, \tag{6.32}$$

where $\widehat{U}_{P_Q} = \widehat{U}_{P_Q \otimes I_{n-k}}$. Similarly, we can expand $|v(V)\rangle \otimes |v(I^{\otimes(k-\ell)})\rangle$ as follows

$$|v(V)\rangle \otimes |v(I^{\otimes(k-\ell)})\rangle = \sum_{\substack{P_Q \in \mathscr{P}_k \\ \mathrm{supp}(P_Q)\cap \overline{T} = \emptyset}} \widehat{U}_{P_Q} |v(P_Q)\rangle + \sum_{\substack{P_Q \in \mathscr{P}_k: \\ \mathrm{supp}(P_Q)\cap \overline{T} \neq \emptyset}} \widehat{U}_{P_Q} |v(I^{\otimes k})\rangle \tag{6.33}$$

Recall that the total influence of the qubits in $\overline{T}$ is at most $\varepsilon^2/8$. This immediately implies a lower bound on the inner product between $|v(V)\rangle \otimes |v(I^{\otimes(k-\ell)})\rangle$ and $|v(U_Q)\rangle$.

$$\begin{aligned}\left|\left(\langle v(V)| \otimes \langle v(I^{\otimes(k-\ell)})|\right)|v(U_Q)\rangle\right| &= \sum_{\substack{P_Q \in \mathscr{P}_k: \\ \mathrm{supp}(P)\cap \overline{T} \neq \emptyset}} |\widehat{U}_{P_Q}|^2 \\ &= 1 - \sum_{\substack{P_Q \in \mathscr{P}_k: \\ \mathrm{supp}(P)\cap \overline{T} \neq \emptyset}} |\widehat{U}_{P_Q}|^2 \geq 1 - \frac{\varepsilon^2}{8},\end{aligned}$$

where the inequality is a direct application of Eq. 6.31. We can now prove the desired result

$$D^2(U, V \otimes I^{\otimes(n-\ell)}) = D^2(U_Q, V \otimes I^{\otimes(k-\ell)}) = 1 - |\langle v(V)\rangle |v(U_Q)|^2 \leq \frac{\varepsilon^2}{4},$$

where we used the stability of $D(\cdot,\cdot)$ under tensor product. ■

Proof of Claim 6.2 We just need to ensure the following:

$$\|\widehat{\rho}_{T_2} - \rho_{T_2}\|_2 \leq \frac{\varepsilon}{2}. \tag{6.34}$$

We first make a preliminary observation. Let $c_P := \mathrm{Tr}[P\rho_T]$. Then,

$$(x_p - c_p)^2 \leq \left(c_p \frac{\varepsilon^2}{8} + \tau\right)^2 \tag{6.35}$$

This allow to upper bound the distance between the partial state ρ_{T_2} and its estimate $\widehat{\rho}_{T_2}$.

$$\|\rho_{T_2} - \widehat{\rho}_{T_2}\|_2^2 = \mathrm{Tr}[(\rho_T - \widehat{\rho}_T)^2] = \frac{1}{16^\ell}\,\mathrm{Tr}\left[\left(\sum_{P \in \mathscr{P}_{2\ell} \setminus I^{\otimes 2\ell}} (c_P - x_P) P\right)^2\right] \tag{6.36}$$

$$= \frac{1}{4^\ell} \sum_{P \in \mathscr{P}_{2\ell} \setminus I^{\otimes 2\ell}} (c_P - x_P)^2 \leq \frac{2}{4^\ell}\left(\sum_{P \in \mathscr{P}_{2\ell} \setminus I^{\otimes 2\ell}} \frac{\varepsilon^4}{64} c_p^2 + \tau^2\right) \leq \frac{\varepsilon^4}{32} + 4^\ell \tau^2, \tag{6.37}$$

where we used the inequality $(x+y)^2 \leq 2(x^2+y^2)$ and the fact that the purity $\mathrm{Tr}\left[\rho_{T_2}^2\right] = 4^{-\ell} \sum_{P \in \mathscr{P}_{2\ell}} c_P^2$ is bounded by 1. Then picking $\tau = 2^{-\ell}\varepsilon/3$ ensures the desired upper bound. By proceeding as in the proof of Lemma 6.4, we have

$$D(V, W) \leq \frac{\varepsilon}{2}, \tag{6.38}$$

as desired. ■

6.3.3 Learning quantum Boolean functions

A quantum Boolean function A is defined as a Hermitian unitary operator [MO10], i.e. an operator satisfying

$$AA^\dagger = A^\dagger A = A^2 = I. \tag{6.39}$$

Notably, Pauli strings $P \in \mathscr{P}_n$ are Quantum Boolean and the unitary evolution (in the Heisenberg picture) of a Quantum Boolean function A is also Quantum Boolean. This can be easily checked by replacing A with $U^\dagger AU$ into the above equation. A key property of quantum Boolean functions is that their Fourier coefficients are all real, i.e.

$$\forall P \in \mathscr{P}_n : \widehat{A}_P \in \mathbb{R}. \tag{6.40}$$

We will now demonstrate that the quantum Goldreich-Levin (GL) algorithm ([MO10], Theorem 26) can be implemented via quantum statistical queries. Whereas the original algorithm requires oracles queries to the target unitary U and its adjoint, we show that the weaker access to QStat_U suffices. A similar result was also established for uniform quantum examples ([AGY20], Theorem 4.4), which are quantum encodings of *classical* Boolean functions. While we will employ Theorem 6.4 for learning quantum Boolean functions, we remark that it does not require the target operator to be Hermitian and it could find broader applications for learning other classes of unitaries.

Theorem 6.4 (Quantum Goldreich-Levin using QSQs)**.** *Let $A \in \mathscr{U}_n$ be a unitary operator and QStat_A be the quantum statistical query oracle associated to the Choi state $|v(A)\rangle$. There is a $\mathrm{poly}(n, 1/\gamma)$-time algorithm that accesses A via queries to QStat_A with tolerance at least $\gamma^2/4$ and outputs a list $L = \{P^{(1)}, P^{(2)}, \ldots, P^{(m)}\} \subseteq \mathscr{P}_n$ such that:*

1. *if $|\widehat{A}_P| \geq \gamma$, then $P \in L$;*
2. *and for all $P \in L$, $|\widehat{A}_P| \geq \gamma/2$.*

Proof. Our algorithm closely follows the one proposed in [MO10]. The only difference is that, for each subset $T \subseteq \{0,1,2,3\}^n$, the oracle queries to A and $A^\dagger$ are replaced by a Qstat_A query that outputs a $(\gamma^2/4)$-estimate of $\sum_{P\in T} |\widehat{A}_P|^2$, as in Lemma 6.3. The remaining part of the quantum Goldreich-Levin algorithm does not involve oracle access to A or $A^\dagger$, thus the rest of the proof coincides with the one of Theorem 26 in [MO10]. ■

Algorithm 2 Quantum Goldreich-Levin algorithm with statistical queries

$L \leftarrow (*, *, \ldots, *)$
for $k = 1$ to n **do**
  for each $\mathscr{S} \in L, \mathscr{S} = (P_1, P_2, \ldots, P_{k-1}, *, *, \ldots, *)$ **do**
    for P_k in $\{I, X, Y, Z\}$ **do**
      Let $\mathscr{S}_{P_k} = (P_1, P_2, \ldots, P_{k-1}, P_k, *, *, \ldots, *)$.
      Estimate $\sum_{P \in \mathscr{S}_{P_k}} |\widehat{A}_P|^2$ to within $\gamma^2/4$ with a QStat query.
      Add $\mathscr{S}_{P_k}$ to L if the estimate of $\sum_{P \in \mathscr{S}_{P_k}} |\widehat{A}_P|^2$ is at least $\gamma^2/2$.
    end for
    Remove $\mathscr{S}$ from L.
  end for
end for
return L

The GL algorithm returns a list of "heavy-weight" Fourier coefficients. If A is a quantum Boolean function, we can easily recover the values of those coefficients, up to a global sign. We prove this result in the following lemma.

Lemma 6.5. *Let $A = \sum_P \widehat{A}_P P$ a quantum Boolean function and let $L \subseteq \mathscr{P}_n$ a list of Pauli strings. Assume that $|\widehat{A}_P| > \tau/2$ for all P. There is a procedure running in time $\mathscr{O}(|L|)$ that accesses the state $|v(A)\rangle$ via QStat_A queries with tolerance at least τ^2 and outputs some estimates $\{\widehat{B}_P | P \in L\}$ such that*

1. *for all $P \in L$, $\widehat{B}_P = \pm\widehat{A}_P \pm \tau$*
2. *for all $P, Q \in L$, $\mathsf{sgn}\left(\widehat{B}_P \widehat{B}_Q\right) = \mathsf{sgn}\left(\widehat{A}_P \widehat{A}_Q\right)$,*

where $\mathsf{sgn}(\cdot)$ is that function that on input $x \in \mathbb{R}$ returns the sign of x.

Proof. By Lemma 6.3, we can estimate the values of $\widehat{A}_P^2$ up to error τ^2 via a QStat query with tolerance τ^2. Let $\widehat{B}_P^2$ be such estimates. Then we have that

$$|\widehat{B}_P| \leq \sqrt{A_P^2 + \tau^2} \leq |\widehat{A}_P| + \tau, \tag{6.41}$$

which proves the first part of the lemma. It remains to estimate the signs of the coefficients, up to a global sign. Let $P^* = \arg\max \widehat{B}_{P^*}^2$, that is the largest estimated squared coefficient. We arbitrarily assign the positive sign to this coefficient, i.e. we let $\widehat{B}_{P^*} = \sqrt{\widehat{B}_{P^*}^2}$. For each other coefficient $P \neq P^*$,we assign the sign with the following procedure. We first define the following observables M^+ and M^-,

$$M^+ := \left(|v(P^*)\rangle + |v(P)\rangle\right)\left(\langle v(P^*)| + \langle v(P)|\right), \tag{6.42}$$

$$M^- := \left(|v(P^*)\rangle - |v(P)\rangle\right)\left(\langle v(P^*)| - \langle v(P)|\right). \tag{6.43}$$

We now compute the expected values of M^+ with respect to $|v(A)\rangle$:

$$\mu^+ := \langle v(A)| M^+ |v(A)\rangle = \tag{6.44}$$

$$= \left(\sum_{Q\in\mathscr{P}_n} \widehat{A}_Q \langle v(Q)| \left(|v(P^*)\rangle + |v(P)\rangle \right) \right) \left(\left(\langle v(P^*)| + \langle v(P)| \right) \sum_{Q\in\mathscr{P}_n} \widehat{A}_Q |v(Q)\rangle \right) \tag{6.45}$$

$$= (\widehat{A}_{P^*} + \widehat{A}_P)^2, \tag{6.46}$$

and, similarly, for M^-,

$$\mu^- := \langle v(A)| M^- |v(A)\rangle = \tag{6.47}$$

$$= \left(\sum_{Q\in\mathscr{P}_n} \widehat{A}_Q \langle v(Q)| \left(|v(P^*)\rangle - |v(P)\rangle \right) \right) \left(\left(\langle v(P^*)| - \langle v(P)| \right) \sum_{Q\in\mathscr{P}_n} \widehat{A}_Q |v(Q)\rangle \right) \tag{6.48}$$

$$= (\widehat{A}_{P^*} - \widehat{A}_P)^2. \tag{6.49}$$

So if $\widehat{A}_{P^*}$ and $\widehat{A}_P$ have the same sign, $\mu^+ > \mu^-$ and vice-versa. Moreover, $|\mu^+ - \mu^-| = 4\left|\widehat{A}_P \widehat{A}_{P^*}\right| > \tau^2$. Then we can tell whether $\mu^+ > \mu^-$ by querying the oracle QStat_A with the observable $M^+ - M^-$ and tolerance τ^2. If the output is positive, then we can conclude that $\mu^+ > \mu^-$ and assign $\widehat{B}_P$ positive sign, and vice-versa if the output is negative. This proves the second part of the theorem.

■

We can now finally provide a QSQ algorithm for learning quantum Boolean functions. We closely follow the proof of ([RWZ22], Proposition 6.7), which provide an analogous learning algorithm for quantum Boolean functions under oracle query access.

Theorem 6.5 (Learning Quantum Boolean Functions with QSQs)**.** *Let A be a quantum Boolean function. There is a* $\mathrm{poly}(n, 2^k)$*-time algorithm that accesses the state $|v(A)\rangle$ via QStat_A queries with tolerance at least $\Omega(4^{-k})$ and outputs a quantum Boolean function A' such that* $\min\{\|A - A'\|_2, \|A + A'\|_2\} \leq \varepsilon$*, where*

$$k \leq k(\varepsilon) = \begin{cases} \mathrm{Inf}^1(A)^2 \cdot e^{\frac{48\mathrm{Inf}^2(A)}{\varepsilon^2} \log \frac{2\mathrm{Inf}^2(A)}{\varepsilon}} & \text{if } \mathrm{Inf}^2(A) \geq 1, \\ \mathrm{Inf}^1(A)^2 \cdot \mathrm{Inf}^2(A)^{-1} \cdot e^{\frac{48\mathrm{Inf}^2(A)}{\varepsilon^2} \log \frac{2\sqrt{\mathrm{Inf}^2(A)}}{\varepsilon}} & \text{else.} \end{cases} \tag{6.50}$$

Proof. We can adapt the proof of Proposition 6.7 in [RWZ22] to the QSQ setting by replacing all the oracle access queries to A with queries to QStat_A. In particular, this involves the implementation of the GL algorithm with the parameter $\gamma = \Theta(\varepsilon 2^{-k})$. This can be done in time $\mathrm{poly}(n, 2^k, \varepsilon^{-1})$ via quantum statistical queries with tolerance $\Theta(\varepsilon^2 4^{-k})$ by Theorem 6.4. Moreover, we need to evaluate $\mathscr{O}(4^k)$ Fourier coefficients with accuracy $\varepsilon 4^{-k}$. By Lemma 6.5, this can be done, up to a global sign, in time $\mathscr{O}(4^k)$ with quantum statistical queries queries with tolerance $\mathscr{O}(\varepsilon^2 4^{-k})$. The remaining part of the proof doesn't involve oracle access queries, and then is identical to the one of ([RWZ22], Proposition 6.7). ■

Remark 6.2. Theorem 6.5 allows us to learn a quantum Boolean function in Hilbert-Schimdt distance, up to a global sign. In other terms, given a target observable A, we can estimate B such that either B or $-B$ is close to A in Hilbert-Schmidt distance. This enables the prediction of the norm of the expected value for an arbitrary state. This follows by an application of Holder's inequality.

$$\left||\operatorname{Tr}[A\rho]| - |\operatorname{Tr}[B\rho]|\right| \leq \min\left\{\left|\operatorname{Tr}[(A-B)\rho]\right|, \left|\operatorname{Tr}[(A+B)\rho]\right|\right\} \tag{6.51}$$

$$\leq \min\{\|A-B\|_2, \|A+B\|_2\} \cdot \|\rho\|_2 \leq \varepsilon. \tag{6.52}$$

If instead we are interested to the unitary evolution performed by A on a random state, we can observe that:

$$D(A,B) \leq \frac{1}{\sqrt{2^n}} \min_{\theta\in[0,2\pi)} \|e^{i\theta}A - B\|_2 \leq \frac{1}{\sqrt{2^n}} \min\{\|A-B\|_2, \|A+B\|_2\}, \tag{6.53}$$

where the first inequality is proven in (Lemma 14, [BY23]). Moreover, the accuracy guarantees of Theorem 6.5 are cast in terms of $\mathrm{Inf}^1(A), \mathrm{Inf}^2(A)$. These parameters can be bounded for an observable evolved by a shallow circuit (in the Heisenberg picture), by using a variant of the light-cone argument, as done in ([RWZ22], Section 6.1). We now introduce some further notation to state their claim. For any $j \in [n]$, let $N_j \subseteq [m]$ be the minimal set of qubits such that $\frac{\mathrm{Tr}_j}{2}\left(U\frac{\mathrm{Tr}_{N_j}}{2^{|N_j|}}(O)U^\dagger\right) = \left(U\frac{\mathrm{Tr}_{N_j}}{2^{|N_j|}}(O)U^\dagger\right)$ for any $O \in \mathscr{L}_n$ and denote $L := \max_i |\{j : i \in N_j\}|$. Then, if O is a quantum Boolean function with $\mathrm{Inf}^1(O), \mathrm{Inf}^2(O), \|O\|_2 = \mathscr{O}(1)$, and U is a unitary with $L = \mathscr{O}(1)$, we can learn evolution in the Heisenberg picture $U^\dagger OU$ by means of Theorem 6.5 by picking $k = \mathscr{O}(1)$. This ensures that the algorithm runs in poly(n) time and that the statistical queries have constant tolerance.

6.3.3.1 Numerical result

We complement our analysis with a numerical simulation of the proposed algorithm for learning quantum Boolean functions. Given a 4-qubit random unitary U, implemented by a circuit consisting in 2 layers of Haar-random gates and a Pauli string P, we considered the quantum Boolean function $U^\dagger PU$. We implemented the quantum Goldreich-Levin algorithm with quantum statistical queries to estimate the high-weight Pauli coefficients of $U^\dagger PU$, and then we estimated their values, up to a global sign, by means of Lemma 6.5. Finally, we used the estimated quantum Boolean function to output an approximation of $|\operatorname{Tr}[PU|\mathbf{0}\rangle\langle\mathbf{0}|U^\dagger]|$, as depicted in Figure 6.1. For each quantum statistical query with tolerance τ, we computed the expected value exactly and added a noisy perturbation, which we sampled from a normal distribution with mean zero and variance $\tau^2/4$. We tested our algorithm on the observables in $\{I,Z\}^{\otimes 4}$, and we did not witness a significant dependence between the performance and the locality of the observable. The choice of a shallow circuit is motivated by the results in [RWZ22], which establish a connection between the performance of the Goldreich-Levin algorithm to the complexity of the underlying circuit, as also discussed in Remark 6.2.

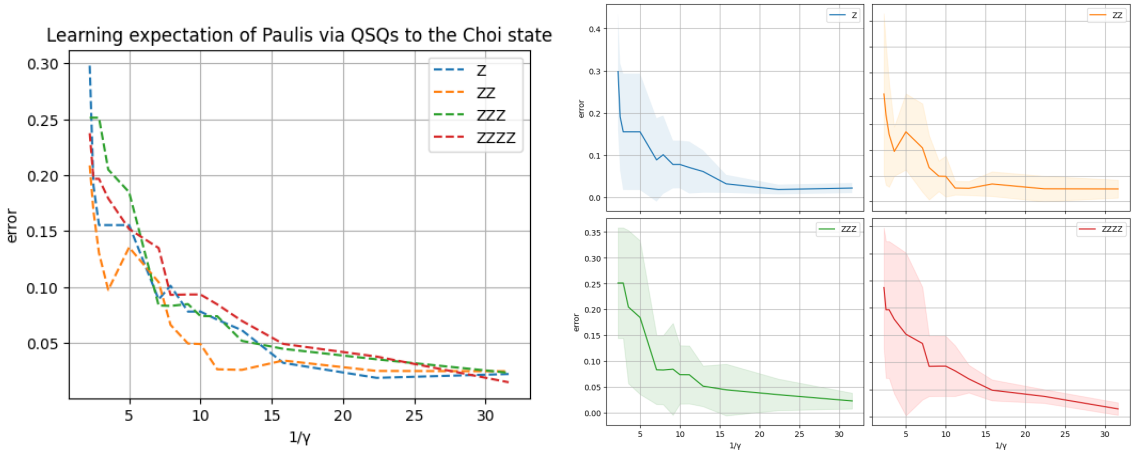


Figure 6.1: Average performance of the Goldreich-Levin algorithm implemented with quantum statistical queries to the Choi-Jamiolkowski state. We tested the algorithm on 10 random 4-qubit random unitaries, in predicting the absolute value of the outcome of Z observables on the unitary evolution of computational basis states. Each random unitary consists in 2 layers of Haar-random gates. We plotted the average error as a function of $1/\gamma$, i.e. the inverse of the threshold of Algorithm 2. We set the tolerance of the quantum statistical queries as $\gamma^2/4$.

6.4 Exponential separations between QSQs and Choi state access

We will now prove a lower bound for learning Choi states with QSQs, and derive from it an exponential separation between learning unitaries from QSQs and learning unitaries with Choi state access. To this end, we combine Lemma 6.1 with an argument given in [AHS23] and based on the following concept class (of classical functions):

$$\mathscr{C} = \left\{ f_A : \{0,1\}^n \to \{0,1\}, f_A(x) = x^\top A x \mod 2 \mid A \in \mathbb{F}_2^{n\times n} \right\} \tag{6.54}$$

Theorem 6.6 (Hardness of learning phase oracles)**.** *The concept class of phase oracle unitaries V_{f_A}, i.e.*

$$\{V_{f_A} \mid A \in \mathbb{F}_2^{n\times n}\} \tag{6.55}$$

requires $2^{\Omega(n)}$ many quantum statistical queries to $\mathsf{QStat}_{V_{f_A}}$ of tolerance $1/\operatorname{poly}(n)$ to be learnt below distance $D < 0.05$ with high probability.

Proof. Our proof is based on the one of ([AHS23], Theorem 17). Their statement is analogous, with the class of quantum examples $|\psi_{f_A}\rangle$ replacing that of unitaries V_{f_A}. The only things we need to prove are the following

$$\| |v(V_{f_A})\rangle\langle v(V_{f_A})| - \mathbb{E}_B |v(V_{f_B})\rangle\langle v(V_{f_B})| \|_{\mathrm{tr}} \geq 1 - \sqrt{17/32}, \tag{6.56}$$

$$\max_{M:\|M\|=1} \mathbb{V}_A \operatorname{Tr}\left[M |v(V_{f_A})\rangle\langle v(V_{f_A})|\right] = 2^{-\Omega(n)} \tag{6.57}$$

and then the result follows from ([AHS23], Theorem 16). The first line follows by checking that

$$\| |v(V_{f_A})\rangle\langle v(V_{f_A})| - \mathbb{E}_B |v(V_{f_B})\rangle\langle v(V_{f_B})| \|_{\mathrm{tr}} = \| |\psi_{f_A}\rangle\langle\psi_{f_A}| - \mathbb{E}_B |\psi_{f_B}\rangle\langle\psi_{f_B}| \|_{\mathrm{tr}} \geq 1 - \sqrt{17/32}, \tag{6.58}$$

Where the lower bound is proven in [AHS23]. As for the variance, we notice the following

$$\begin{aligned} \mathbb{V}_A \operatorname{Tr}\left[M|v(V_{f_A})\rangle\langle v(V_{f_A})|\right] =& \mathbb{E}_A \operatorname{Tr}\left[M|v(V_{f_A})\rangle\langle v(V_{f_A})|^2\right] - \mathbb{E}_A \operatorname{Tr}\left[M|v(V_{f_A})\rangle\langle v(V_{f_A})|\right]^2 & (6.59)\\ =& \mathbb{E}_A \operatorname{Tr}\left[M'|\psi_{f_A}\rangle\langle\psi_{f_A}|^2\right] - \mathbb{E}_A \operatorname{Tr}\left[M'|\psi_{f_A}\rangle\langle\psi_{f_A}|\right]^2 = 2^{-\Omega(n)}, & (6.60) \end{aligned}$$

where the observable M' is the one obtained following the procedure of Lemma 6.1 and the upper bound follows again from [AHS23]. ■

On the other hand, the unitary V_{f_A} is efficiently learnable from separable measurements to Choi states. This is an immediate consequence of a result given in [ABDY22], saying that the function f_A is efficiently learnable from separable measurements to phase states, defined as $|\phi_{f_A}\rangle = 2^{-n/2}\sum_{x\in\{0,1\}^n}(-1)^{f_A(x)}|x\rangle$. We observe the $|v(V_{f_A})\rangle = 2^{-n/2}\sum_{x\in\{0,1\}^n}(-1)^{f_A(x)}|x,x\rangle$, then adapting the argument to Choi states is straightforward.

We also provide a double exponential lower bound for testing properties of channels, hinging on a lower bound for testing purity of states given in [AHS23]. First, recall that the *unitarity* [WGHF15, CDWE19] of a quantum channel is defined as

$$u(\mathcal{N}) := \frac{2^n}{2^n-1}\mathbb{E}_{|\psi\rangle\sim\mu_n} \operatorname{Tr}\left[\mathcal{N}(|\psi\rangle\langle\psi|)^2\right] - \frac{2^n}{2^n-1}\operatorname{Tr}\left[\mathcal{N}\left(\frac{I}{2^n}\right)^2\right] \tag{6.61}$$

Theorem 6.7 (Hardness of testing unitarity)**.** *Let $\mathscr{A}$ be an algorithm that estimates with high probability the unitarity of a quantum channel $\mathcal{N}$ with error smaller than* 0.24 *using* $\mathsf{Qstat}_{\mathcal{N}}$ *queries with tolerance at least τ. Then $\mathscr{A}$ must make at least $2^{\Omega(\tau^2 2^n)}$ such queries.*

Proof. Assume the existence of an algorithm $\mathscr{A}$ contradicting the statement of the theorem. We will prove the theorem by contradiction, by first showing that the unitarity is closely related to the purity of the Choi state $\mathscr{J}(\mathcal{N})$, and then applying the lower bound for testing purity given in ([AHS23], Theorem 25).

$$\begin{aligned} \mathbb{E}_{|\psi\rangle\sim\mu_n} \operatorname{Tr}\left[\mathcal{N}(|\psi\rangle\langle\psi|)^2\right] &= \mathbb{E}_{|\psi\rangle\sim\mu_n} \operatorname{Tr}\left[\mathbb{F}\mathcal{N}^{\otimes 2}(|\psi\rangle\langle\psi|^{\otimes 2})\right] & (6.62)\\ = \operatorname{Tr}\left[\mathbb{F}\mathcal{N}^{\otimes 2}(\mathbb{E}_{|\psi\rangle\sim\mu_n}|\psi\rangle\langle\psi|^{\otimes 2})\right] &= \operatorname{Tr}\left[\mathbb{F}\mathcal{N}^{\otimes 2}\left(\frac{\mathbb{I}+\mathbb{F}}{2^n(2^n+1)}\right)\right] & (6.63)\\ &= \operatorname{Tr}\left[\mathbb{F}\mathcal{N}^{\otimes 2}\left(\frac{\mathbb{F}}{2^n(2^n+1)}\right)\right] + \operatorname{Tr}\left[\mathbb{F}\mathcal{N}^{\otimes 2}\left(\frac{\mathbb{I}}{2^n(2^n+1)}\right)\right] & (6.64)\\ &= \operatorname{Tr}\left[\mathbb{F}\mathcal{N}^{\otimes 2}\left(\frac{\mathbb{F}}{2^n(2^n+1)}\right)\right] + \operatorname{Tr}\left[\mathcal{N}\left(\frac{I}{2^n}\right)^2\right]\frac{2^n}{(2^n+1)} & (6.65) \end{aligned}$$

Then we can rearrange the unitarity as follows

$$u(\mathcal{N}) = \frac{1}{4^n-1}\operatorname{Tr}\left[\mathbb{F}\mathcal{N}^{\otimes 2}(\mathbb{F})\right] - \frac{1}{4^n-1}\operatorname{Tr}\left[\mathcal{N}\left(\frac{I}{2^n}\right)^2\right] \tag{6.66}$$

We can also use the Kraus representation $\mathscr{N}(\cdot) = \sum_\ell K_\ell(\cdot)K_\ell^\dagger$ and write

$$\mathrm{Tr}\left[\mathbb{F}\mathscr{N}^{\otimes 2}(\mathbb{F})\right] = \sum_{\ell,\ell'} \mathrm{Tr}\left[\mathbb{F}(K_\ell \otimes K_{\ell'})\mathbb{F}(K_\ell^\dagger \otimes K_{\ell'}^\dagger)\right] \tag{6.67}$$

$$= \sum_{\ell,\ell'} |\mathrm{Tr}[K_\ell K_{\ell'}^\dagger]|^2 = 4^n \mathrm{Tr}[J(\mathscr{N})^2], \tag{6.68}$$

where the last two identities are proven in ([QFK+22], Eqs. 160-164). Putting all together, we obtain:

$$\frac{4^n}{4^n-1}\mathrm{Tr}[J(\mathscr{N})^2] - \frac{1}{4^n-1} \leq u(\mathscr{N}) \leq \frac{4^n}{4^n-1}\mathrm{Tr}[J(\mathscr{N})^2] \tag{6.69}$$

Thus the unitarity of $\mathscr{N}$ and the purity of $\mathscr{J}(\mathscr{N})$ are within an exponentially small additive terms. Then the algorithm $\mathscr{A}$ would estimate the purity of $\mathscr{J}(\mathscr{N})$ with error smaller than $0.24 + 1/(4^n-1)$ with less than $2^{\Omega(\tau^2 2^n)}$ queries, contradicting ([AHS23], Theorem 25). ■

It's easy to see that the unitarity can be estimated with $\mathscr{O}(1)$ joint measurements to the Choi state or $\mathscr{O}(2^n)$ separable measurements to the Choi state . This can be shown invoking previous upper bounds for purity estimation [MdW13, CCHL22b] and exploiting again the connection between unitarity and the purity of the Choi state.

6.5 Application: Classical Surrogates

In this section we discuss a potential application of our results to quantum machine learning. We will consider particularly variational quantum algorithms for approximating a classical function $f : \mathscr{X} \to \mathbb{R}$. For a broad class of such algorithms [SK19, Sch21], the prediction phase can be cast as follows: the input $\boldsymbol{x} \in \mathscr{X}$ is encoded into a quantum state with a suitable feature map $\boldsymbol{x} \mapsto \rho(\boldsymbol{x})$, which evolves according to a parametric channel $\mathscr{U}_{\boldsymbol{\theta}}$ and subsequently is measured with a local observable O. Hence, the parametric circuit induces a hypothesis function $h(\cdot)$, which associates $\boldsymbol{x}$ to the following label

$$h(\boldsymbol{x}) = \mathrm{Tr}[O\mathscr{U}_{\boldsymbol{\theta}}(\rho(\boldsymbol{x}))]. \tag{6.70}$$

Thus, given a distribution $\mathscr{D}$ over $\mathscr{X}$, the goal is to find a parameter $\boldsymbol{\theta}^*$ satisfying the following:

$$\mathbb{E}_{\boldsymbol{x}\sim\mathscr{D}}|h(\boldsymbol{x}) - f(\boldsymbol{x})| = \mathbb{E}_{\boldsymbol{x}\sim\mathscr{D}}|\mathrm{Tr}[O\mathscr{U}_{\boldsymbol{\theta}^*}(\rho(\boldsymbol{x}))] - f(\boldsymbol{x})| \leq \varepsilon, \tag{6.71}$$

where ε is a small positive constant. Given a set of examples $(\boldsymbol{x}_1, f(\boldsymbol{x}_1)), (\boldsymbol{x}_2, f(\boldsymbol{x}_2)), \ldots, (\boldsymbol{x}_m, f(\boldsymbol{x}_m))$ one can then train this model in a hybrid fashion and select a parameter $\boldsymbol{\theta}$. Then the label of an unseen instance $\boldsymbol{x}_{m+1}$ can be predicted with accuracy ε preparing $\mathscr{O}(\varepsilon^{-2})$ copies of the state $\mathscr{U}_{\boldsymbol{\theta}}(\rho(\boldsymbol{x}))$ and measuring the observable O.

A recent line of research showed that, in some cases, one can fruitfully perform the prediction phase with a purely classical algorithm, that goes under the name of *classical surrogate* [SEM23]. So far, the proposed approaches rely on the classical shadow tomography [JGM+23] and the Fourier

analysis of real functions [LTD+22, SEM23], which can be applied to the general expression of quantum models as trigonometric polynomials. Here we argue that the QSQ learning framework can find application in the quest for surrogate models, introducing more flexibility in the surrogation process. Particularly, [JGM+23] resorts to a flipped model of quantum circuit where the parameter $\boldsymbol{\theta}$ is encoded in a quantum state, subsequently measured by a variational measurements depending on the $\boldsymbol{x}$. While this model can provide quantum advantage for specific tasks, it would be interesting to obtain similar results beyond the flipped circuit model, and specifically for the setting where the instance $\boldsymbol{x}$ is encoded before the parameter $\boldsymbol{\theta}$. This goal can be achieved through the algorithms discussed in the present chapter, since they do not require the unitary to be a flipped a circuit. However, the distance over unitaries we adopted brings accuracy guarantees for the prediction only when the input state is sampled from a locally scrambled ensemble. Thus, we need to extend the definition given in [SEM23] to incorporate the input distribution $\mathscr{D}$.

Definition 6.2 (Worst-case and average-case surrogate models)**.** Let $\varepsilon \geq 0$ and $0 \leq \delta \leq 1$. A hypothesis class of quantum learning models $\mathscr{F}$ has a worst-case (ε,δ)-classical surrogate if there exists a process $\mathscr{S}$ that upon input of a learning model $f \in \mathscr{F}$ produces a classical model $g \in \mathscr{G}$ such that

$$\Pr\left[\sup_{\boldsymbol{x}\in\mathscr{X}} \|f(\boldsymbol{x}) - g(\boldsymbol{x})\| \leq \varepsilon\right] \geq 1-\delta, \tag{6.72}$$

for a suitable norm on the output space $\mathscr{Y}$. Similarly, we say that $\mathscr{F}$ has an average-case (ε,δ)-classical surrogate if there exists a process $\mathscr{S}$ that upon input of a learning model $f \in \mathscr{F}$ produces a classical model $g \in \mathscr{G}$ such that

$$\Pr[\mathbb{E}_{\boldsymbol{x}\sim\mathscr{D}} \|f(\boldsymbol{x}) - g(\boldsymbol{x})\| \leq \varepsilon] \geq 1-\delta. \tag{6.73}$$

The process $\mathscr{S}$ must be efficient in the size of the quantum learning model, the error bound ε and the failure probability δ.

In particular, it is easy to see that if the conditional distribution of the states $\rho(\boldsymbol{x})$ is locally scrambled, then we can produce an average-case classical surrogate of $f(\boldsymbol{x}) = \mathrm{Tr}[O\mathscr{U}_{\boldsymbol{\theta}}\rho(\boldsymbol{x})]$ via QSQs by means of Theorems 6.2,6.3,6.5 . For instance, if $\mathscr{D}$ is the uniform distribution over $[6]^n$, the ensemble $\{|\phi(\boldsymbol{x})\rangle\}_{\boldsymbol{x}}$ defined as follows is locally scrambled. We have:

$$|\phi(\boldsymbol{x})\rangle = \bigotimes_{i=1}^{n} |\phi(x_i)\rangle \quad \text{where} \quad |\phi(x_i)\rangle = \begin{cases} |0\rangle & \text{if } x_i = 1 \\ |1\rangle & \text{if } x_i = 2 \\ |+\rangle & \text{if } x_i = 3 \\ |-\rangle & \text{if } x_i = 4 \\ |y_+\rangle & \text{if } x_i = 5 \\ |y_-\rangle & \text{if } x_i = 6. \end{cases} \tag{6.74}$$

While this example is just meant to motivate our definition of average-case surrogate models, the quest for quantum encodings mapping a target distribution over $\mathscr{X}$ to a locally scrambled distribution

would be of primary importance for the design of surrogation processes. We also remark that worst-case surrogate models could be found by means of the quantum Goldreich-Levin algorithm, and in particular by exploiting the fact the unitary evolution in the Heisenberg picture of a Pauli string $P \in \mathscr{P}_n$, i.e. $\mathscr{U}_{\boldsymbol{\theta}}^{\dagger}(P) = U_{\boldsymbol{\theta}}^{\dagger} P U_{\boldsymbol{\theta}}$, is a quantum Boolean function. This follows from the fact that the accuracy guarantees of Theorem 6.5 expressed in Hilbert-Schimdt distance can be transferred to an arbitrary state, as noted in Remark 6.2. This will allow learning $\mathscr{U}_{\boldsymbol{\theta}}^{\dagger}(P)$, up to a multiplicative sign, and hence to predict functions of the form

$$h(\boldsymbol{x}) = \left|\mathrm{Tr}[P\mathscr{U}_{\boldsymbol{\theta}}(\rho(\boldsymbol{x}))]\right|. \tag{6.75}$$

CHAPTER 7

DIFFERENTIAL PRIVACY: AN OVERVIEW

Is a secret still a secret if everyone knows it?

- George R. R. Martin, *A Clash of Kings* – Tyrion Lannister

In recent years, the availability of large datasets and advanced computational tools has sparked progress across various fields, including natural sciences, medicine, finance, and social sciences. This advance came also with privacy concerns since even the release of aggregated data can compromise the sensitive information contained in the original dataset. This poses a significant challenge for the researcher, who must adopt privacy-preserving techniques to avoid the exposure of private data. This motivated the quest for a robust framework to assess privacy, that eventually led to the wide adoption of differential privacy as the *de facto* standard for ensuring privacy both in statistical data analysis and machine learning applications [DMNS06, DR14, CDE+23, CMS11b, ACG+16, PAE+16, BTT18]. Notably, the Census Bureau of United States adopted differential privacy [Abo18, AACM+22], and several industrial applications were also deployed [CJK+18, RSP+21, XZA+23].

In this Chapter we will first review some competing notions of privacy and show that they can lead to impressive privacy breaches [NS07], and thus provide a self-contained introduction to differential privacy, highlighting some previous applications to quantum computation.

Intuitively, a differentially private algorithm $\mathscr{A}(\cdot)$ can learn a statistical property of a dataset consisting of n elements, yet it leaks *almost* nothing about each individual element. In other words, given two inputs x and x' which are very close according to some chosen metric, the output distributions $\mathscr{A}(x)$ and $\mathscr{A}(x')$ should be almost indistinguishable. We call x and x' neighbouring inputs. If x and x' represent datatsets about n individuals, then it's customary to consider x and x' neighbouring if one of such individuals is present in x and absent in x'. Then, if $\mathscr{A}(\cdot)$ is differentially private, the output alone doesn't allow for inferring whether the input contained a given individual. This goal is pursued by combining various techniques, that usually involve randomising the input or perturbing the output by adding noise. The challenge is then to achieve the desired level of privacy by adding less noise as possible, hence preserving accuracy.

Apart from privacy-preserving data analysis and machine learning, differential privacy has also found several applications in other fields of computer science such as statistical learning theory [KLN$^+$11a, WLF16, BLM20, AQS21], adaptive data analysis [DFH$^+$15, BNS$^+$21, FS17] and mechanism design [MT07].

7.1 Anonymization or pseudonymization?

One of the main goals of privacy-preserving techniques is to protect *personal identifiable information*, or *personal data*, within a publicly accessible dataset. The EU General Data Protection Regulation (GDPR)[gdp] defines personal data as follows:

> 'Personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

Thus, the wide scope of personal data accounts for biometric, demographic and financial information, but also less explicit information such as daily habits, judgements and opinions, including political beliefs. Then privacy-preserving techniques aim at anonymizing personal data, that is preventing any possible association between the individuals and their personal data. To this end, privacy specialists have devised a number of different techniques. Whereas some of them, such as differential privacy, come with robust security guarantees, several methods are based on non-rigorous heuristics. The latter category includes the so-called *pseudonymization* technique. As the name suggests, this technique does not achieve the goal of fully anonymizying the personal data, but instead it provides a partial obfuscation. Pseudonymization can be achieved by replacing all the personal identifiable

information with artificial identifiers, called pseudonyms. This heuristic makes the record less identifiable, without compromising its utility. Moreover, the European Data Protection Board and the European Commission endorsed the adoption of pseudonymization as a state-of-the-art measure for the compliance of GDPR. In particular, GDPR provides the following definition of pseudonymization:

> 'Pseudonymization' means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

Thus, by the definition, this technique is effective when the malicious party does not hold any additional information, which is a rather optimistic scenario. To see why, we can consider one of the most spectacular de-anonymization attack of the last few years, namely the one performed on the Netflix Prize Dataset by Arvind Narayanan and Vitaly Shmatikov [NS06]. The Netflix Prize was a competition held from 2006 to 2009, whose participants were asked to design an algorithm to predict user ratings for films, based exclusively on previous ratings. To this end, a "pseudonymized" training set was publicly released, where the users' identities were replaced by random numbers. The competition was cancelled in 2010 due to a class action lawsuit against Netflix, based on privacy concerns, in particular those arising from the vulnerabilities exposed by the attack performed in [NS06]. Such attack is based on a fairly simple intuition: in order to break pseudonymization, one only needs to retrieve some background knowledge about the user ratings. This can be easily obtained, for instance, from other publicly accessible ratings datasets, such as the Internet Movie Database (IMDb). Comparing the ratings on IMDb with the Netflix dataset, the researchers were able to perform a so-called *linkage attack*, and therefore recognise the Netflix records of known users, disclosing potentially sensitive information, such as their political beliefs.

Privacy breaches akin to the one described have been documented on other occasions as well. For instance, incidents involving the Massachusetts Group Insurance Commission in the mid-1990s and the web portal America Online in 2006 serve as cautionary tales. For a more comprehensive discussion on this topic, we refer the reader to [Ohm09].

The obvious shortcomings of pseudonymization led to the establishment of differential privacy as a more robust framework to protect sensitive data. In the words of Cynthia Dwork, differential privacy can be succinctly described as follows:

> Differential privacy describes a promise, made by a data curator to a data subject: you will not be affected, adversely or otherwise, by allowing your data to be used in any study, no matter what other studies, data sets, or information from other sources is available.

In the following section we will introduce the basic notions of differential privacy, showing how the informal definition above can be translated in a rigorous mathematical theory. Whereas no technique achieve a perfect anonymization while keeping the usability of the data, we will see that differential

privacy has the advantage to provide a quantitative measure of the tradeoff between anonymization and usability.

7.2 Mathematical foundations of differential privacy

We concisely introduce the definition of differential privacy. For a comprehensive introduction to the topic, we refer to [DR14], [Vad17] and [CDE+23]. Throughout this thesis, we'll denote by $\sim$ the neighbouring condition, i.e. a relationship between two inputs, consisting of either classical vectors or quantum states. We'll write $\overset{Q}{\sim}$ when we want to emphasise that the neighbouring relationship refers to quantum states. The choice of the relationship is problem-dependent. In many practical cases, it's convenient to say that two binary vectors $x, x' \in \{0,1\}^n$ are neighbouring if their Hamming distance is at most one, i.e.

$$x \sim x' \iff d_H(x,x') \leq 1.$$

In alternative, we can select a p-norm and a threshold $\gamma \geq 0$ and opt for the following neighbouring relationship:

$$x \sim x' \iff \|x - x'\|_p \leq \gamma.$$

We say that a randomised algorithm $\mathscr{A}(\cdot)$ is (ε,δ)-differentially private (DP) if for all $x \sim x'$ and for all $S \subseteq \mathsf{range}(\mathscr{A})$, it satisfies

$$\Pr[\mathscr{A}(x) \in S] \leq e^{\varepsilon} \Pr[\mathscr{A}(x') \in S] + \delta.$$

We say that $\mathscr{A}(\cdot)$ is ε-DP when it is $(\varepsilon,0)$-DP. Equivalently, differential privacy can be defined in terms of hockey-stick divergence E_γ and the smooth max-relative entropy (or smooth max-divergence) D_∞^δ:

$$\mathscr{A} \text{ is } (\varepsilon,\delta)\text{-DP} \iff \forall x \sim x' : E_{e^\varepsilon}(\mathscr{A}(x)\|\mathscr{A}(x')) \leq \delta \iff \forall x \sim x' : D_\infty^\delta(\mathscr{A}(x)\|\mathscr{A}(x')) \leq \varepsilon, \tag{7.1}$$

where the (classical) hockey-stick divergence E_γ between two distributions P and Q is defined as follows [PPV10]:

$$E_\gamma(P\|Q) := \frac{1}{2}\int |\mathrm{d}P - \gamma \mathrm{d}Q| - \frac{1}{2}(\gamma - 1),$$

for $\gamma \geq 1$. These information-theoretic divergences can be thought of as a measure of closeness between distributions, thus these reformulations are consistent with the intuition that private algorithms map neighbouring inputs to "close" output distributions. Differential privacy with $\delta = 0$ is also referred to as *pure* differential privacy, whereas the case with $\delta \neq 0$ is referred to as *approximate* differential privacy. Roughly speaking, an (ε,δ)-DP algorithm can be thought of as an algorithm that is ε-DP with probability $1-\delta$. We remark that this intuition is slightly imprecise, and thus we refer to the following references for a more detailed explanation [BS16, Mei18, Vad17].

It's also worth noticing that the max-divergence corresponds to the Rényi divergence of order ∞. Thus, it's possible to relax pure differential privacy by replacing the max-divergence with the Rényi

divergence of order α, for $\alpha \geq 1$ [Mir17]. We say that $\mathscr{A}$ is (α,ε)-RDP (Rényi differentially private) if for all $x \sim x'$,

$$D_\alpha(\mathscr{A}(x)\|\mathscr{A}(x')) \leq \varepsilon.$$

As a consequence, for all $S \subseteq \mathsf{range}(\mathscr{A})$, we have

$$\Pr[\mathscr{A}(x) \in S] \leq e^\varepsilon \Pr[\mathscr{A}(x') \in S]^{(\alpha-1)/\alpha}.$$

If $\mathscr{A}$ is (α,ε)-RDP then it is also $(\varepsilon + \frac{\log(1/\delta)}{\alpha-1},\delta)$-DP for any $0 < \delta < 1$. Similarly, if $\mathscr{A}$ is $(\varepsilon,0)$-DP then it is also $(\alpha, 2\alpha\varepsilon^2)$-RDP for any $\alpha \geq 1$.

7.2.1 Privacy via classical noisy channels

Now we present two widely used mechanisms that ensure differential privacy by injecting noise into the output. To this end, we introduce two classical channels $\Lambda_{\mathscr{L},b} : \mathbb{R} \to \mathbb{R}$ and $\Lambda_{\mathscr{G},\sigma} : \mathbb{R} \to \mathbb{R}$, that corresponds to an additive noise coming from either the Laplace distribution of scale b or the Gaussian distribution of variance σ^2, both centred in zero. The channels are defined as follows:

$$\Lambda_{\mathscr{L},b}(x) = x + \eta \qquad \text{where} \quad \eta \sim \frac{1}{2b}\exp\left(-\frac{|\eta|}{b}\right) \quad \text{(Laplace channel)} \tag{7.2}$$

$$\Lambda_{\mathscr{G},\sigma}(x) = x + \zeta \qquad \text{where} \quad \zeta \sim \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{\zeta^2}{2\sigma^2}\right) \quad \text{(Gaussian channel)} \tag{7.3}$$

Let $f : \mathscr{X} \to \mathbb{R}$ be a scalar function. We define the *sensitivity* of f as

$$\Delta_f := \max_{\substack{x,x' \in \mathscr{X} \\ x \sim x'}} |f(x) - f(x')|. \tag{7.4}$$

We can use either the Laplace or the Gaussian channel to ensure differential privacy, by calibrating the noise rate with respect to the sensitivity of the target function. Then $\Lambda_{\mathscr{L},b}(f(\cdot))$ is ε-DP if $b \geq \Delta/\varepsilon$. Similarly, $\Lambda_{\mathscr{G},\sigma}(f(\cdot))$ is (ε,δ)-DP if $\sigma^2 \geq 2\ln(1.25/\delta)\Delta^2/\varepsilon^2$. The addition of Laplace noise is referred to as *Laplace mechanism* [DMNS06], whereas the addition of Gaussian noise is referred to as *Gaussian mechanism* [DR14]. Both mechanisms can also be analysed within the relaxed framework of Rényi differential privacy [Mir17].

7.3 Local differential privacy

An algorithm $\mathscr{A}$ is (ε,δ)-locally differentially private (LDP) if it is (ε,δ)-differentially private and, moreover, every possible pair of inputs x, y are considered neighboring. In other terms, for all x, y and for all $S \subseteq \mathrm{range}(\mathscr{A})$, we have

$$\Pr[\mathscr{A}(x) \in S] \leq e^\varepsilon \Pr[\mathscr{A}(y) \in S] + \delta, \tag{7.5}$$

and therefore

$$\mathscr{A} \text{ is } (\varepsilon,\delta)\text{-LDP} \iff \forall x, y : E_{e^\varepsilon}(\mathscr{A}(x)\|\mathscr{A}(y)) \leq \delta \iff \forall x, y : D_\infty^\delta(\mathscr{A}(x)\|\mathscr{A}(y)) \leq \varepsilon. \tag{7.6}$$

Local differential privacy brings a considerably stronger notion of security and it usually employed when a dataset is managed by an untrusted curator. For instance, let $x_1, x_2, \dots, x_k$ be the personal data of k distinct parties, and let $\mathscr{C}$ be a curator that wants to collect such data to conduct a research. To this end, each party will randomize its own personal data via an (ε, δ)-LDP algorithm $\mathscr{A}$. Denote by z_i the output obtained by the i-th party, i.e. $z_i \sim \mathscr{A}(x_i)$. Then the curator receives as input $z_1, z_2, \dots, z_k$ and outputs a value Y. By robustness to post-processing, we have, for all $i \in [k]$, and for all x_i, y_i,

$$\Pr[\mathscr{C}(z_1, z_2, \dots, z_k) = Y | x_1, x_2, \dots, x_i, \dots, x_k] \tag{7.7}$$

$$\leq e^{\varepsilon} \Pr[\mathscr{C}(z_1, z_2, \dots, z_k) = Y | x_1, x_2, \dots, \underbrace{y_i}_{\text{replaced element}}, \dots, x_k] + \delta. \tag{7.8}$$

In particular, this holds even if the curator releases the entire randomized dataset, i.e. if $\mathscr{C}$ is the identity function. In this case, we can regard $z_1, z_2, \dots, z_k$ as a synthetic dataset, that is a new dataset generated from the original one, which can be reused multiple times without compromising the privacy the individuals represented in the original dataset. Clearly, synthetic data generation is useful if the old and new datasets share some meaningful statistical property, for instance if k-th moments of the associated distribution are sufficiently close.

7.3.1 Randomized response

We now introduced the most popular approach to local differential privacy, that is the randomized response mechanism. Surprisingly, randomized response is older than differential privacy itself, as this method as been devised in 1965 in the context of structured survey interviews by the economist and statistician Stanley L. Warner [War65]. As Warner observed, the individuals in a sample survey may be reluctant to respond faithfully for several reasons, and may even deliberately provide a false information. Then, the author of the survey could instead propose the interviewee to perform the following experiment. Assume that the interview consists in a yes-no question and that the correct answer is "yes" and let $p > \frac{1}{2}$.

1. With probability p, the interviewee responds "yes", therefore providing the correct information.
2. With the remaining probability $1 - p$, the interviewee responds "no".

Collecting answers from a sufficiently large sample of N individuals, the interviewer can approximately estimate the frequency f of the population for which "yes" is the correct answer. To this end, it suffices to observe that the expected number of "yes" answers in the survey is

$$\frac{1}{N} \cdot \mathbb{E}[\text{number of participants responding "yes"}] = pf + (1-p)(1-f) \tag{7.9}$$

$$= f(2p-1) + 1 - p. \tag{7.10}$$

Thus, by standard concentration of measure, it is easy to see that f can be estimated up to additive error ε with $N = \Theta(\varepsilon^{-2}(2p-1)^{-1})$ samples. We will now turn to the analysis of the privacy guarantees.

For convenience, we set $p = \frac{e^\varepsilon}{1+e^\varepsilon}$. Then we have

$$\Pr[\text{“yes” answer} \mid \text{ground truth is “yes”}] = \Pr[\text{ “no” answer}\mid \text{ground truth is “no”}] = \frac{e^\varepsilon}{1+e^\varepsilon}, \tag{7.11}$$

$$\Pr[\text{“yes” answer} \mid \text{ground truth is “no”}] = \Pr[\text{ “no” answer} \mid \text{ground truth is “yes”}] = \frac{1}{1+e^\varepsilon}. \tag{7.12}$$

Rearranging, we obtains

$$\Pr[\text{ “yes” answer} \mid \text{ground truth is “yes”}] = e^\varepsilon \Pr[\text{ “no” answer}\mid \text{ground truth is “yes”}], \tag{7.13}$$

$$\Pr[\text{ “yes” answer}\mid \text{ground truth is “no”}] = e^\varepsilon \Pr[\text{“no” answer} \mid \text{ground truth is “no”}], \tag{7.14}$$

which prove that the randomized response mechanism with $p = \frac{e^\varepsilon}{1+e^\varepsilon}$ is ε-LDP.

This method can be easily extended beyond the binary case. Assume that the possible answers belong to the set $[k] = \{1,2,\dots,k\}$ and that the correct answer is j. Then the k-ary randomized response works as follows.

1. With probability $\frac{e^\varepsilon}{k-1+e^\varepsilon}$, the interviewee provides the correct answer j.
2. With the remaining probability, the interviewee provides a random answer sampled uniformly from $[k] \setminus \{j\}$.

The privacy analysis is identical to the one of the binary case. As for the analysis of the accuracy, we refer to [KOV14, KBR16] for an extensive treatment.

7.3.2 Information-theoretic interpretation of local privacy

Local differential differential privacy poses a severe constraint on the amount of information that can be conveyed to the curator. From an information theory perspective, such data shrinkage admits a strikingly simple interpretation: local differential privacy is equivalent to a contraction of the hockey-stick divergence [AAC21a]. In particular, for two discrete distribution P and Q with equal support on a set $\mathscr{X}$, we have

$$\mathscr{A} \text{ is } (\varepsilon,\delta)\text{-LDP} \iff E_{e^\varepsilon}(\mathscr{A}(P)\|\mathscr{A}(Q)) \le \delta E_{e^\varepsilon}(\mathscr{A}(P)\|\mathscr{A}(Q)), \tag{7.15}$$

where $\mathscr{A}(P)$ is the distribution obtained by applying the algorithm $\mathscr{A}$ on an input sampled from P. Crucially, the hockey-stick divergence underlies a wide family of divergences, known as f-divergences, in a sense that any arbitrary f-divergence can be represented by an integral sum of hockey-stick divergences [Csi64]. Recall that, for a twice differentiable convex function f on non-negative reals with $f(1) = 0$, the associated f-divergence is defined as

$$D_f(P\|Q) := \sum_{x\in\mathscr{X}} Q(x) f(P(x)/Q(x)). \tag{7.16}$$

Then by ([SV16], Proposition 3) we have

$$D_f(P\|Q) = \int_1^\infty f''(\gamma) E_\gamma(P\|Q) + \gamma^{-3} f''(\gamma^{-1}) E_\gamma(Q\|P) \mathrm{d}\gamma. \tag{7.17}$$

Hinging on this result, it is possible to extend Equation 7.15 to all f-divergences, as showed in ([AAC21a], Lemma 1). For any pair of distributions P, Q over $\mathscr{X}$, we have

$$\mathscr{A} \text{ is } (\varepsilon,\delta)\text{-LDP} \implies D_f(\mathscr{A}(P)\|\mathscr{A}(Q)) \leq (1-(1-\delta)e^{-\varepsilon}) \cdot D_f(\mathscr{A}(P)\|\mathscr{A}(Q)). \tag{7.18}$$

This result has far-reaching implication. In particular, we recall that the relative entropy, also referred as Kullback-Leibler (KL) divergence, is an f-divergence with $f(x) = x\log x$. Since the relative entropy plays a central role in hypothesis testing, Equation 7.18 can be used to show a "differentially private" version the celebrated Chernoff-Stein lemma, as we detail below. Furthermore, several implications to private estimation theory are discussed in [AAC21a].

7.3.3 Private hypothesis testing

We will now focus on the widely recognized problem of binary hypothesis testing with the additional constraint of local differential privacy. Let P, Q two distributions over $\mathscr{X}$, which are usually referred as the *null* hypothesis and the *alternative* hypothesis, respectively. Consider a scenario where we have n independent and identically distributed (i.i.d.) samples denoted as $x_1, x_2, \ldots, x_n$, drawn either from P or Q, and we want to distinguish between these two possibilities. Moreover, we assume that each x_i is subsequently transformed into $z_i = \mathscr{A}_i(x_i)$ where $\mathscr{A}_i$ is a suitable (ε,δ)-LDP algorithm. The algorithms $\mathscr{A}_1, \mathscr{A}_2, \ldots, \mathscr{A}_n$ can be chosen in an interactive fashion. Given a decision rule, that is a randomized algorithm $\mathscr{T}$ that takes as input $z_1, z_2, \ldots, z_n$ and outputs either P or Q. We denote by $\alpha_n^{\varepsilon,\delta}(\mathscr{T})$ the probability of outputting Q when the underlying distribution is P (type I error), and vice versa we $\beta_n^{\varepsilon,\delta}(\mathscr{T})$ the probability of outputting P when the underlying distribution is Q (type II error). To achieve the ideal balance between type I and type II error probabilities, we set a constant $\tau \in (0,1)$ and define the following quantity:

$$\beta_{n,\tau}^{\varepsilon,\delta} := \min_{\mathscr{T}} \left\{ \beta_n^{\varepsilon,\delta}(\mathscr{T}) \middle| \alpha_n^{\varepsilon,\delta}(\mathscr{T}) \leq \tau \right\}. \tag{7.19}$$

The following asymptotic lower bound on $\beta_{n,\tau}^{\varepsilon,\delta}$ was provided in ([AAC21a], Corollary 4),

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_{n,\tau}^{\varepsilon,\delta} \geq -(1-(1-\delta)e^{-\varepsilon}) \cdot D(P\|Q) \qquad \text{(Private Stein's lemma).} \tag{7.20}$$

It is insightful to compare this result with the asymptotic limit obtained without the privacy constraint, i.e. when $\delta = 1$. Denote by $\beta_{n,\tau} := \beta_{n,\tau}^{\varepsilon,1}$ the coefficient associated to the non-private case. Following ([Cov99], Theorem 11.8.3), we have

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_{n,\tau} = -D(P\|Q) \qquad \text{(Chernoff-Stein's lemma).} \tag{7.21}$$

Combining these two results, we obtain that, for sufficiently large values of n, the constraint of local differential privacy leads to an increase in sample complexity of a factor $1-(1-\delta)e^{-\varepsilon}$. In Chapter 9, we will prove a quantum analogue of the private Stein's lemma, which holds for small value of the privacy parameter ε. We achieve this goal by proving new entropy inequality for the quantum relative entropy under local privacy.

7.3.4 Equivalence with statistical query learning

As demonstrated in the work of Kasiviswanathan et al. [KLN+11b], local differentially private algorithms can be effectively characterized by means of statistical queries, a concept originally introduced in [Kea98b]. To elucidate this finding, we shall first provide some essential definitions. Consider a probability distribution $\mathscr{D}$ defined over a domain $\mathscr{X}$.

A statistical query (SQ) oracle, denoted as $\mathsf{SQ}_{\mathscr{D}}$, accepts as input a function $g : \mathscr{X} \to [-b, b]$ and a tolerance parameter $\tau \in (0, 1)$, yielding an output value v that adheres to the following condition:

$$|v - \mathbb{E}_{u\sim\mathscr{D}}[g(u)]| \leq \tau. \tag{7.22}$$

A statistical query (SQ) algorithm interacts with the distribution $\mathscr{D}$ by employing the SQ oracle $\mathsf{SQ}_{\mathscr{D}}$. We are now ready to state the equivalence result of [KLN+11b].

Theorem 7.1. *Let $(x_1, x_2, \ldots, x_n) \sim \mathscr{D}^m$ be a set of points sampled i.i.d. from the distribution $\mathscr{D}$. Then if $m \geq c \cdot \frac{\log(1/\beta)b^2}{\varepsilon^2\tau^2}$, there exists a pair of algorithm $\mathscr{A}$ and $\mathscr{B}$ such that:*

1. *$\mathscr{A}$ satisfies ε-local differential privacy;*
2. *$\mathscr{B}$ receives as input $\mathscr{A}(x_1), \mathscr{A}(x_2), \ldots, \mathscr{A}(x_n)$ and approximates $\mathbb{E}_{u\sim\mathscr{D}}[g(u)]$ within additive error $\pm\tau$ with probability at least $1 - \beta$.*

Theorem 7.2. *Let $\mathscr{A}$ be an ε-LDP algorithm that takes as input a point x sampled from the distribution $\mathscr{D}$. Then there exists a statistical query (SQ) algorithm that in expectation makes e^{ε} queries to $\mathsf{SQ}_{\mathscr{D}}$ with accuracy $\tau = \Theta(\beta/(e^{2\varepsilon}))$, such that the total variation distance between $\mathscr{A}$'s and $\mathscr{B}$'s output distributions is at most β.*

7.4 Quantum differential privacy

More recently, the major influence of quantum computing and quantum information has led to the exploration of differentially private quantum algorithms. Since many near-term quantum algorithms involve a classical optimiser as a subroutine, one possible approach consists in privatising such optimiser and leaving the rest of the algorithm unchanged. This strategy is adopted in [SMT17, LLD21, DHL+22, WCY23].

Alternatively, we can rely on several notions of *quantum* differential privacy. Quantum differential privacy allows the design of private measurements and channels combining classical and quantum noise. This is extremely relevant with the emergence of Noisy Intermediate Scale Quantum devices (NISQ) today [Pre18b]. The noisy nature of these devices on the one hand, and the potential capabilities of quantum algorithms, on the other hand, make such quantum or hybrid quantum-classical mechanisms, an interesting subject of study from the point of view of privacy. Several efforts has been made in this area of research, including [ZY17b, AR19, HRF23, Far23, NGW23]. Furthermore, the connection between machine learning and differential privacy [FS17, LAG+19] suggests that

exploring quantum differential privacy can lead to intriguing insights into the capabilities of quantum machine learning.

One of the main challenges in translating the definition of DP in the quantum setting is to characterise the notion of neighbouring quantum states, i.e. choose the right metric to measure the similarity between the input states. The first notion of quantum differential privacy was proposed in [ZY17b] and it's based on bounded trace distance, whereas the definition introduced in [AR19] is based on reachability by a single-qudit operation. Another possible definition is based on the quantum Wasserstein distance of order 1. This metric was introduced in [DPMTL21b] and the authors mention quantum differential privacy as one potential application of their work. Furthermore, quantum private PAC learning has been defined in [AGY20] and a quantum analogue of the equivalence between private classification and online prediction has been shown in [AQS21]. Moreover, an equivalence between learning with quantum local differential privacy and quantum statistical query (QSQ) learning was provided in [AK22b]. Other authors compared classical and quantum mechanisms in the context of local differential privacy [YH20, Yos21]. Building upon these prior contributions, we aim at establishing a general framework for differentially private quantum algorithms, providing a more general definition of neighbouring quantum states and attaining better privacy guarantees combining classical and quantum noisy channels.

7.4.1 Definition and properties

Let ρ, σ two neighbouring quantum states, i.e. $\rho \overset{Q}{\sim} \sigma$. We'll discuss appropriate neighbouring conditions for quantum states in the next sections and for the moment we use the letter Q as a placeholder. We also say that ρ and σ are Q-neighbouring in order to emphasise that we selected a suitable relationship Q over quantum states. Following [ZY17b, HRF23], we say that a quantum channel $\mathscr{C}(\cdot)$ is (ε,δ)-DP if for all $\rho \overset{Q}{\sim} \sigma$, for all POVM $M = \{M_m\}$ and for all m, we have that

$$\mathrm{Tr}[M_m \mathscr{C}(\rho)] \leq e^{\varepsilon} \mathrm{Tr}[M_m \mathscr{C}(\sigma)] + \delta.$$

As in the classical case, this can be equivalently expressed in terms of the quantum hockey-stick divergence or the quantum smooth max-relative entropy:

$$\mathscr{C} \text{ is } (\varepsilon,\delta)\text{-DP} \iff \forall \rho,\sigma : \rho \overset{Q}{\sim} \sigma : E_{e^{\varepsilon}}(\mathscr{C}(\rho) \| \mathscr{C}(\sigma)) \leq \delta \tag{7.23}$$

$$\iff \forall \rho,\sigma : \rho \overset{Q}{\sim} \sigma : D_{\infty}^{\delta}(\mathscr{C}(\rho) \| \mathscr{C}(\sigma)) \leq \varepsilon, \tag{7.24}$$

where the quantum hockey-stick divergence E_γ is defined as follows:

$$E_\gamma(\rho \| \sigma) := \mathrm{Tr}(\rho - \gamma\sigma)^+, \tag{7.25}$$

for $\gamma \geq 1$. Here X^+ denotes the positive part of the eigendecomposition of a Hermitian matrix $X = X^+ - X^-$. We refer to Lemma III.2 in [HRF23] for more details. A special case of particular interest is the one of quantum-to-classical channels (i.e. POVM measurements), mapping states

to probability distributions. For a measurement $\mathscr{M}$, denote as $\mathscr{M}(\rho)$ the probability distribution induced by measuring $\mathscr{M}$ on input ρ. Quantum differential privacy shares many useful properties with classical differential privacy. Notably, it is robust to parallel composition and post-processing (also referred to as sequential composition).

Proposition 7.1 (Adapted from Corollary III.3, [HRF23])**.** *The following properties hold.*

- *(Post-processing) Let $\mathscr{A}$ be (ε,δ)-differentially private and $\mathscr{N}$ be an arbitrary quantum channel, then $\mathscr{N}\circ\mathscr{A}$ is also (ε,δ)-differentially private.*
- *(Parallel composition) Let $\mathscr{A}_1$ be (ε_1,δ_1)-differentially private and $\mathscr{A}_2$ be (ε_2,δ)-differentially private. Define that $\rho_1\otimes\rho_2 \overset{Q}{\sim} \sigma_1\otimes\sigma_2$ if $\rho_1 \overset{Q}{\sim} \sigma_1$ and $\rho_2 \overset{Q}{\sim} \sigma_2$. Then $\mathscr{A}_1\otimes\mathscr{A}_2$ is $(\varepsilon_1+\varepsilon_2,\overline{\delta})$-differentially private on such product states, with $\overline{\delta}=\min\{\delta_1+e^{\varepsilon_1}\delta_1, e^{\varepsilon_2}\delta_1+\delta_2\}$.*

  Moreover, if $\mathscr{A}_1$ and $\mathscr{A}_2$ are quantum-classical channels (measurements), we have that $\mathscr{A}_1\otimes\mathscr{A}_2$ is $(\varepsilon_1+\varepsilon_2,\delta_1+\delta_2)$-differentially private.

Proof. The proposition coincides with Corollary III.3 in [HRF23], except for the final statement about the parallel composition of differentially private measurements. Since the output of a measurement is a classical distribution, the proof of this part is identical to the one of Theorem 3.16 in [DR14]. ■

In short, the composition theorem ensures that performing k times an ε-DP algorithm is (εk)-differentially private, and then the privacy budget scales as the number of repetitions k. However, under mild assumptions, this scaling can be improved to $O(\sqrt{k})$. This result is called *advanced composition* (we refer to Theorem 3.20 in [DR14] for the classical case). Moreover, advanced composition holds also for quantum measurements under suitable assumptions (Theorem 6, [ZY17b]).

Rényi quantum differential privacy has also been defined in [HRF23]. Due to the non-commutative nature of quantum mechanics, the quantum generalisation of the Rényi divergence is not unique. However, we don't need to fix a particular definition of the quantum Rényi divergence, since we can define Rényi quantum differential privacy in terms of an arbitrary family of Rényi divergences $\mathbb{D}_\alpha$, as defined in [Tom15]. Thus, a quantum channel $\mathscr{C}$ is (ε,α)-Rényi differentially private if

$$\sup_{\rho\sim\sigma}\mathbb{D}_\alpha(\mathscr{C}(\rho)\|\mathscr{C}(\sigma))\leq\varepsilon.$$

7.5 Relation with gentle measurements

The growing interest in quantum differential privacy is notably spurred by its remarkable connection with the concept of quantum gentle measurements, as discovered by Aaronson and Rothblum in their work [AR19]. In this section, we will delve into their findings.

At the heart of quantum theory lies the fundamental principle known as the "information-disturbance" tradeoff. It is a well-established principle that any measurement of an arbitrary quantum state necessarily perturbs or disturbs that state to some extent. However, exceptions exist when

dealing with specific state families, such as when we have a guarantee that a state, say $|\psi\rangle$, belongs to a particular set, like $|0\rangle, |1\rangle$. In such cases, a computational measurement can discriminate between the states without perturbing them.

In the realm of quantum measurements, one often encounters the terms "gentle" for measurements causing low disturbance and "weak" or "trivial" for those yielding minimal information. The literature offers various rigorous definitions for these concepts. In particular, Aaronson and Rothblum provide the following definition.

Definition 7.1 (Gentleness)**.** Given a set $S \subseteq \mathscr{S}_n$ of quantum mixed states and a parameter $\alpha \in [0,1]$, we say that an implementation of a measurement M is α-gentle on S if for all states $\rho \in S$, and all possible outcomes y of applying M to ρ, we have

$$\|\rho - \rho'\|_{\mathrm{tr}} \leq \alpha, \tag{7.26}$$

where ρ' is the post-measurement state.

The following theorem shows a connection between gentleness and quantum local differential privacy. We remark that in [AR19], local differentially private measurements are referred as trivial measurements.

Theorem 7.3 ([AR19], Lemmas 23, 26)**.** *Let $\varepsilon \leq 1$ and $\alpha \leq \frac{1}{4.01}$.*

1. *M satisfies ε-local differential privacy $\implies$ M is $O(\varepsilon)$-gentle on all states.*
2. *M is α-gentle on product states $\implies$ M satisfies $O(\alpha)$-local differential privacy on product states.*

Moreover, we also review the following connection between gentle measurements and quantum differential privacy beyond the local model. It's important to note that, in contrast to the local differential privacy scenario, this connection is applicable exclusively to product states. In this particular context, the concept of neighboring quantum states is based on the convertibility via local operations.

Theorem 7.4 ([AR19], Lemmas 28, 32)**.** *Let $\rho \sim \sigma$ if there exists $i \in [n]$ such that $\mathrm{Tr}_i \rho = \mathrm{Tr}_i \sigma$. Let $\varepsilon \leq 1$ and $\alpha \leq \frac{1}{4.01}$.*

1. *M is a product measurement and it satisfies ε-differential privacy on product states $\implies$ M is $O(\varepsilon\sqrt{n})$-gentle on product states*
2. *M is a α-gentle on all states $\implies$ M satisfies $O(\alpha)$- differential privacy.*

Building upon this profound connection, Aaronson and Rothblum introduced a novel algorithm for shadow tomography of quantum states. We recall that the task of shadow tomography consists in the estimation of the expected outcomes of a predetermined set of measurements on an unknown quantum state. Similar ideas based on gentleness were also employed in subsquent works on learning quantum states and channels [BO21, FQR22].

7.6 From quantum to classical differential privacy

We will now elucidate how quantum differential privacy can serve as a proxy for safeguarding the privacy of classical data when it is encoded within a quantum state. To begin, we present an initial definition.

Definition 7.2 (Privacy-preserving quantum encodings)**.** Let $\mathscr{X}$ a set equipped with a neighboring relationship $\sim$. A quantum encoding $\rho(\cdot)$ is Q-neighboring-preserving if

$$x \sim x' \implies \rho(x) \overset{Q}{\sim} \rho(x').$$

The following proposition formalizes the intuitive fact that Q-neighboring-preserving encodings can be used to transfer privacy guarantees and ensure the privacy of the underlying classical input.

Proposition 7.2 (Transferring privacy guarantees)**.** *Let $\rho(\cdot)$ a quantum encoding, i.e. a function mapping a classical vector $x \in \mathscr{X}$ to a quantum state $\rho(x)$. Assume $\mathscr{X}$ is equipped with a neighboring relationship $\sim$ and $\mathscr{S}_n$ is equipped with a neighboring relationship $\overset{Q}{\sim}$. Assume that $\rho(\cdot)$ is Q-neighboring-preserving. Let $\mathscr{M}$ be a measurement. We have,*

$$\mathscr{M} \text{ is } (\varepsilon,\delta)\text{-DP with respect to } \overset{Q}{\sim} \implies \mathscr{M}(\rho(\cdot)) \text{ is } (\varepsilon,\delta)\text{-DP with respect to } \sim.$$

Proof. The proposition follows from the definition of differential privacy. Assuming $\mathscr{M}(\cdot)$ is (ε,δ)-DP, we have

$$\forall \sigma,\sigma' : \sigma \overset{Q}{\sim} \sigma',\ \forall S \subseteq \mathsf{range}(\mathscr{M}) : \Pr[\mathscr{M}(\sigma) \in S] \leq e^{\varepsilon} \Pr[\mathscr{M}(\sigma') \in S] + \delta.$$

Since $\rho(\cdot)$ is Q-neighboring-preserving, the above inequality still holds if we set $\sigma := \rho(x)$ and $\sigma' := \rho(x')$ for $x \sim x'$. Moreover, we replace $\mathsf{range}(\mathscr{M})$ with $\mathsf{range}(\mathscr{M} \circ \rho(\cdot))$ (we can do it since $\mathsf{range}(\mathscr{M} \circ \rho(\cdot))$ is a subset of $\mathsf{range}(\mathscr{M})$). The result readily follows.

$$\forall x, x' : x \sim x',\ \forall S \subseteq \mathsf{range}(\mathscr{M} \circ \rho(\cdot)) : \Pr[\mathscr{M}(\rho(x)) \in S] \leq e^{\varepsilon} \Pr[\mathscr{M}(\rho(x')) \in S] + \delta.$$

■

The above theorem suggests a path for deploying quantum private algorithms on classical data. Notably, building upon this result, the authors of [YLLP23] explored the level of DP ensured by the global depolarizing channel for several noise magnitudes and multiple quantum encodings.

7.7 Certified adversarial robustness

Now, we outline the connection between differential privacy and adversarial robustness, which has been previously established in [LAG+19] and extended to the quantum setting in [DHL+21b, Hir23, HTY+23]. We consider a slightly different setting, known as *k-class classification*, where a classification algorithm $\mathscr{A}$ outputs a label $y \in [k]$ on input $\boldsymbol{x}$. For instance, for $k = 2$, we can consider an algorithm that outputs label 1 if $\boldsymbol{x}$ represents a dog and 2 if $\boldsymbol{x}$ represents a cat. Consider k observables $O_1, \ldots, O_k$, and assume, for simplicity, that their spectrum lies in $[0,1]$. The algorithm $\mathscr{A}$ works as follows.

1. On input $\boldsymbol{x}$, for each $i \in [k]$, the algorithm measures the observable O_i on the state $\rho(\boldsymbol{x},\boldsymbol{\theta})$ m times and stores the outcomes in $y_1^{(i)},\dots,y_m^{(i)}$.
2. For each $i \in [k]$, let $y^{(i)} = \sum_{j=1}^m y_j^{(i)}$.
3. $\mathscr{A}$ returns the index $i^* \in [k]$ such that $i^* = \arg\max y^{(i)}$.

We adopt Proposition 1 in [LAG+19] to the quantum setting.

Proposition 7.3 (Robustness condition)**.** *Let $\beta \in (0,1]$. Let $\rho(\cdot,\boldsymbol{\theta})$ be Q-neighbouring-preserving and assume that each of the m measurements in step (1) satisfies (ε,δ)-DP with respect to Q-neighbouring quantum states. For any input $\boldsymbol{x}$, if for some $i \in [k]$,*

$$y^{(i)} > e^{2\varepsilon} \max_{j\neq i} y^{(j)} + (1+e^{\varepsilon})\delta + \sqrt{\frac{2}{m}\log\left(\frac{4k}{\beta}\right)}, \tag{7.27}$$

then the algorithm $\mathscr{A}$ satisfies, for all $\boldsymbol{x} \sim \boldsymbol{x'}$

$$\Pr[\mathscr{A}(\boldsymbol{x}) = \mathscr{A}(\boldsymbol{x'})] \geq 1-\beta.$$

In this case, we say that the classifier $\mathscr{A}$ is β-robust to adversarial attacks.

Proof. Let $\boldsymbol{x} \sim \boldsymbol{x'}$. Since $\rho(\cdot,\boldsymbol{\theta})$ is Q-neighbouring-preserving, $\rho(\boldsymbol{x},\boldsymbol{\theta}) \overset{Q}{\sim} \rho(\boldsymbol{x'},\boldsymbol{\theta})$. The assumption that each measurement satisfies (ε,δ)-DP implies

$$\forall i \in [k], \forall F \subseteq \mathsf{range}(O_i) : \Pr_{\rho(\boldsymbol{x},\boldsymbol{\theta})}[O_i \in F] \leq e^{\varepsilon} \Pr_{\rho(\boldsymbol{x'},\boldsymbol{\theta})}[O_i \in F] + \delta.$$

We first need to prove the following inequality. For all i,

$$\mathrm{Tr}[O_i\rho(\boldsymbol{x},\boldsymbol{\theta})] \leq e^{\varepsilon}\,\mathrm{Tr}[O_i\rho(\boldsymbol{x'},\boldsymbol{\theta})] + \delta. \tag{7.28}$$

Recall that the expectation of a non-negative random variable X can be expressed as

$$\mathbb{E}(X) = \int_{t\geq 0} \Pr[X > t]\,dt.$$

Combining this with differential privacy, we obtain

$$\mathrm{Tr}[O_i\rho(\boldsymbol{x},\boldsymbol{\theta})] = \int_{t\geq 0} \Pr_{\rho(\boldsymbol{x},\boldsymbol{\theta})}[O_i > t]\,dt$$
$$\leq e^{\varepsilon} \int_{t\geq 0} \Pr_{\rho(\boldsymbol{x'},\boldsymbol{\theta})}[O_i > t]\,dt + \delta = e^{\varepsilon}\,\mathrm{Tr}[O_i\rho(\boldsymbol{x},\boldsymbol{\theta})] + \delta,$$

which proves (7.28). It remains to show that the discrepancy between $y^{(i)} = \frac{1}{m}\sum_{j=1}^m y_j^{(i)}$ and of $\mathrm{Tr}[O_i\rho(\boldsymbol{x},\boldsymbol{\theta})]$ is small enough with high probability. To this end, we can use concentration of measure. By Chernoff-Hoeffding's bound,

$$\Pr\left[\left|\frac{1}{m}\sum_{j=1}^m y_j^{(i)} - \mathrm{Tr}[O_i\rho(\boldsymbol{x},\boldsymbol{\theta})]\right| \geq t\right] \leq 2e^{-2mt^2}.$$

and thus $y^{(i)} = \mathrm{Tr}[O_i\rho(\boldsymbol{x},\boldsymbol{\theta})] \pm t$ with probability at least $1-2e^{-2mt^2}$. Denote by $\widetilde{y}^{(1)},\ldots,\widetilde{y}^{(k)}$ the average of the measurements on the state $\rho(\boldsymbol{x}',\boldsymbol{\theta})$. By union bound, with probability at least $1-4ke^{-2mt^2} = 1-\beta$ we have that

$$\forall i \in [k] : \Big(y^{(i)} = \mathrm{Tr}[O_i\rho(\boldsymbol{x},\boldsymbol{\theta})] \pm t\Big) \wedge \Big(\widetilde{y}^{(i)} = \mathrm{Tr}[O_i\rho(\boldsymbol{x}',\boldsymbol{\theta})] \pm t\Big). \tag{7.29}$$

Assume, by contradiction, that $\mathscr{A}(x) \neq \mathscr{A}(x')$ and (7.29) hold simultaneously Since $\mathscr{A}(x) \neq \mathscr{A}(x')$, there exists $i \neq i'$ such that

$$y^{(i)} > \max_{j\neq i} y^{(j)} \text{ and } \widetilde{y}^{(i')} > \max_{j\neq i} \widetilde{y}^{(j)}.$$

Putting them all together, we have

$$\begin{aligned}\widetilde{y}^{(i)} &\geq \mathrm{Tr}[O_i\rho(\boldsymbol{x}',\boldsymbol{\theta})] - t \geq e^{-\varepsilon}(\mathrm{Tr}[O_i\rho(\boldsymbol{x},\boldsymbol{\theta})] - t) - e^{-\varepsilon}\delta \\ &\geq e^{-\varepsilon}(y^{(i)} - 2t) - e^{-\varepsilon}\delta > \max_{j\neq i} e^{\varepsilon} y^{(j)} + \delta \\ &\geq \max_{j\neq i} \widetilde{y}^{(j)} \geq \widetilde{y}^{(i')}\end{aligned}$$

Thus we obtained $\widetilde{y}^{(i)} > \widetilde{y}^{(i')}$ contradicting the assumptions $\mathscr{A}(x) \neq \mathscr{A}(x')$. This proves that $\mathscr{A}(x) = \mathscr{A}(x')$ with probability at least $1-\beta$. ■

It's easy to see how the above proposition is related to adversarial attacks. Assume that an adversary has the capabilities of tampering with the input $\boldsymbol{x}$ by replacing it with $\boldsymbol{x}'$ such that $\boldsymbol{x} \sim \boldsymbol{x}'$. We remark that there's no unique way of choosing the neighbouring relationship in this context, as it is closely related to the capabilities of the adversary. Under the same assumptions of Proposition 7.3, the adversarial attack doesn't alter the output with high probability. The condition expressed in (7.27) can be interpreted as the classifier being "fairly confident" about its prediction. We also remark that Proposition 7.3 can be applied to virtually any algorithm $\mathscr{A}$, even in the absence of an explicit private mechanism, since all algorithms are by default $(0,\tau)$-DP with respect to neighbouring states with trace distance bounded by τ. This can be easily checked from the properties of the trace distance.

Following [LAG+19], given a distribution $\mathscr{D}$ over labeled inputs of the form $(\boldsymbol{x}, f(\boldsymbol{x}))$, we can define the *certified accuracy* $\mathscr{R}(\mathscr{A})$ of an (ε,δ)-DP algorithm $\mathscr{A}$ as follows

$$\mathscr{R}(\mathscr{A}) := \Pr_{(\boldsymbol{x},f(\boldsymbol{x}))\sim\mathscr{D}}\left[\Big(i^* = f(\boldsymbol{x})\Big) \wedge \left(\delta < \frac{y^{(i^*)} - e^{2\varepsilon}\max_{j\neq i^*} y^{(j)} - g(k,\beta,m)}{1+e^{\varepsilon}}\right)\right],$$

where $g(k,\beta,m) := \sqrt{2m^{-1}\log\big(4k/\beta\big)}$ and $i^* = \arg\max y^{(i)}$. In other terms, $\mathscr{R}$ is a lower bound on the probability that an instance is classified correctly and the classification is β-robust to adversarial attacks. We remark that $\mathscr{R}$ can be easily estimated by computing the fraction of the test set that is classified correctly and, simultaneously, satisfies (7.27).

7.8 Generalization

We conclude by recalling the connection between differential privacy and generalization. Given a randomised algorithm $M : \mathscr{X}^m \times \mathscr{X} \to [0, B]$ and two datasets $S, S' \in \mathscr{X}^m$ we define the following quantity:

$$\mathscr{E}_S[M(S)] := \frac{1}{m}\sum_{z\in S}\mathbb{E}_M[M(S,z)], \quad \mathscr{E}_{S'}[M(S)] := \frac{1}{m}\sum_{z'\in S'}\mathbb{E}_M[M(S,z')].$$

Lemma 7.1 (Lemma 6.4, [FS17])**.** *Let $S \in \mathscr{X}^m$ and $\boldsymbol{x} \in \mathscr{X}$. Let M be an algorithm that on input $(S, \boldsymbol{x})$ outputs a value $y \in [0, B]$. Assume that M is (ε, δ)-differentially private with respect to S, where $S \sim S'$ if they differ in at most one entry. Let $\mathscr{P}$ be an arbitrary distribution over $\mathscr{X}$. Then:*

$$\mathbb{E}_{S,S'\sim\mathscr{P}^m}[(\mathscr{E}_{S'}[M(S)])^k] \le e^{k^2\varepsilon}\mathbb{E}_{S\sim\mathscr{P}^m}[(\mathscr{E}_{S'}[M(S)] + k\delta B)^k].$$

We also define $\mathscr{E}_{\mathscr{P}}[M(S)] := \mathbb{E}_{z\sim\mathscr{P},\,M}[M(S,z)]$. Clearly,

$$\mathbb{E}_{S'\sim\mathscr{P}^n}[\mathscr{E}_{S'}[M(S)]] = \mathscr{E}_{\mathscr{P}}[M(S)].$$

Moreover, as noted in [FS17], standard concentration inequalities implies that $\mathscr{E}_{S'}[M(S)]$ is strongly concentrated around $\mathscr{E}_{\mathscr{P}}[M(S)]$. Note that for $M(S,(x,y)) = \ell(M'(S,x),y)$, $\mathscr{E}_S[M(S)] = \mathscr{E}_S[\ell(M'(S))]$ and $\mathscr{E}_{\mathscr{P}}[M(S)] = \mathscr{E}_{\mathscr{P}}[\ell(M'(S))]$, in other words these are exactly the empirical and the expected loss of the predictor given by M'.

CHAPTER 8

A UNIFYING FRAMEWORK FOR QUANTUM DIFFERENTIAL PRIVACY

Poetry is the synthesis of hyacinths and biscuits.

- Carl Sandburg

In this Chapter, we propose a novel and general definition of neighboring quantum states. We demonstrate that this definition captures the underlying structure of quantum encodings and can be used to provide exponentially tighter privacy guarantees for quantum measurements. Our approach combines the addition of classical and quantum noise and is motivated by the noisy nature of near-term quantum devices.

Moreover, we also investigate an alternative setting where we are provided with multiple copies of the input state. In this case, differential privacy can be ensured with little loss in accuracy combining concentration of measure and noise-adding mechanisms. Finally, we complement our theoretical findings with an empirical estimation of the certified adversarial robustness ensured by differentially private measurements.

8.1 Motivation: connecting neighboring relationships with quantum encodings

Our motivation stems from a practical goal: the development of quantum algorithms satisfying differential privacy with respect to a classical input x. We assume that this input belongs to a set equipped with a neighboring relationship. Furthermore, we focus on quantum algorithms that incorporate a quantum encoding subroutine, wherein the classical input x undergoes a transformation into a quantum state denoted as $\rho(x)$.

As a result, we aim to establish a quantum neighboring relationship that mirrors the inherent classical neighboring relationship. Specifically, we demand the following key property:

$$x \text{ and } x' \text{ are neighboring} \implies \rho(x) \text{ and } \rho(x') \text{ are neighboring.} \tag{8.1}$$

It is easy to see why the above property is extremely useful. If an algorithm $\mathscr{A}$ is ε-differentially private with respect to $\rho(x)$, then $\mathscr{A} \circ \rho$ is ε-differentially private with respect to x (this is stated more formally in Proposition 7.2). In the meantime, we want to avoid neighboring relationships that are excessively loose, as this would make the output almost independent of the input. Certainly, the simplistic notion of a universal neighboring relationship that designates all states as neighbors may technically meet the criteria defined in Equation 8.1. However, it would significantly compromise the overall accuracy, as the outputs of any pair of inputs ρ, σ would be be ε-close in quantum max-relative entropy. Another paradigmatic example of a "pathological" relationship is the one based on *constant* trace distance:

$$\rho \text{ and } \rho' \text{ are neighboring} \iff \left\|\rho - \rho'\right\|_{\mathrm{tr}} \leq \tau = \Theta(1). \tag{8.2}$$

To fix the ideas, let $\tau = 0.1$. It's easy to see that for any pair of states ρ, σ we can build a sequence $\rho_0, \rho_2, \ldots, \rho_{10}$, such that

$$\begin{cases} \rho_0 = \rho \\ \rho_{10} = \sigma \end{cases} \qquad \text{and for all } i,\ \rho_i \sim \rho_{i+1}. \tag{8.3}$$

In particular, it suffices to let:

$$\rho_i = \left(1 - \frac{i}{10}\right)\rho + \frac{i}{10}\sigma. \tag{8.4}$$

By triangle inequality, the outputs of ρ and σ will be (10ε)-close in quantum max-relative entropy. This severely curtails the potential of private algorithms. Irrespective of the input states, the resulting distribution of outputs would exhibit an extreme concentration around a single value. A more comprehensive examination of this issue is provided in Section 8.7.

Remarkably, existing quantum neighboring relationships satisfy these essential criteria for only a select few quantum encodings. This presents a significant limitation, as this property is fundamental to the framework of differential privacy, and the quantum domain involves a diverse array of quantum and classical data types. To address this challenge, it has become imperative to devise a more

inclusive approach capable of accommodating various encodings. Thus, we introduce a generalized neighboring relationship designed to facilitate the application of a broad spectrum of algorithms, spanning both near-term and long-term quantum computing contexts. Our findings shed new light on the following fundamental question.

Question 3. *Can we leverage quantum noise to guarantee properties like differential privacy and robustness to adversarial attacks?*

8.2 Overview of main results

Within the domain of quantum differential privacy, we tackle several technical problems, and our approach is to resolve them within a comprehensive framework, leveraging a diverse set of tools and techniques drawn from the field of quantum information. The following provides an overview of our principal contributions.

Improved privacy bounds for noisy channels. Our first contribution consists of tighter privacy guarantees for a general family of noisy channels, which includes local Pauli noise and particularly as a special case, the depolarising channel. To this end, we prove the advanced joint convexity of the quantum hockey-stick divergence. Moreover, we provide a tighter analysis of the privacy of quantum measurements post-processed with classical stochastic channels, such as the Laplace or Gaussian noise. This approach allows us to be able to study both classical and quantum noisy mechanisms for differential privacy, within a unified framework.

Generalized neighboring relationship. Our second contribution is a generalized neighboring relationship, that allows us to recover the previous definition as special cases. We demonstrate how to design differentially private measurements according to this definition by introducing both classical and quantum noise into the computation. Notably, we show that local measurements can be made differentially private by adding a modest amount of noise. Our work is the first to incorporate the locality and in the analysis of quantum differential privacy.

Privacy-utility tradeoff for quantum differential privacy. There exists an unavoidable tradeoff between the desired level of privacy and the resulting loss in accuracy. Here, we make a crucial observation: different neighboring relationships have different tradeoffs. In particular, this limits the applicability of neighboring relationships based solely on the bounded trace distance. We also show no-go results for pure quantum differential privacy under the Wasserstein distance of order 1.

Private estimation with multiple copies. We provide differentially private mechanisms for estimating the expected values of observables given m copies of a quantum state. These mechanisms can find applications in privatising the results of experiments on physical devices where estimating the expectation value is the main figure of merit.

Applications. Our results can be applied to variational quantum algorithms and other quantum machine learning models to enhance or certify privacy. We specifically focus on certified adversarial robustness through differential privacy and we perform numerical simulations to assess the robustness to adversarial attacks of private quantum classifiers.

8.3 Organization

This Chapter is organised as follows. In Section 8.4 we introduce our generalized framework for quantum differential privacy and we discuss its properties. Within this formal framework, we prove several results. Starting with Section 8.5, we provide several improved privacy bounds for the case where the neighboring relationship is specified with a bounded trace distance between quantum states. We then turn to the unique properties of our framework in Section 8.6 which allows us to study local measurements as quantum differentially private mechanisms, as well as addressing the question of how quantum and classical noise can be studied together in the context of differential privacy. In Section 8.7 we define the cost of differential privacy and benchmark different approaches and notions of neighboring, under this lens, providing negative and positive results which clarify and justify the applicability of our framework. In Section 8.8 we introduce mechanisms for privately estimating expectation values. Finally, in Section 8.9, we discuss applications of some of our results in quantum machine learning, particularly for certified adversarial robustness, and we support our theoretical findings with numerical simulations.

8.4 Generalized neighboring relationship

In this section, we present the cornerstone of our work, which is a general definition of neighboring quantum states.

Definition 8.1. Let $\rho,\sigma \in \mathscr{S}_n$ and let $\Xi \subset P([n])$, i.e. let Ξ be a collection of subsets of $[n]$. Let $\tau > 0$ be a parameter. We say that ρ and σ are (Ξ,τ)-neighboring and we write $\rho \overset{(\Xi,\tau)}{\sim} \sigma$ if

$$\exists \mathscr{I} \in \Xi : \mathrm{Tr}_{\mathscr{I}} \rho = \mathrm{Tr}_{\mathscr{I}} \sigma \ \wedge \ \|\rho - \sigma\|_{\mathrm{tr}} \leq \tau. \tag{8.5}$$

If $\Xi = \{\mathscr{I} : \mathscr{I} = \{i, i+1, \ldots, i+\ell\} \text{ for some } i\}$, i.e. each subset $\mathscr{I}$ is a collection of ℓ consecutive integers (modulo n), we say that ρ and σ are (ℓ,τ)-neighboring and we write $\rho \overset{(\ell,\tau)}{\sim} \sigma$. When $\Xi = \{[n]\}$, we simply write $\rho \overset{\tau}{\sim} \sigma$ and we say that ρ and σ are τ-neighboring.

This definition of neighboring states extends those used in previous works. In [ZY17b, DHL$^+$21b, HRF23], two states are neighboring if they have bounded trace distance τ, i.e. if they are τ-neighboring. Moreover, setting $\ell = 1$ and $\tau = 1$ we recover the definition of quantum differential privacy based on convertibility by local measurements, used in [AR19].

This notion is particularly suitable to handle local measurements, i.e. measurements expressible as sums of local terms, as we show in Section 8.6. We remark that local measurements are of particular

Table 8.1: As we discuss in details in Appendix 11.2, the encodings above are (Ξ,τ)-neighboring-preserving for appropriate Ξ and τ depending on the encodings. We assume that the initial vectors $\boldsymbol{x}$ and $\boldsymbol{x'}$ are neighboring if $\|\boldsymbol{x}-\boldsymbol{x'}\|_0 \le \gamma_0$, $\|\boldsymbol{x}-\boldsymbol{x'}\|_1 \le \gamma_1$ and $\|\boldsymbol{x}-\boldsymbol{x'}\|_2 \le \gamma_2$. We also assumed that the Hamiltonian encoding is implemented by a $1D$ circuit of depth at most L.

ENCODING $\rho(\cdot)$	$\max_{\mathscr{I}\in\Xi}\lvert\mathscr{I}\rvert$	τ
AMPLITUDE ENCODING	n	γ_2
ROTATION ENCODING	γ_0	1
1D-HAMILTONIAN ENCODING	$2L\gamma_0$	$O(1)\gamma_1$
1D-HAMILTONIAN ENCODING (LOW NOISE)	$2L\gamma_0$	$O(1)\sqrt{n}\exp(-L)$
1D-HAMILTONIAN ENCODING (HIGH NOISE)	$2L\gamma_0$	$O(1)\exp(-L)\gamma_1$

interest since they can be considered practically feasible measurements for extracting classical information from quantum data (or quantum systems) [HKP20]. They also play a major role in variational learning algorithms as they are provably resilient to barren plateaus [CSV$^+$21].

On the other hand, several encodings widely used in quantum machine learning are (Ξ,τ)-neighboring-preserving, for appropriate choices of (Ξ,τ). We include upper bounds for $\max_{\mathscr{I}\in\Xi}\lvert\mathscr{I}\rvert$ and τ in Table 8.1. We delay to Section 11.2 the definition of the various encodings and the proof of upper bounds.

We also show that the notion of (Ξ,τ)-neighboring states degrades gently under quantum post-processing, assuming that the post-processing channel has a bounded light-cone, as defined below. Given a quantum channel Φ acting on n qubits, we define its light-cone as follows: first, for any qubit i, we denote by $\mathscr{I}_i$ the minimal subset of qubits such that $\mathrm{Tr}_{\mathscr{I}_i}\Phi(\rho)=\mathrm{Tr}_{\mathscr{I}_i}\Phi(\sigma)$ for any two n-qubit states ρ and σ such that $\mathrm{Tr}_i(\rho)=\mathrm{Tr}_i(\sigma)$. Then, the light-cone of Φ is defined as

$$\lvert\mathscr{I}\rvert := \max_{i\in[k]}\lvert\mathscr{I}_i\rvert. \tag{8.6}$$

Proposition 8.1 (Robustness to quantum post-processing)**.** *Let ρ and σ be two (Ξ,τ)-neighboring states and consider a channel Φ with light-cone bounded by K. Then $\Phi(\rho)$ and $\Phi(\sigma)$ are (Ξ',τ)-neighboring, where*

$$\max_{\mathscr{I}\in\Xi'}\lvert\mathscr{I}\rvert \le K\max_{\mathscr{I}\in\Xi}\lvert\mathscr{I}\rvert. \tag{8.7}$$

Proof. The proposition follows from the fact that the trace distance is non-increasing and from the definition of light-cone provided above. We have

$$\frac{1}{2}\|\Phi(\rho)-\Phi(\sigma)\|_1 \le \frac{1}{2}\|\rho-\sigma\|_1 \le \tau \tag{8.8}$$

Moreover,

$$\mathrm{Tr}_{\mathscr{J}}\rho = \mathrm{Tr}_{\mathscr{J}}\sigma \tag{8.9}$$

for $\mathscr{J}\in\Xi$. Since the channel Φ has bounded light-cone K, there exists $\mathscr{J}'\subseteq[n]$

$$\mathrm{Tr}_{\mathscr{J}'}\rho = \mathrm{Tr}_{\mathscr{J}'}\sigma \tag{8.10}$$

where $|\mathscr{J}'| \leq K|\mathscr{J}|$. This implies the desired result. ■

We conclude this section by observing that our definition can be easily related to the quantum Wasserstein distance of order 1. Combining Lemma 3.3 and Eq. 3.40, we obtain

$$\rho \overset{(\Xi,\tau)}{\sim} \sigma \implies W_1(\rho,\sigma) \leq \min\left\{\max_{\mathscr{J}\in\Xi}|\mathscr{J}|\frac{3}{2}\tau, n\tau\right\}. \tag{8.11}$$

It's natural to ask whether it would be convenient to define neighboring quantum states in terms of the W_1 distance. The answer to this question is twofold. On the one hand, when we dispose of a single copy of the input state, the W_1 distance leads to a suboptimal tradeoff between privacy and accuracy, as we show in Theorem 8.7. On the other hand, when we dispose of multiple copies of the input state, neighboring quantum states can be suitably defined in terms of the W_1 distance. We will discuss this alternative setting in Section 8.8.

8.5 Improved privacy for states with bounded trace distance

Before dealing with the general case of (Ξ,τ)-neighbouring states, we provide several new results for τ-neighbouring states, i.e. states with trace distance bounded by τ. This corresponds to the definition previously explored in [ZY17b, HRF23]. In particular, we provide tighter guarantees for two private mechanisms, namely a generalized noisy channel and the addition of classical noise on the output of a quantum measurement. Following the convention used in [HRF23], we state the results of this section using the quantum hockey-stick divergence.

Lemma 8.1. *Let $\mathscr{N}_p(\cdot) = p\frac{I}{2^n} + (1-p)\mathscr{M}(\cdot)$ a channel. For $0 \leq p \leq 1$ and $\gamma \geq 1$ we have*

$$E_{\gamma'}(\mathscr{N}_p(\rho)\|\mathscr{N}_p(\sigma)) \leq (1-p)(1-\beta)E_\gamma(\rho\|I/2^n) + (1-p)\beta E_\gamma(\rho\|\sigma), \tag{8.12}$$

where $\gamma' = 1+(1-p)(\gamma-1)$ and $\beta = \gamma'/\gamma$.

Proof. The result follows from Lemma 11.2 by plugging $\rho_0 = I/2^n$, $\rho_1 = \rho$ and $\rho_2 = \sigma$. ■

Recall that from Lemma IV.1 in [HRF23] we have that for the depolarising noise (hence for $\mathscr{M} = \mathrm{Id}$) and for any $\gamma \geq 1$,

$$E_\gamma(\mathscr{N}_p(\rho)\|\mathscr{N}_p(\sigma)) \leq \max\left\{0, (1-\gamma)\frac{p}{2^n} + (1-p)E_\gamma(\rho\|\sigma)\right\}. \tag{8.13}$$

In the following theorem, we extend this previous bound to an arbitrary channel $\mathscr{M}$ and we combine it with Lemma 8.1.

Theorem 8.1. *Let $\mathscr{N}_p(\cdot) = p\frac{I}{2^n} + (1-p)\mathscr{M}(\cdot)$ a channel. For $0 \leq p \leq 1$ and $\gamma' \geq 1$ we have*

$$E_{\gamma'}(\mathscr{N}_p(\rho)\|\mathscr{N}_p(\sigma)) \leq \tag{8.14}$$

$$\min\left\{(1-p)(1-\beta)E_\gamma(\rho\|I/2^n) + (1-p)\beta E_\gamma(\rho\|\sigma), \max\left\{0, (1-\gamma')\frac{p}{2^n} + (1-p)E_{\gamma'}(\rho\|\sigma)\right\}\right\}. \tag{8.15}$$

where $\gamma = 1+(\gamma'-1)/(1-p)$ and $\beta = \gamma'/\gamma$.

Proof. Lemma 8.1 implies that

$$E_{\gamma'}(\mathcal{N}_p(\rho)\|\mathcal{N}_p(\sigma)) \leq (1-p)(1-\beta)E_\gamma(\rho\|I/2^n) + (1-p)\beta E_\gamma(\rho\|\sigma), \tag{8.16}$$

Then it remains to show that

$$E_{\gamma'}(\mathcal{N}_p(\rho)\|\mathcal{N}_p(\sigma)) \leq \max\left\{0, (1-\gamma')\frac{p}{2^n} + (1-p)E_{\gamma'}(\rho\|\sigma)\right\}. \tag{8.17}$$

The proof closely follows the one of Lemma IV.1 and Lemma IV.4 in [HRF23]. We have

$$E_{\gamma'}(\mathcal{N}_p(\rho)\|\mathcal{N}_p(\sigma)) \tag{8.18}$$

$$= \mathrm{Tr}((1-\gamma')p\frac{I}{2^n} + (1-p)\mathcal{M}((\rho-\gamma'\sigma)))^+ \tag{8.19}$$

$$= \mathrm{Tr}\,P^+((1-\gamma')p\frac{I}{2^n} + (1-p)\mathcal{M}((\rho-\gamma'\sigma))), \tag{8.20}$$

where P^+ is the projector onto the positive subspace of $((1-\gamma')p\frac{I}{2^n} + (1-p)\mathcal{M}((\rho-\gamma'\sigma)))$. Observe that

$$E_{\gamma'}(\mathcal{N}_p(\rho)\|\mathcal{N}_p(\sigma)) > 0 \quad \Rightarrow \quad \mathrm{Tr}\,P^+ \geq 1. \tag{8.21}$$

Considering this case we get

$$E_{\gamma'}(\mathcal{N}_p(\rho)\|\mathcal{N}_p(\sigma)) \tag{8.22}$$

$$= (1-\gamma')\frac{p}{2^n}\mathrm{Tr}\,P^+ + (1-p)(\mathrm{Tr}\,P^+(\mathcal{M}(\rho-\gamma'\sigma))) \tag{8.23}$$

$$\leq (1-\gamma')\frac{p}{2^n} + (1-p)E_{\gamma'}(\mathcal{M}(\rho)\|\mathcal{M}(\sigma)) \tag{8.24}$$

$$\leq (1-\gamma')\frac{p}{2^n} + (1-p)E_{\gamma'}(\rho\|\sigma) \tag{8.25}$$

$$\leq (1-\gamma')\frac{p}{2^n} + (1-p). \tag{8.26}$$

Note that for sufficiently large γ' the upper bound could become negative, but one can easily check that in this case $E_{\gamma'}(\mathcal{N}_p(\rho)\|\mathcal{N}_p(\sigma)) = 0$ implying that we are in the other case.

■

For single-qubit product channels, we give the following bound:

Theorem 8.2. *Let $\mathcal{N}_p(\cdot) = p\frac{I}{2} + (1-p)\mathcal{M}(\cdot)$ a single-qubit channel. For $0 \leq p \leq 1$ and $\gamma' \geq 1$ we have*

$$E_{\gamma'}(\mathcal{N}_p^{\otimes k}(\rho)\|\mathcal{N}_p^{\otimes k}(\sigma)) \leq \tag{8.27}$$

$$\min\left\{(1-p^k)(1-\beta)E_\gamma(\rho\|I/2^k) + (1-p^k)\beta E_\gamma(\rho\|\sigma), \max\left\{0, (1-\gamma')\frac{p^k}{2^k} + (1-p^k)E_{\gamma'}(\rho\|\sigma)\right\}\right\}. \tag{8.28}$$

where $\gamma = 1 + (\gamma'-1)/(1-p)$ and $\beta = \gamma'/\gamma$.

Proof. It suffices to note that $\mathcal{N}_p^{\otimes k}$ can be rearranged as:

$$\mathcal{N}_p^{\otimes k}(\cdot) = p^k \frac{I}{2^k} + (1-p^k)\mathcal{M}'(\cdot), \tag{8.29}$$

where $\mathcal{M}'$ is a quantum channel. Then the result follows from Theorem 8.1. ■

These first two technical results show that several quantum noisy channels contract the quantum hockey-stick divergence. This can be used to prove that those channels ensure quantum differential privacy for τ-neighbouring states. In particular, we derive the following corollaries, that improve Lemma IV.2 and Lemma IV.5 in [HRF23].

Corollary 8.1. *Let $\mathcal{N}_p(\cdot) = p\frac{I}{2^n} + (1-p)\mathcal{M}(\cdot)$ a channel. $\mathcal{N}_p$ is (ε,δ)-DP with respect to τ-neighbouring states with*

$$\delta \leq \max\left\{0, (1-e^{\varepsilon})\frac{p}{2^n} + (1-p)\tau\right\}. \tag{8.30}$$

Let $\gamma = 1+(e^{\varepsilon}-1)/(1-p)$ and $\beta = e^{\varepsilon}/\gamma$. Under the additional assumption that the input state ρ satisfies $E_\gamma(\rho\|\frac{\rho}{I/2^n}) \leq \eta$, we also have

$$\delta \leq (1-p)(1-\beta)\eta + (1-p)\beta\tau. \tag{8.31}$$

Determining whether Equation 8.31 yields a clear advantage over Equation 8.30 is not a straightforward task. To shed light on this matter, we have plotted Figure 8.1, in which both bounds for δ are graphically represented as functions of ε, considering a specific set of parameters. Our observation reveals that neither bound consistently outperforms the other. Therefore, the selection of the appropriate bound will depend on the specific value of ε. An upper bound of δ as a function of ε is also referred to as *privacy profile*, a concept introduced in [BBG18].

Corollary 8.2. *Let $\mathcal{N}_p(\cdot) = p\frac{I}{2} + (1-p)\mathcal{M}(\cdot)$ single-qubit a channel. $\mathcal{N}_p^{\otimes k}$ is (ε,δ)-DP with respect to τ-neighbouring states with*

$$\delta \leq \max\left\{0, (1-e^{\varepsilon})\frac{p^k}{2^k} + (1-p^k)\tau\right\}. \tag{8.32}$$

Let $\gamma = 1+(e^{\varepsilon}-1)/(1-p^k)$ and $\beta = e^{\varepsilon}/\gamma$. Under the additional assumption that the input state ρ satisfies $E_\gamma\left(\rho\middle\|\frac{I}{2^k}\right) \leq \eta$, we also have

$$\delta \leq (1-p^k)(1-\beta)\eta + (1-p^k)\beta\tau. \tag{8.33}$$

Bounding privacy with the purity. Our results improve the prior bounds under the additional assumption that the divergence $E_\gamma\left(\rho\|I/2^n\right)$ is relatively small. The value of $E_\gamma\left(\rho\|I/2^n\right)$ can be thought as a "distance" between the state ρ and the maximally mixed state, thus small values of $E_\gamma\left(\rho\|I/2^n\right)$ are associated to high levels of noise. Hence, we can connect it to the purity $\mathrm{Tr}\left[\rho^2\right]$ of the state ρ, or the related D_2 divergence. By definition, we have

$$\mathrm{Tr}\left[\rho^2\right] = 2^{-n+D_2(\rho\|I/2^n)}. \tag{8.34}$$

The hockey stick divergence and the Rényi divergence satisfy the following relationship ([Tom15], Proposition 6.22)

$$E_{e^\varepsilon}\left(\rho \| I/2^n\right) \leq \delta, \tag{8.35}$$

where $\varepsilon = D_2\left(\rho \| I/2^n\right) - \log\left(1-\sqrt{1-\delta^2}\right) \leq D_2\left(\rho \| I/2^n\right) + \log\left(2/\delta^2\right)$. We also note that two states with low purity are also close in hockey-stick divergence:

$$E_\gamma(\rho\|\sigma) \leq E_1(\rho\|\sigma) \leq E_1\left(\rho \| I/2^n\right) + E_1\left(\sigma \| I/2^n\right). \tag{8.36}$$

And then $E_1\left(\rho \| I/2^n\right) = \frac{1}{2}\left\|\rho - I/2^n\right\|_1$ can be bounded either with the quantum Bretagnolle Huber inequality (Lemma 11.1) or the Pinsker's inequality. Now, we show how Corollary 8.1 and Corollary 8.2 can be rephrased in terms of the purity of the input state.

Corollary 8.3. *Let $\mathcal{N}_p(\cdot) = p\frac{I}{2^n} + (1-p)\mathcal{M}(\cdot)$ a channel that acts on state ρ with bounded purity $\mathrm{Tr}[\rho^2] \leq \zeta < 1$. Let $\gamma = 1 + (e^\varepsilon - 1)/(1-p)$, $\beta = e^\varepsilon/\gamma$ and $\eta = \sqrt{2n\zeta^{\frac{1}{\log 2}}\gamma^{-1}}$. Then $\mathcal{N}_p$ is (ε,δ)-DP with respect to τ-neighbouring states with*

$$\delta \leq (1-p)(1-\beta)\eta + (1-p)\beta\tau. \tag{8.37}$$

Proof. The proof follows by plugging the relation between purity and hockey-stick divergence into Corollary 8.1. We have

$$D_2(\rho\|I/2^n) \leq \log_2(\zeta) + n, \tag{8.38}$$

and hence, by Equation 8.35,

$$E_\gamma(\rho\|I/2^n) \leq \sqrt{2n\zeta^{\frac{1}{\log 2}}\gamma^{-1}} := \eta, \tag{8.39}$$

which satisfies the hypothesis of Corollary 8.1. ■

Proceeding in a similar way can also prove a purity-based bound for local channels.

Corollary 8.4. *Let $\mathcal{N}_p(\cdot) = p\frac{I}{2^n} + (1-p)\mathcal{M}(\cdot)$ a single-qubit channel and assume that $\mathcal{N}_p^{\otimes k}$ acts on state ρ with bounded purity $\mathrm{Tr}[\rho^2] \leq \zeta < 1$. Let $\gamma = 1 + (e^\varepsilon - 1)/(1-p^k)$, $\beta = e^\varepsilon/\gamma$ and $\eta = \sqrt{2n\zeta^{\frac{1}{\log 2}}\gamma^{-1}}$. Then $\mathcal{N}_p$ is (ε,δ)-DP with respect to τ-neighbouring states with*

$$\delta \leq (1-p^k)(1-\beta)\eta + (1-p^k)\beta\tau. \tag{8.40}$$

8.5.1 Privacy via classical post-processing

Now, we show that the output of a quantum measurement can be privatised by adding classical noise. This finding is especially intriguing for two primary reasons. First, it offers a practical method for applying well-established tools and techniques from classical differential privacy to protect the outputs of quantum algorithms. Second, it enables us to combine classical noise with the output distributions resulting from quantum measurements.

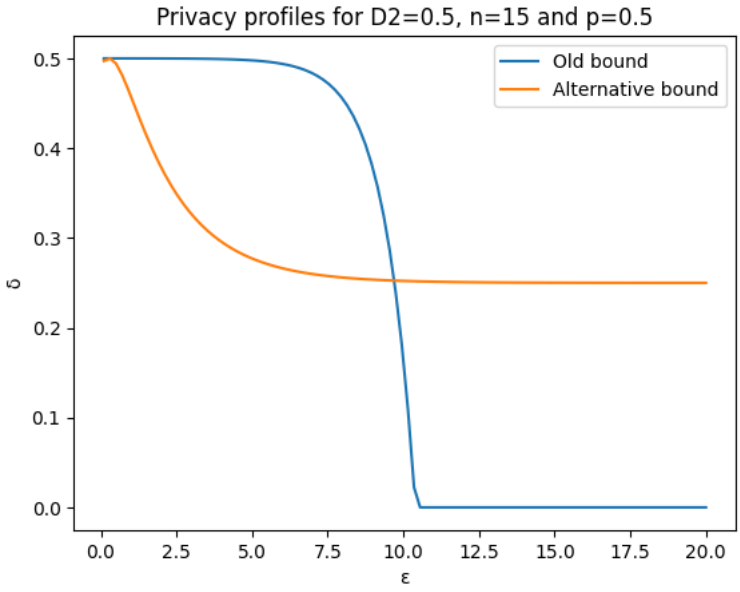


Figure 8.1: In this figure we compare the former upper bound from [HRF23] (Equation 8.30) with the novel upper bound provided in this section (Equation 8.31). We emphasise that each bound outperforms the other for some values of ε. We assumed that the input state satisfy $D_2(\rho\|I/2^n) \leq 0.5$, $n = 15$ and $p = 0.5$. The upper bound on τ is derived from $\|\rho-\sigma\|_1 \leq \|\rho - I/2^n\|_1 + \|\rho - I/2^n\|_1 \leq 2\sqrt{2D_2(\rho\|I/2^n)}$, i.e. combining the triangle inequality and the Pinsker's inequality.

Moreover, we remark that our analsys accounts for both quantum and classical noise. We do so by recognizing that quantum noisy channels reduce the trace distance between any two quantum states. Furthermore, the level of differential privacy achieved by adding classical noise is inversely proportional to the trace distance between neighboring states.

Lemma 8.2. *Let ρ, σ such that $\|\rho-\sigma\|_{\mathrm{tr}} \leq \tau$. Let M be a POVM measurement and Λ a classical channel such that $\forall x, x' \in \mathsf{range}(M) : E_{e^\varepsilon}(\Lambda(x)\|\Lambda(x')) \leq \delta$. Then we have that*

$$E_{e^{\varepsilon'}}(\Lambda(M(\rho))\|\Lambda(M(\sigma))) \leq \tau\delta, \tag{8.41}$$

where $\varepsilon' = \log(1+\tau(e^\varepsilon - 1))$, which for small ε gives $\varepsilon' \simeq \tau\varepsilon$.

Proof. Let $\nu := M(\rho)$ and $\nu' := M(\sigma)$. We have that

$$|\nu - \nu'|_{\mathrm{tv}} := \eta \leq \tau, \tag{8.42}$$

which follows from the data processing inequality. Moreover, there always exists some distributions ν_0, ν_1, ν_1' such that

$$\nu = (1-\eta)\nu_0 + \eta\nu_1, \qquad \nu' = (1-\eta)\nu_0 + \eta\nu_1'. \tag{8.43}$$

The above identities are discussed in detail in ([BBG18], Section 3). We also have,

$$\max\{E_{e^\varepsilon}(\Lambda(\nu_1)\|\Lambda(\nu_0)), E_{e^\varepsilon}(\Lambda(\nu_1)\|\Lambda(\nu_1'))\} \leq \delta \tag{8.44}$$

This follows by noting that ν_0, ν_1, ν_1' are supported in range(M) and applying the (standard) joint-convexity of the hockey-stick divergence. By advanced joint convexity (Lemma 11.2), we have that for all states ρ_0, ρ_1, ρ_2 and $\gamma' = 1 + (1-p)(\gamma - 1)$,

$$E_{\gamma'}(p\rho_0 + (1-p)\rho_1 \| p\rho_0 + (1-p)\rho_2) \leq (1-p)(1-\beta)E_\gamma(\rho_1\|\rho_0) + (1-p)\beta E_\gamma(\rho_1\|\rho_2), \tag{8.45}$$

Then,

$$E_{e^{\varepsilon'}}(\Lambda(M(\rho))\|\Lambda(M(\sigma))) \leq \tau\delta. \tag{8.46}$$

■

Lemma 8.2 is stated in terms of a general classical noisy channel. In the following theorem we consider the special cases of the Laplace and Gaussian mechanisms, two noisy channels widely used in many differentially private classical algorithms and defined in Section 7.2.

Theorem 8.3. *Let M a measurement whose possible outcomes are in the range $[a, a+\Delta]$ for $a \in \mathbb{R}$.*

- *(Laplace mechanism) Let $\Lambda_{\mathscr{L},b}$ the Laplace noise of scale b. Then $\Lambda_{\mathscr{L},b}(M(\cdot))$ is ε'-DP with respect to τ-neighbouring states, where*

$$\varepsilon' = \log(1 + \tau(e^{\Delta/b} - 1)). \tag{8.47}$$

- *(Gaussian mechanism) Let $\Lambda_{\mathscr{G},\sigma}$ the Gaussian noise of variance $\sigma^2 \geq 2\ln(1.25/\delta)\Delta^2/\varepsilon^2$. Then $\Lambda_{\mathscr{G},\sigma}(M(\cdot))$ is (ε', δ')-DP with respect to τ-neighbouring states, where*

$$\varepsilon' = \log(1 + \tau(e^{\varepsilon} - 1)) \quad \textit{and} \quad \delta' = \tau\delta. \tag{8.48}$$

Proof. The theorem follows by replacing the channel Λ in Lemma 8.2 with the Laplace and Gaussian noise, respectively. ■

8.5.2 Implications for quantum-inspired sampling

As the trace distance generalizes the total variation distance, the range of applicability of Theorem 8.3 includes also classical algorithms. In particular, we show here an application for private quantum-inspired sampling. In quantum-inspired algorithms [Tan19, Tan21, GLT18, CLW18, CGL+20], a classical vector $u \in \mathbb{C}^{2^n}$ is accessed through quantum-inspired sampling: i.e. an entry u_i is sampled with probability proportional to $|u_i|^2$. This is equivalent to encoding u into the state

$$|u\rangle = \frac{1}{\|u\|_2} \sum_{i\in\{0,1\}^n} u_i |i\rangle, \tag{8.49}$$

and performing a computational-basis measurement. Let p_u be the distribution induced by such measurements. Say that $u \sim u'$ if u and u' differ in only one entry. In particular, let $u_i = u_i'$ for all $i \neq j$.

$$|\|u\|_2^2 - \|u'\|_2^2| = \left|\sum_i |u_i|^2 - \sum_i |u_i'|^2\right| \tag{8.50}$$

$$\leq \left|\sum_{i\neq j} |u_i|^2 - \sum_{i\neq j} |u_i'|^2 + |u_j|^2 - |u_j'|^2\right| \leq \max\{|u_j|^2, |u_j'|^2\} \tag{8.51}$$

It's easy to see that p_u and $p_{u'}$ are close in total variation distance.

$$|p_u - p_{u'}|_{\text{tv}} = \frac{1}{2}\sum_i \left| \frac{|u_i|^2}{\|u\|_2^2} - \frac{|u'_i|^2}{\|u'\|_2^2} \right| \tag{8.52}$$

$$\leq \frac{1}{2}\left(\sum_{i\neq j} |u_i|^2 \left| \frac{1}{\|u\|_2^2} - \frac{1}{\|u'\|_2^2} \right| + \left| \frac{|u_j|^2}{\|u\|_2^2} - \frac{|u'_j|^2}{\|u'\|_2^2} \right| \right) \tag{8.53}$$

$$\leq \frac{1}{2}\left(\min\{\|u\|_2^2, \|u'\|_2^2\} \frac{\max\{|u_j|^2, |u'_j|^2\}}{\|u\|_2^2\|u'\|_2^2} + \frac{|u_j|^2}{\|u\|_2^2} + \frac{|u'_j|^2}{\|u'\|_2^2} \right) \tag{8.54}$$

$$\leq \frac{3}{2}\max\left\{ \frac{|u_j|^2}{\|u\|_2^2}, \frac{|u'_j|^2}{\|u'\|_2^2} \right\} := \alpha. \tag{8.55}$$

Then, by subadditivity of the total variation distance,

$$|p_u^{\otimes m} - p_{u'}^{\otimes m}|_{\text{tv}} \leq m\alpha. \tag{8.56}$$

We will show the intuitive fact that quantum-inspired sampling amplifies differential privacy. First, we can consider the encoding $u \mapsto p_u^{\otimes m}$ and derive the following special case of Theorem 8.3.

Corollary 8.5. *Let u, u' be neighbouring if they differ in at most one entry. Consider the oracle O_u that returns a u_i with probability $\frac{|u_i|^2}{\|u\|_2^2}$. For $a \in \mathbb{R}$ and $\Delta \geq 0$, let $\mathscr{S}$ a randomised algorithm with range $[a, a+\Delta]$ that makes m queries to O_u and assume that $\frac{3}{2}\frac{|u_j|^2}{\|u\|_2^2} \leq \alpha$.*

- *(Laplace mechanism) Let $\Lambda_{\mathscr{L},b}$ the Laplace noise of scale b. Then $\Lambda_{\mathscr{L},b}(\mathscr{S}(\cdot))$ is ε'-DP, where*

$$\varepsilon' = \log(1 + \alpha m(e^{\Delta/b} - 1)). \tag{8.57}$$

- *(Gaussian mechanism) Let $\Lambda_{\mathscr{G},\sigma}$ the Gaussian noise of variance $\sigma^2 \geq 2\ln(1.25/\delta)\Delta^2/\varepsilon^2$. Then $\Lambda_{\mathscr{G},\sigma}(\mathscr{S}(\cdot))$ is (ε', δ')-DP, where*

$$\varepsilon' = \log(1 + \alpha m(e^{\varepsilon} - 1)) \quad \textit{and} \quad \delta' = \alpha m\delta. \tag{8.58}$$

The approach described above is tailored to noise-adding mechanisms. In Section 11.3 we provide a more general result that applies to any private mechanism and it builds upon prior work on privacy amplification by subsampling [BBG18, Ull17].

8.6 Differential privacy for (Ξ, τ)-neighboring states

While in Section 8.5 we provided tighter bounds for quantum differential privacy with respect to states with bounded trace distance, here we add two additional ingredients: the locality of the measurements and the generalized neighboring relationship defined in Section 8.4. Under these stronger assumptions, we can improve the privacy guarantees of local noisy channels and classical post-processing. First, we need to introduce the following quantity.

Definition 8.2 (Worst-case quantum sensitivity)**.** Let O be an observable expressed as a weighted sum of Pauli operators, $O=\sum_{P\in\mathscr{P}_n} c_P P$. Let $\mathscr{I}\subseteq[n]$ and consider the subset $\mathscr{S}_\mathscr{I}$ of all the Pauli strings that act non trivially on $\mathscr{I}$. The worst-case quantum sensitivity of O with respect to $\mathscr{I}$ is defined as

$$\Delta(O;\mathscr{I}) := 2\sum_{P\in\mathscr{S}_\mathscr{I}} |c_P|. \tag{8.59}$$

Let $\Xi\subseteq P([n])$, i.e. Ξ is a collection of subsets of $[n]$. The worst-case quantum sensitivity of O with respect to Ξ is defined as

$$\Delta_\Xi(O) := \max_{\mathscr{I}\in\Xi}\Delta(O;\mathscr{I}). \tag{8.60}$$

We will omit the index Ξ and simply write $\Delta(O)$ when there is no ambiguity.

So, if $O=\sum_{i=1}^n Z_i$ and $\Xi=\{\{1\},\{2\},\dots,\{n\}\}$, the worst-case quantum sensitivity equals $\Delta(O)=2$. This is consistent with the fact that, if ρ and σ satisfy $\mathrm{Tr}_j\,\rho=\mathrm{Tr}_j\,\sigma$, then all the terms but Z_j induce the same distributions when measured on either ρ or σ. Moreover, the outcome of term Z_j will be either 1 or -1, then it belongs to an interval of length 2. We can also consider the more general case where $O_\ell=\sum_{i=1}^n\bigotimes_{j=i}^{i+\ell-1} Z_j$ and $\Xi=\{\{i,i+1,\dots,i+k\}|$ for $i=1,2,\dots,n-k\}$. It's easy to see that $\Delta(O_\ell)=2k+4\ell-4$.

We can now state the first result of this section, concerning a class of local noisy channels, which includes the local Pauli noise.

Theorem 8.4 (Generalized private measurement via local noisy channels)**.** *Let $O=\sum_P c_P P$ be an observable consisting of a weighted sum of commuting Pauli operators, and let $\mathscr{O}$ the quantum-to-classical channel implementing a measurement of O. Let $\mathscr{M}$ an arbitrary single qubit channel and let $\mathscr{N}(\cdot)=pI/2+(1-p)\mathscr{M}(\cdot)$. Let $k=\max_{\mathscr{I}\in\Xi}|\mathscr{I}|$. Then $\mathscr{O}\circ\mathscr{N}^{\otimes n}$ satisfies (ε,δ_k)-DP with respect to (Ξ,τ)-neighboring states, where*

$$\delta_k \le \max\left\{0,(1-e^\varepsilon)\frac{p^k}{2^k}+(1-p^k)\tau\right\}. \tag{8.61}$$

Let $\gamma=1+(e^\varepsilon-1)/(1-p)$ and $\beta=e^\varepsilon/\gamma$. Under the additional assumption the the input state ρ satisfies $E_\gamma(\rho\|I/2^n)\le\eta$, the following inequality also holds

$$\delta_k \le (1-p^k)(1-\beta)\eta+(1-p^k)\beta\tau. \tag{8.62}$$

Proof. Since $\rho\overset{(\Xi,\tau)}{\sim}\sigma$, there exists $\mathscr{I}\in\Xi$ such that

$$\mathrm{Tr}_\mathscr{I}\,\rho=\mathrm{Tr}_\mathscr{I}\,\sigma \quad\text{and}\quad |\mathscr{I}|\le k. \tag{8.63}$$

We also have

$$\mathscr{N}^{\otimes n}(\rho)=p^{|\mathscr{I}|}\left(\mathrm{Tr}_\mathscr{I}\,\mathscr{M}(\rho)\otimes\frac{I}{2^{|\mathscr{I}|}}\right)+(1-p^{|\mathscr{I}|})\mathscr{M}'(\rho), \tag{8.64}$$

for a suitable channel $\mathscr{M}'$. This stems from the fact that with probability $p^{|\mathscr{I}|}$, the reduced state of the subset $\mathscr{I}$ is mapped to the maximally mixed state $I/2^{|\mathscr{I}|}$.

Then, the measurement O can be implemented by measuring each qubit in a different Pauli basis and then performing classical postprocessing. As quantum differential privacy is robust to postprocessing, we only need to prove that Pauli measurements preserve (ε, δ_k)-DP. We can assume without loss of generality that the qubits in the subsystem $\mathscr{I}^c := [n] \setminus \mathscr{I}$ are measured first, since we assumed that O is a weighted sum of commuting Pauli operators, and hence the measurement order does not alter the overall statistics.

Assume that measuring the subsystem $\mathscr{I}^c$ produces the outcome $\boldsymbol{y} \in \{\pm 1\}^{n-|\mathscr{I}|}$. Equation 8.63 implies that

$$p_{\boldsymbol{y}} := \Pr[\boldsymbol{y} \text{ is obtained on input } \rho] = \Pr[\boldsymbol{y} \text{ is obtained on input } \sigma]. \tag{8.65}$$

So measuring the qubits in $\mathscr{I}^c$ does not allow to distinguish between ρ and σ.

Denote by $\rho_{\boldsymbol{y}}$ the post-measurement state produced by measuring the system $\mathscr{I}^c$ and obtaining outcome $\boldsymbol{y}$. Let $\mathscr{T}_{\boldsymbol{y}}$ be the quantum channel mapping ρ to $\rho_{\boldsymbol{y}}$. We now show that the channel $\operatorname{Tr}_{\mathscr{I}^c} \circ \mathscr{T}_{\boldsymbol{y}} \circ \mathscr{N}^{\otimes n}$ preserves (ε, δ_k)-differential privacy.

$$\operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\rho)) \tag{8.66}$$

$$= p^{|\mathscr{I}|} \operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}} \left(\operatorname{Tr}_{\mathscr{I}} \mathscr{M}(\rho) \otimes \frac{I}{2^{|\mathscr{I}|}} \right) + (1 - p^{|\mathscr{I}|}) \operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{M}'(\rho)) \tag{8.67}$$

$$= p^{|\mathscr{I}|} \frac{I}{2^{|\mathscr{I}|}} + (1 - p^{|\mathscr{I}|}) \operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{M}'(\rho)) \tag{8.68}$$

$$:= p^{|\mathscr{I}|} \frac{I}{2^{|\mathscr{I}|}} + (1 - p^{|\mathscr{I}|}) \mathscr{M}''(\rho). \tag{8.69}$$

where the second equality follows from $\mathscr{T}_{\boldsymbol{y}} \left(\operatorname{Tr}_{\mathscr{I}} \mathscr{M}(\rho) \otimes \frac{I}{2^{|\mathscr{I}|}} \right) = \operatorname{Tr}_{\mathscr{I}} \mathscr{T}_{\boldsymbol{y}}(\mathscr{M}(\rho)) \otimes \frac{I}{2^{|\mathscr{I}|}}$ and we defined $\mathscr{M}'' := \operatorname{Tr}_{\mathscr{I}^c} \circ \mathscr{T}_{\boldsymbol{y}} \circ \mathscr{M}'$. By Corollary 8.2,

$$E_{e^\varepsilon}(\operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\rho)) \| \operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\sigma))) \leq \delta_k. \tag{8.70}$$

So far, we have proved (ε, δ_k)-differential privacy conditioning to a fixed value of $\boldsymbol{y}$. In order to prove that measuring O on $\mathscr{N}^{\otimes n}(\rho)$ preserves (ε, δ_k)-DP, it is sufficient consider the outcome $\boldsymbol{y}$ and the partial post-measurement state $\operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\rho))$. Thus we need to ensure that

$$E_{e^\varepsilon} \left(\sum_{\boldsymbol{y}} p_{\boldsymbol{y}} (\operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\rho)) \otimes |\boldsymbol{y}\rangle \langle \boldsymbol{y}|) \middle\| \sum_{\boldsymbol{y}} p_{\boldsymbol{y}} (\operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\sigma)) \otimes |\boldsymbol{y}\rangle \langle \boldsymbol{y}|) \right) \leq \delta(\varepsilon, k) \tag{8.71}$$

We also have, for all $\gamma \geq 1$,

$$E_\gamma \left(\sum_{\boldsymbol{y}} p_{\boldsymbol{y}} (\operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\rho)) \otimes |\boldsymbol{y}\rangle \langle \boldsymbol{y}|) \middle\| \sum_{\boldsymbol{y}} p_{\boldsymbol{y}} (\operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\sigma)) \otimes |\boldsymbol{y}\rangle \langle \boldsymbol{y}|) \right) \tag{8.72}$$

$$\leq \sum_{\boldsymbol{y}} p_{\boldsymbol{y}} E_\gamma \left(\operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\rho)) \otimes |\boldsymbol{y}\rangle \langle \boldsymbol{y}| \,\|\, \operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\sigma)) \otimes |\boldsymbol{y}\rangle \langle \boldsymbol{y}| \right) \tag{8.73}$$

$$\leq \sum_{\boldsymbol{y}} p_{\boldsymbol{y}} E_\gamma \left(\operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\rho)) \| \operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\sigma)) \right) \tag{8.74}$$

$$\leq \max_{\boldsymbol{y}} E_\gamma \left(\operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\rho)) \| \operatorname{Tr}_{\mathscr{I}^c} \mathscr{T}_{\boldsymbol{y}}(\mathscr{N}^{\otimes n}(\sigma)) \right), \tag{8.75}$$

where the second line follows from the convexity of the hockey-stick divergence (Equation 3.54) and the third line follows from the stability of the hockey-stick divergence (Equation 3.55). Combining Equation 8.70 with Equation 8.72 gives the desired result:

$$E_{e^{\varepsilon}}(\mathscr{O}(\mathscr{N}^{\otimes n}(\rho))\|\mathscr{O}(\mathscr{N}^{\otimes n}(\sigma))) \leq \delta_k. \tag{8.76}$$

■

We emphasize that the number of qubits n appearing in the guarantees of Corollary 8.2 is now replaced by $k = \max_{\mathscr{I}\in\Xi}|\mathscr{I}|$. Thus if $k = \text{polylog}(n)$, this new bound is exponentially tighter than the previous one. In a similar fashion, we can adapt Theorem 8.3 to the generalized neighboring relationship, by employing the worst-case quantum sensitivity introduced in Definition 8.2.

Theorem 8.5 (Generalized private measurement via classical post-processing)**.** *Let ρ and σ two (Ξ,τ)-neighboring quantum states, i.e. $\rho \overset{(\Xi,\tau)}{\sim} \sigma$. Let O be an observable, and denote $\mathscr{O}$ as a quantum-to-classical channel implementing a measurement of O.*

- *(Laplace mechanism) Let $\Lambda_{\mathscr{L},b}$ the Laplace noise of scale b. Then $\Lambda_{\mathscr{L},b}(\mathscr{O}(\cdot))$ is ε'-DP with respect to (Ξ,τ)-neighboring states, where*

$$\varepsilon' = \log(1+\tau(e^{\Delta(O)/b}-1)). \tag{8.77}$$

- *(Gaussian mechanism) Let $\Lambda_{\mathscr{G},\sigma}$ the Gaussian noise of variance $\sigma^2 \geq 2\log(1.25/\delta)\Delta(O)^2/\varepsilon^2$. Then $\Lambda_{\mathscr{G},\sigma}(\mathscr{O}(\cdot))$ is (ε',δ')-DP with respect to (Ξ,τ)-neighboring states, where*

$$\varepsilon' = \log(1+\tau(e^{\varepsilon}-1)) \quad \textit{and} \quad \delta' = \tau\delta. \tag{8.78}$$

Proof. Proceeding as in the proof of Theorem 8.4, consider $\mathscr{I}\in\Xi$ such that $\text{Tr}_{\mathscr{I}}\rho = \text{Tr}_{\mathscr{I}}\sigma$ and let $\mathscr{S}_{\mathscr{I}}$ be the subset of all the Pauli strings that act non trivially on $\mathscr{I}$. Thus, we can decompose O as $O = O_1 + O_2$, where $O_1 = \sum_{P\notin\mathscr{S}_{\mathscr{I}}} c_P P$ and $O_2 = O - O_1 = \sum_{P\in\mathscr{S}_{\mathscr{I}}} c_P P$. Assume without loss of generality that O_1 is measured first. Since $\text{Tr}_{\mathscr{I}}\rho = \text{Tr}_{\mathscr{I}}\sigma$ and O_1 acts non trivially only on $\mathscr{I}^c = [n]\setminus\mathscr{I}$, then this measurement produces no loss of privacy, i.e.

$$\forall y : p(y) := \Pr_{\rho}[O_1 = y] = \Pr_{\sigma}[O_1 = y]. \tag{8.79}$$

Observe that O_2 is a measurement whose output is comprised into $[-\Delta(O)/2, \Delta(O)/2]$. Moreover, let ρ_y be the post-measurement state obtained when O_1 returns outcome y. As the trace distance is non-increasing, we have,

$$\|\rho_y - \sigma_y\|_{\text{tr}} \leq \|\rho-\sigma\|_{\text{tr}} \leq \tau, \tag{8.80}$$

Conditioning on input y, the output of $O = O_1 + O_2$ lies in $[y-\Delta/2, y+\Delta/2]$. Then Theorem 8.3 yields

$$E_{e^{\varepsilon'}}\left(\sum_y p(y)\Lambda_{\mathscr{L},b}(\mathscr{O}(\rho_y))\middle\|\sum_y p(y)\Lambda_{\mathscr{L},b}(\mathscr{O}(\rho_y))\right) \tag{8.81}$$

$$\leq \max_y E_{e^{\varepsilon'}}\left(\Lambda_{\mathscr{L},b}(\mathscr{O}(\rho_y))\|\Lambda_{\mathscr{L},b}(\mathscr{O}(\rho_y))\right) \leq 0. \tag{8.82}$$

for $\varepsilon' = \log(1+\tau(e^{\Delta(O)/b}-1))$. Similarly, replacing the Laplace noise with the Gaussian noise and applying again Theorem 8.3,

$$E_{e^\varepsilon}\left(\sum_y p(y)\Lambda_{\mathscr{G},\sigma}(\mathscr{O}(\rho_y))\middle\|\sum_y p(y)\Lambda_{\mathscr{G},\sigma}(\mathscr{O}(\rho_y))\right) \tag{8.83}$$

$$\leq \max_y E_{e^\varepsilon}\left(\Lambda_{\mathscr{G},\sigma}(\mathscr{O}(\rho_y))\|\Lambda_{\mathscr{G},\sigma}(\mathscr{O}(\rho_y))\right) \leq \delta', \tag{8.84}$$

where $\sigma^2 \geq 2\log(1.25/\delta)\Delta(O)^2/\varepsilon^2$, $\varepsilon' = \log(1+\tau(e^\varepsilon-1))$ and $\delta' = \tau\delta$. ■

We observe that similar results can be derived for multiple sources of noise, beyond the Laplace or the Gaussian channels, along the lines of Lemma 8.2. We leave it to the reader to extend Theorem 8.5 to alternative stochastic channels.

8.7 The cost of quantum differential privacy

Differential privacy, both in the classical and in the quantum setting, can be achieved by introducing noise into the computation, thus reducing the final accuracy. Intuitively, large values of ε can be attained with little loss in accuracy, while for $\varepsilon = 0$ the output is totally independent of the input. In particular, if an algorithm is ε-DP with respect to Hamming distance, we have that

$$\forall x, x' : D_\infty(\mathscr{A}(x)\|\mathscr{A}(x')) \leq \varepsilon n, \tag{8.85}$$

thus if $\varepsilon = O(1/n)$, any pair of inputs (not necessarily neighboring) are mapped to outputs $O(1)$-close in max-divergence. This result follows from the fact that the max-relative entropy satisfies the triangle inequality (both in the classical and in the quantum cases), i.e. $\forall \rho_1, \rho_2, \sigma : D_\infty(\rho_1\|\rho_2) \leq D_\infty(\rho_1\|\sigma) + D_\infty(\sigma\|\rho_2)$. We can pick a sequence of $n+1$ inputs $x_0, x_1, \ldots, x_n$ such that $x = x_0$, $x' = x_n$ and $x_i \sim x_{i+1}$. Then iterating the triangle inequality yields Equation 8.85. However, for most applications ε can be chosen as a constant independent of n, avoiding this undesired concentration of the output around a unique value.

A vast portion of the literature about differential privacy is devoted to optimising the tradeoff between the value of ε and the loss in utility. In this section we make a crucial observation: the privacy-utility tradeoff doesn't depend solely on the value of ε, but also on the notion of neighboring inputs. Thus, the privacy-utility tradeoff is an important figure of merit for the comparison of different approaches to quantum differential privacy.

In particular, we argue that some prior definitions of neighboring quantum states suffer from a poor tradeoff between privacy and accuracy, leading to a suboptimal scaling with respect to the number of qubits n. This is the case, for instance, if we require two neighboring states to have bounded trace distance $\tau = \Theta(1)$. We also provide a similar result for the Wasserstein distance of order 1.

8.7.1 Concentration inequalities for private measurements

It's well known that noisy quantum algorithms suffer from severe limitations, that often hinder quantum advantage. Prior works [FGP21, DPMRF23] showed that, if the noise exceeds a given threshold, the output of noisy devices is concentrated around the maximally mixed state, and then it can be efficiently approximated with a classical computer. Since quantum differential privacy involves the injection of noise, it's not surprising that similar concentration inequalities hold for quantum private algorithms. In the remainder of this section, we will show how this concentration affects the accuracy of private measurements. For the sake of simplicity, we will state our results in terms of simple, local observables such as $O = \sum_{i=1}^{n} Z_i$. Similar results can be obtained for any observable with bounded Lipschitz constant, as also discussed in [DPMRF23], but our choice is sufficient to display the shortcomings of a poor choice of the neighboring relationship. If we measure O on the maximally mixed state $I/2^n$, the outcome satisfies a Gaussian concentration inequality [DPMRF23]:

$$\Pr_{I/2^n}(|O| \geq an) \leq Ke^{-a^2 n}, \tag{8.86}$$

for $K = 1$. So, if a state ρ satisfies $D_\infty(\rho \| I/2^n) \leq \varepsilon$, the definition of the quantum max-relative entropy yields,

$$\Pr_{\rho}(|O| \geq an) \leq e^{\varepsilon} \Pr_{I/2^n}(|O| \geq an) \leq K' e^{-a^2 n}, \tag{8.87}$$

where $K' = e^{\varepsilon}$. For the sake of simplicity, throughout this section, we consider the special case of *pure* differential privacy, i.e. $(\varepsilon, 0)$-DP, but our results can be suitably extended to the more general *approximate* differential privacy, i.e. (ε, δ)-DP, under the assumption that $\delta \ll 1$.

Consider a quantum channel $\mathscr{A}(\cdot)$ and assume for the sake of simplicity that $\mathscr{A}$ is unital, i.e. $\mathscr{A}(I) = I$. We show that different neighboring relationships $\overset{Q}{\sim}$ have a disparate impact on the accuracy. The first result is devoted to states with bounded trace distances.

Theorem 8.6 (Concentration inequality for bounded trace distance)**.** *Consider the observable $O = \sum_{i=1}^{n} Z_i$ and let $\mathscr{A}$ be a unital quantum channel satisfying ε-DP with respect to τ-neighboring states, i.e. $D_\infty(\mathscr{A}(\rho)\|\mathscr{A}(\sigma)) \leq \varepsilon$ if $\frac{1}{2}\|\rho - \sigma\|_1 \leq \tau$. Assume $\tau = \Theta(1)$. Then, for any input state ρ, the output $\mathscr{A}(\rho)$ satisfies the following concentration inequality:*

$$\Pr_{\mathscr{A}(\rho)}(|O| \geq an) \leq K' e^{-a^2 n}, \tag{8.88}$$

where $K' = e^{O(\varepsilon)}$.

Proof. For two arbitrary quantum states, we have

$$\forall \rho, \sigma : D_\infty(\mathscr{A}(\rho)\|\mathscr{A}(\sigma)) \leq \varepsilon/\tau. \tag{8.89}$$

This can be seen by building the following chain :

$$\rho_i = \rho \max(0, 1 - i\tau) + \sigma \min(1, i\tau) \tag{8.90}$$

We note that $\frac{1}{2}\|\rho_i - \rho_{i+1}\|_1 \leq \tau$ which implies $D_\infty(\mathscr{A}(\rho_i)\|\mathscr{A}(\rho_{i+1})) \leq \varepsilon$. Then Equation 8.89 can be deduced by iterating the triangle inequality. Combining it with Equation 8.87, we obtain

$$\forall \rho : \Pr_{\mathscr{A}(\rho)}(|O| \geq an) \leq Ke^{-a^2 n}, \tag{8.91}$$

where $K = e^{\varepsilon/\tau} = e^{O(\varepsilon)}$. ■

To showcase the implications of the Theorem 8.6, we set $\tau = 0.1$ and we consider $\rho := |1^n\rangle\langle 1^n|$. We remark that ρ is an eigenvector of O, with eigenvalue n. However, instead of measuring O directly, we can post-process ρ with a ε-DP channel $\mathscr{A}$ as defined in the statement of the theorem. Set $\varepsilon = 1$. In order to achieve an error smaller than, say, $0.5n$, we need to ensure that the outcome is larger than $0.9n$. Then Theorem 8.6 implies that the error is larger than $0.5n$ with high probability:

$$\Pr_{\mathscr{A}(\rho)}(|n - O| \leq 0.5n) = \Pr_{\mathscr{A}(\rho)}(O \geq 0.5n) \leq \Pr_{\mathscr{A}(\rho)}(|O| \geq 0.5n) \leq e^{10-0.25n} \tag{8.92}$$

and hence setting $n = 100$ we obtain

$$\Pr_{\mathscr{A}(\rho)}(|n - O| \leq 0.5n) \leq 3 \times 10^{-7}. \tag{8.93}$$

Now, we provide a similar result for another neighboring definition. In [DPMTL21b], the authors extend the Wasserstein distance of order 1 (or W_1 distance) to quantum states and suggest quantum differential privacy as a potential application of their work. Recall that the W_1 distance between the quantum states ρ and σ of $\mathscr{H}_n$ is defined as

$$W_1(\rho,\sigma) = \min\Big(\sum_{i=1}^n c_i : c_i \geq 0, \rho - \sigma = \sum_{i=1}^n c_i\left(\rho^{(i)} - \sigma^{(i)}\right), \tag{8.94}$$

$$\rho^{(i)}, \sigma^{(i)} \in \mathscr{S}_n, \mathrm{Tr}_i\,\rho^{(i)} = \mathrm{Tr}_i\,\sigma^{(i)}\Big). \tag{8.95}$$

The following theorem shows that the W_1 distance leads to the following undesired concentration inequality.

Theorem 8.7 (Concentration inequality for bounded W_1 distance)**.** *Consider the observable $O = \sum_{i=1}^n Z_i$ and let $\mathscr{A}$ be a unital quantum channel satisfying ε-DP with respect stated with W_1 distance bounded by 1, i.e. $D_\infty(\mathscr{A}(\rho_1)\|\mathscr{A}(\rho_2)) \leq \varepsilon$ if $W_1(\rho_1,\rho_2) \leq 1$. Then, for any input state ρ, the output $\mathscr{A}(\rho)$ satisfies the following concentration inequality:*

$$\Pr_{\mathscr{A}(\rho)}(|O| \geq an) \leq K' e^{-a^2 n}, \tag{8.96}$$

where $K' = e^{\varepsilon}(n - e^{-\varepsilon}(n-1))$.

Proof. Quantum differential privacy with respect to bounded Wasserstein distance of order 1 can be expressed as:

$$W_1(\rho_1,\rho_2) \leq 1 \implies D_\infty(\mathscr{A}(\rho_1)\|\mathscr{A}(\rho_2)) \leq \varepsilon. \tag{8.97}$$

We show that even this definition causes the output state to be highly concentrated around zero, independent of the input state. In particular, we show that for two arbitrary quantum states ρ and σ, we have

$$\forall \rho, \sigma : D_\infty(\mathscr{A}(\rho) \| \mathscr{A}(\sigma)) \leq \varepsilon', \tag{8.98}$$

where $\varepsilon' = \varepsilon + \log(n - ne^{-\varepsilon} + e^{-\varepsilon})$. This can be seen considering the mixture $\rho' := \left(1 - \frac{1}{n}\right)\rho + \frac{\sigma}{n}$ and noting that $W_1(\rho, \rho') \leq 1$. Then, by the definition of ε-differential privacy,

$$\left(1 - \frac{1}{n}\right) \mathrm{Tr}[M_m \mathscr{A}(\rho)] + \frac{1}{n} \mathrm{Tr}[M_m \mathscr{A}(\sigma)] \tag{8.99}$$

$$= \mathrm{Tr}[M_m \mathscr{A}(\rho')] \leq e^{\varepsilon} \mathrm{Tr}[M_m \mathscr{A}(\rho)] \tag{8.100}$$

And thus

$$\mathrm{Tr}[M_m \mathscr{A}(\sigma)] \leq e^{\varepsilon}(n - e^{-\varepsilon}(n-1)) \mathrm{Tr}[M_m \mathscr{A}(\rho)] \tag{8.101}$$

$$= e^{\varepsilon'} \mathrm{Tr}[M_m \mathscr{A}(\rho)], \tag{8.102}$$

which implies Equation 8.98. Then, for any input ρ, $\mathscr{A}(\rho)$ is ε-close to the maximally mixed state in quantum max-relative entropy, up to additive logarithmic factors. Applying Equation 8.87 yields

$$\Pr_{\mathscr{A}(\rho)}(|O| \geq an) \leq K' e^{-a^2 n}, \tag{8.103}$$

where where $K = e^{\varepsilon'} = e^{\varepsilon}(n - e^{-\varepsilon}(n-1))$. ■

Proceeding similarly as for the trace distance, set $\rho := |1^n\rangle\langle 1^n|$ and $\varepsilon = 1$. Theorem 8.7 implies that

$$\Pr_{\mathscr{A}(\rho)}(|n - O| \leq 0.5n) = \Pr_{\mathscr{A}(\rho)}(O \geq 0.5n) \leq \Pr_{\mathscr{A}(\rho)}(|O| \geq 0.5n) \leq (en - (n-1))e^{-0.25n} \tag{8.104}$$

and hence setting $n = 100$ we obtain

$$\Pr_{\mathscr{A}(\rho)}(|n - O| \leq 0.5n) \leq 2.4 \times 10^{-9}. \tag{8.105}$$

Then the above example can be considered as a no-go result concerning $(\varepsilon, 0)$-DP under Wasserstein distance of order 1. We emphasise that the main argument of Theorem 8.7 is based on the construction of a classical mixed state, and then it holds both for the classical and the quantum W_1 distance. On the other hand, one could define the neighboring relationship solely on pure states and hence overcome our no-go result. However, it is not obvious whether this definition can lead to a good privacy-utility tradeoff. We leave this possibility as an open problem for future explorations.

We also remark that $(0, \delta)$-DP under the W_1 distance is equivalent to $(0, \delta)$-DP with respect to $(1,1)$-neighboring quantum states. Assume that a channel $\mathscr{A}$ is $(0, \delta)$-DP with respect to $(1,1)$-neighboring quantum states and let $M = (M_1, \ldots, M_k)$ be a POVM measurement

$$\forall \rho_1 \overset{(1,1)}{\sim} \rho_2 \forall S \subseteq [k] \sum_{j \in S} \mathrm{Tr}\left[M_j(\mathscr{A}(\rho_1) - \mathscr{A}(\rho_2))\right] \leq \delta. \tag{8.106}$$

Then,

$$\sum_{j\in S} \mathrm{Tr}\left[M_j(\mathscr{A}(\rho)-\mathscr{A}(\sigma))\right] \leq \sum_{j\in S}\sum_{i=1}^{n} c_i \,\mathrm{Tr}\left[M_j(\mathscr{A}(\rho^{(i)})-\mathscr{A}(\sigma^{(i)}))\right] \tag{8.107}$$

$$= \sum_{i=1}^{n} c_i \sum_{j\in S} \mathrm{Tr}\left[M_j(\mathscr{A}(\rho^{(i)})-\mathscr{A}(\sigma^{(i)}))\right] \leq \sum_{i=1}^{n} c_i\delta = W_1(\rho,\sigma)\delta, \tag{8.108}$$

where the last inequality follows from $\rho^{(i)} \overset{(1,1)}{\sim} \sigma^{(i)}$. Since $(1,1)$-neighboring states satisfies $W_1(\rho,\sigma) \leq 1$, the equivalence follows.

8.7.2 A positive result for (ℓ,τ)-neighboring states

We conclude this section with a positive result: adopting the definition introduced in Section 8.6, we can privately sample from an observable that approximates $O = \sum_{i=1}^{n} Z_i$, with a small loss in accuracy. We remark that the special case $\ell = \tau = 1$ has already been studied in [AR19].

Theorem 8.8 (Efficient private measurement for (ℓ,τ)-neighboring states)**.** *Let $\mathscr{O}$ be the quantum to classical channel implementing a measurement of the observable $O = \sum_{i=1}^{n} Z_i$. Assume that a state ρ satisfies*

$$\Pr_{\rho}[|O - \langle O\rangle_\rho| > a] \leq b. \tag{8.109}$$

and let $\alpha := \frac{2\ell}{\log((e^\varepsilon-1)\tau^{-1}+1)} \approx 2\ell\tau\varepsilon^{-1}$. Then there exits a a quantum-to-classical channel $\mathscr{O}_\varepsilon$ such that:

1. *$\mathscr{O}_\varepsilon$ is ε-DP with respect to (ℓ,τ)-neighboring states.*
2. *The following concentration inequality holds:*

$$\Pr[|\mathscr{O}_\varepsilon(\rho) - \langle O\rangle_\rho| > a + t\alpha] \leq b + e^{-t}. \tag{8.110}$$

Proof. Let $\Lambda_{\mathscr{L}}$ be the Laplace noise of magnitude α. The first part of the theorem follows directly from Theorem 8.3, by choosing $\mathscr{O}_\varepsilon = \Lambda_{\mathscr{L}} \circ \mathscr{O}$. Moreover, if $Y \sim \mathrm{Lap}(\alpha)$, then

$$\Pr[|Y| > t\cdot\alpha] = e^{-t}. \tag{8.111}$$

Define the event $E = \{|Y| \leq t\alpha\}$. Then we have

$$\Pr[|\mathscr{O}_\varepsilon(\rho) - \langle O\rangle_\rho| > a + t\alpha] \tag{8.112}$$

$$\leq \Pr[|\mathscr{O}_\varepsilon(\rho) - \langle O\rangle_\rho| > a + t\alpha|E]\Pr[E] + \Pr[|\mathscr{O}_\varepsilon(\rho) - \langle O\rangle_\rho| > a + t\alpha|\overline{E}]\Pr[\overline{E}] \tag{8.113}$$

$$\leq \Pr_{\rho}[|O - \langle O\rangle_\rho| > a] + \Pr[|Y| > t\cdot\alpha] \leq b + e^{-t}. \tag{8.114}$$

■

So, in particular, $\rho = |1^n\rangle\langle 1^n|$, we have that $\Pr_\rho[O = \langle O\rangle_\rho] = 1$ since ρ is an eigenvector of O. Then Theorem 8.8 yields

$$\Pr[|\mathscr{O}_\varepsilon(\rho) - n| < t\cdot\alpha] \geq 1 - e^{-t}. \tag{8.115}$$

Finally, we plot the upper bounds derived in this section in Figure 8.2 and Figure 8.3.

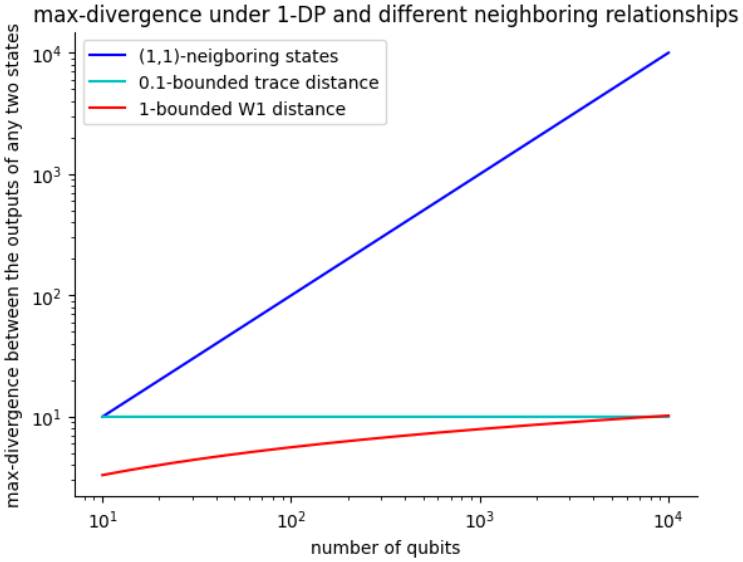


Figure 8.2: Upper bounds on the quantum max-relative entropy between any two states under 1-DP for several neighboring relationships and various values of n.

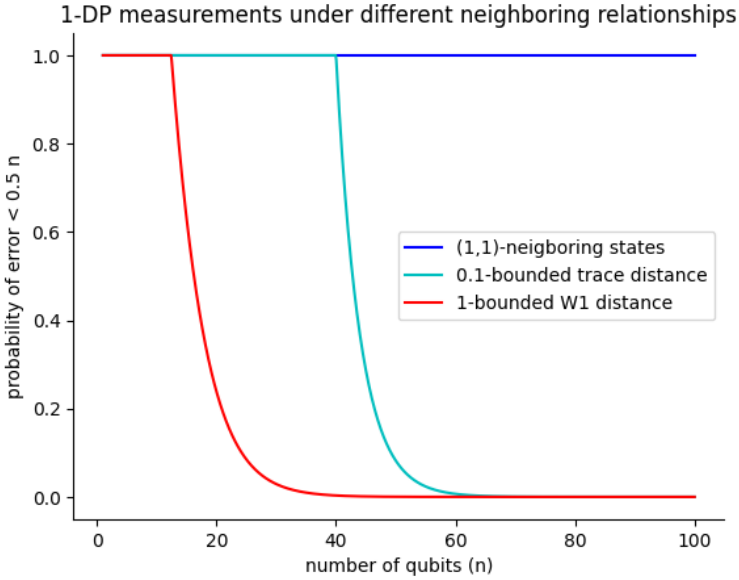


Figure 8.3: Upper bounds on the probability of achieving error lower than $0.5n$ for a measurement of $\frac{1}{n}\sum_{i=1}^{n} Z_i$ on the state $|1^n\rangle$, for several neighboring relationships and various values of n. We assumed the input state undergoes a 1-DP channel.

8.8 Privacy-preserving estimation of expected values

In this section, we provide differentially private mechanisms for estimating the expected values of observables given m copies of a quantum state. Despite their similarities, performing private measurements on a single state and privately estimating the expected value of these measurements given many copies are inherently different tasks. In principle, we could perform an ε-DP measurement on each copy and then average the results. Then the overall algorithm satisfies (ε', δ')-DP with $\varepsilon' \approx \varepsilon\sqrt{m\log(1/\delta')}$ by advanced composition (Theorem 6 in [ZY17b]).

However, this approach is highly suboptimal as the privacy loss (i.e. the parameter ε) grows as $\sqrt{m}$. We present here a simpler and more efficient approach based on the concentration of measure, whose privacy loss decreases as m increases. Given an observable O and set of quantum states equipped with a relationship denoted as $\overset{Q}{\sim}$, we'll define the *average quantum sensitivity* of O as follows:

$$\overline{\Delta}(O) = \max_{\rho \overset{Q}{\sim} \sigma} \mathrm{Tr}\{O(\rho - \sigma)\}. \tag{8.116}$$

Notably, we will present a simple technique whose privacy loss is proportional to $\overline{\Delta}(O) + \sqrt{1/m}$. This newly defined quantity is closely related to other notions introduced in prior work. Remark that the Lipschitz constant [DPMTL21b] can be recovered as a special case by considering as Q-neighboring the states with W_1 distance at most one, i.e. $\rho \overset{Q}{\sim} \sigma \iff W_1(\rho,\sigma) = 1$. Moreover, if a quantum encoding $\rho(\cdot)$ is Q-neighboring-preserving, then the above can be related to the classical definition of sensitivity introduced in Equation 7.4. Consider the function $f(x) = \mathrm{Tr}\{O\rho(x)\}$, then

$$\Delta_f = \max_{x \sim x'} |f(x) - f(x')| \leq \max_{\rho(x) \overset{Q}{\sim} \rho(x')} |\mathrm{Tr}(O(\rho(x) - \rho(x'))| \leq \max_{\rho \overset{Q}{\sim} \sigma} |\mathrm{Tr}(O(\rho - \sigma))| := \overline{\Delta}(O) \tag{8.117}$$

We now prove that there exists a simple differentially private algorithm consisting of measurements and classical post-processing that gives a suitable tradeoff between sensitivity and privacy. We first consider a general post-processing channel and then we provide more concrete bounds for the Laplace and Gaussian noises.

Theorem 8.9. *Consider a neighboring relationship $\overset{Q}{\sim}$ over the set of quantum states $\mathcal{S}_n$. Let $\rho^{\otimes m}$ be a collection of m copies of a quantum state $\rho \in S_n$ and O an observable. Let $\Lambda(\cdot)$ be a classical channel with the following property. For $\delta' \in (0,1]$ and $x, x' \in \mathbb{R}$,*

$$|x - x'| \leq \overline{\Delta}(O) + \sqrt{m^{-1}\log(4/\delta')} \implies E_{e^\varepsilon}(\Lambda(x) \| \Lambda(x')) \leq \delta. \tag{8.118}$$

Consider the following algorithm $\mathscr{A}$:

1. *Measure O on each copy of ρ and collect the outcomes $y_1, \ldots, y_m$.*
2. *Compute the average $\hat{\mu} = \frac{1}{m}\sum_{i=1}^m y_i$ and output $\Lambda(\hat{\mu})$.*

Then the algorithm $\mathscr{A}$ is $(\varepsilon, \delta+\delta')$-DP.

Proof. Consider two neighboring quantum states $\rho \overset{Q}{\sim} \sigma$. For $X \in \{\rho, \sigma\}$, let $\hat{\mu}_X$ the average obtained on input $X^{\otimes m}$. By Chernoff-Hoeffding's bound,

$$\Pr\left[\left|\hat{\mu}_X - \mathrm{Tr}[OX]\right| \geq \frac{t}{2}\right] \leq 2e^{-mt^2}. \tag{8.119}$$

Hence, by union bound,

$$\Pr[E] \leq \delta' := 4e^{-mt^2}, \tag{8.120}$$

where E is the following event:

$$E := \left\{\left(\left|\hat{\mu}_\rho - \mathrm{Tr}[O\rho]\right| \geq \frac{t}{2}\right) \vee \left(|\hat{\mu}_\sigma - \mathrm{Tr}[O\sigma]| \geq \frac{t}{2}\right)\right\}. \tag{8.121}$$

Conditioning on the complementary event $\overline{E}$ and observing that $t = \sqrt{m^{-1}\log(4/\delta')}$, we have,

$$|\hat{\mu}_\rho - \hat{\mu}_\sigma| \leq |\hat{\mu}_\rho - \mathrm{Tr}[O\rho]| + |\mathrm{Tr}[O\rho] - \mathrm{Tr}[O\sigma]| + |\mathrm{Tr}[O\sigma] - \hat{\mu}_\sigma| \tag{8.122}$$

$$\leq \Delta + t = \Delta + \sqrt{m^{-1}\log(4/\delta')}. \tag{8.123}$$

This implies that, conditioning on $\overline{E}$,

$$E_{e^\varepsilon}(\Lambda(\hat{\mu}_\rho) \| \Lambda(\hat{\mu}_\sigma)) \leq \delta, \tag{8.124}$$

equivalently, we have

$$\forall S : \Pr[\Lambda(\hat{\mu}_\rho) \in S|\overline{E}] \leq e^\varepsilon \Pr[\Lambda(\hat{\mu}_\sigma) \in S|\overline{E}] + \delta. \tag{8.125}$$

Then we also have that, for all S

$$\Pr[\Lambda(\hat{\mu}_\rho) \in S] = \Pr[\Lambda(\hat{\mu}_\rho) \in S|E]\Pr[E] + \Pr[\Lambda(\hat{\mu}_\rho) \in S|\overline{E}]\Pr[\overline{E}] \tag{8.126}$$

$$\leq \Pr[\Lambda(\hat{\mu}_\rho) \in S|\overline{E}] + \delta' \leq e^\varepsilon \Pr[\Lambda(\hat{\mu}_\sigma) \in S|\overline{E}] + \delta + \delta' \tag{8.127}$$

$$\leq e^\varepsilon \Pr[\Lambda(\hat{\mu}_\sigma) \in S] + \delta + \delta'. \tag{8.128}$$

■

Finally, plugging the Laplace and the Gaussian channels in Theorem 8.9, we obtain the following corollary.

Corollary 8.6. *Let $\mathscr{A}$, $\rho^{\otimes m}$ and O as in Theorem 8.9 and let $\Delta := \overline{\Delta}(O)$. The following privacy guarantees hold.*

- *(Laplace noise) Let $\Lambda_{\mathscr{L},b}$ the Laplace channel of scale $b := (\Delta + \sqrt{m^{-1}\log(4/\delta')})/\varepsilon$. Then the algorithm $\mathscr{A}$ is (ε, δ')-DP.*
- *(Gaussian noise) Let $\Lambda_{\mathscr{G},\sigma}$ the Gaussian channel of variance*

$$\sigma^2 \geq 2\log(1.25/\delta)(\Delta + \sqrt{m^{-1}\log(4/\delta')})^2/\varepsilon^2. \tag{8.129}$$

Then the algorithm $\mathscr{A}$ is $(\varepsilon, \delta+\delta')$-DP.

Table 8.2: Here we summarize the results of Section 8.8.1. For each neighboring relationship over quantum states, we list the corresponding average quantum sensitivity $\Delta(O)$ of an observable O.

$\rho \overset{Q}{\sim} \sigma$	$\Delta(O)$
$\|\rho-\sigma\|_p \leq \tau$	$\tau\|O\|_q$
$\frac{1}{2}\|\rho-\sigma\|_1 \leq \tau$	$\tau\|O\|_1$
$W_1(\rho,\sigma) \leq \tau$	$\|O\|_{Lip}\tau$
$\rho \overset{(\Xi,\tau)}{\sim} \sigma$	$\min\{\frac{3}{2}\|O\|_{Lip}\max_{\mathscr{I}\in\Xi}\|\mathscr{I}\|\tau, \|O\|_{Lip}n\tau\}$

8.8.1 Bounding the average quantum sensitivity

Here we provide several bounds for the quantum sensitivity based on different neighboring relationships. The first bound is based on Hölder's inequality, i.e. $|\mathrm{Tr}(LR)| \leq \|L\|_p\|R\|_q$ for $p^{-1}+q^{-1}=1$, where $\|\cdot\|_p$ is the Schatten p-norm. Say that $\rho \overset{Q}{\sim} \sigma$ if $\|\rho-\sigma\|_p \leq \tau$. Then applying Hölder's inequality yields

$$\Delta(O) \leq \|O\|_q\tau. \tag{8.130}$$

For the special case of $p=1$ (which corresponds to the trace distance) a stronger bound holds:

$$\Delta(O) = \max_{\rho,\sigma:\|\rho-\sigma\|_1\leq\tau} \mathrm{Tr}[O(\rho-\sigma)] \leq \frac{1}{2}\|O\|_\infty\|\rho-\sigma\|_1 \leq \frac{\tau}{2}\|O\|_\infty. \tag{8.131}$$

We can also consider a neighboring relationship based on the Wasserstein distance of order 1, i.e. $\rho \overset{Q}{\sim} \sigma$ if $W_1(\rho,\sigma)\leq\tau$. Then the quantum sensitivity is proportional to the Lipschitz constant.

$$\Delta(O) = \max_{\rho,\sigma:W_1(\rho,\sigma)\leq\tau} \mathrm{Tr}\{O(\rho-\sigma)\} \leq \|O\|_{Lip}\tau. \tag{8.132}$$

By Lemma 3.3, we also have that if $\rho \overset{(\Xi,\tau)}{\sim} \sigma$, then $W_1(\rho,\sigma) \leq \frac{3}{2}\max_{\mathscr{I}\in\Xi}|\mathscr{I}|\tau$. This implies

$$\Delta(O) = \max_{\rho,\sigma:\rho\overset{(\Xi,\tau)}{\sim}\sigma} \mathrm{Tr}\{O(\rho-\sigma)\} \leq \frac{3}{2}\|O\|_{Lip}\max_{\mathscr{I}\in\Xi}|\mathscr{I}|\tau. \tag{8.133}$$

The above bounds for $\Delta(O)$ are listed concisely in Table 8.2.

8.9 Private quantum machine learning

In this section, we demonstrate the applications of the results and tools we derived so far to variational quantum algorithms for machine learning. Let $\rho(\boldsymbol{\theta};\boldsymbol{x})$ be the output of a variational quantum circuit. We will assume that the parameters $\boldsymbol{\theta}$ are trained using a suitable (classical) dataset $S=(\boldsymbol{s^{(1)}},\ldots,\boldsymbol{s^{(m)}})$. Given a test set $\mathscr{X}$, we're asked to approximate a function $f:\mathbb{R}^d\to\mathbb{R}$. Thus, we can use variational quantum algorithms to find a set of parameters $\boldsymbol{\theta}$ that satisfy

$$\forall \boldsymbol{x}\in\mathscr{X} : f(\boldsymbol{x}) \simeq \mathrm{Tr}(O\rho(\boldsymbol{\theta};\boldsymbol{x})), \tag{8.134}$$

where O is a suitable observable. Given this simple scenario, differential privacy can come in different flavours.

- Let $\boldsymbol{x} = (x_1, \dots, x_d) \in \mathscr{X}$ be the input vector. Given a neighboring relationship $\boldsymbol{x} \sim \boldsymbol{x}'$, we can ensure differential privacy with respect to the input $\boldsymbol{x}$. This is particularly useful when $\boldsymbol{x}$ contains the sensitive information of multiple individuals or when $\boldsymbol{x}$ might be corrupted by an *adversarial attack*.
- In the alternative, we can require differential privacy with respect to the training set $S = (\boldsymbol{s}^{(1)}, \dots, \boldsymbol{s}^{(m)})$, where $S \sim S'$ if they differ only in a single entry $\boldsymbol{s}^{(j)}$. This notion of privacy is meant to protect the sensitive information of the individuals who compose the training set. Furthermore, it also enhances *generalisation*, i.e. it allows to upper bound of the discrepancy between the error on the training set and the generalisation error.

8.9.1 Private evaluation with respect to the input $\boldsymbol{x}$

Given a suitable notion of neighboring inputs $\boldsymbol{x} \sim \boldsymbol{x}'$, we want to find a neighboring relationship over quantum states $\overset{Q}{\sim}$ such that $\rho(\cdot, \boldsymbol{\theta})$ is Q-neighboring-preserving. In other terms, we need to ensure that

$$\boldsymbol{x} \sim \boldsymbol{x}' \implies \rho(\boldsymbol{x}, \boldsymbol{\theta}) \overset{Q}{\sim} \rho(\boldsymbol{x}', \boldsymbol{\theta}). \tag{8.135}$$

First, we select the relationship Q according to Table 8.1. If a single copy of $\rho(\boldsymbol{x}, \boldsymbol{\theta})$ is available, we can make the measurement differentially private either by adding a final quantum noisy channel (Theorem 8.4) or by classical post-processing (Theorem 8.5). If, instead, we're able to prepare multiple copies of $\rho(\boldsymbol{x}, \boldsymbol{\theta})$, it's convenient to post-process the average outcome with classical noise. Then differential privacy is guaranteed by Corollary 8.6.

Numerical results. Finally, we complement our theoretical analysis with a numerical simulation implemented in PennyLane. We consider a classification task based on the first two classes of the famous IRIS dataset and each input $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$ is susceptible to be perturbed by an adversarial attack. We assume that the adversary can select a single entry x_i and map it to x_i' with $|x_i - x_i'| \leq \tau$, for some threshold $0 \leq \tau \leq 1$. We trained a simple 4-qubit binary classifier, based on the variational circuit depicted in Figure 8.4, whose gates are parametrised by a trainable vector $\boldsymbol{\theta}$ and the input vector $\boldsymbol{x}$. Hence, the output is measured according to $O = \frac{1}{8}\sum_{i=1}^{4}(Z_i + 1)$ and the classifier outputs 0 if the outcome is larger than 0.5 and 1 otherwise. It's easy to see that this encoding is $(1, \tau)$-privacy-preserving with respect to the neighboring definition induced by the adversarial attack. The circuit is ended by a final layer of local depolarising noise $\mathscr{N}_p^{\otimes n}$, which ensures (ε, δ_1)-differential privacy with respect to $(1, \tau)$-neighboring states, with δ_1 defined as in Theorem 8.4. We trained the model with the `Adam` optimiser [KB14] with several noise levels p and then we used the test set to estimate the certified accuracy for each p, and we plotted it against the threshold τ in Figure 8.5. The results show that the noise level should be set according to attack threshold τ, as for $\tau \leq 0.2$ the circuit with $p = 0.1$ outperforms the others, while for $\tau \geq 0.2$ the circuit with $p = 0.3$ achieves the best certified accuracy.

Our simulation differs from previous experiments in multiple ways. First, we remark that our simulation combines local noisy channels with the novel neighboring relationship we introduced in

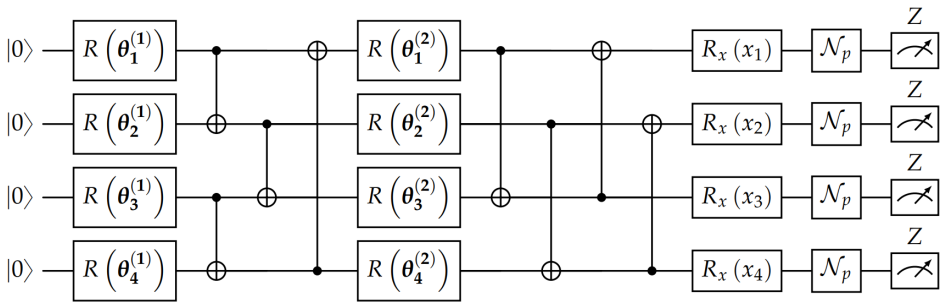


Figure 8.4: The parametric quantum circuit used in the simulation. We placed the encoding gates after the trainable gates in order to produce a $(1,\tau)$-neighboring-preserving encoding. The output state is measured according to the observable $O = \frac{1}{8}\sum_{i=1}^{4}(Z_i + 1)$.

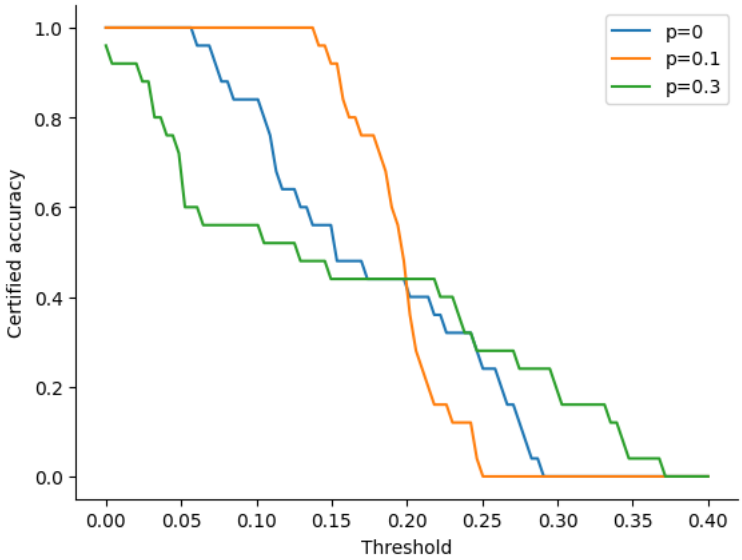


Figure 8.5: This plot contains the values of the certified accuracy estimated for various noise levels p and various attack thresholds τ.

the present Chapter. In contrast to this, the simulation in [DHL+21b] is based on τ-neighboring states and ensures privacy via multiple layers of *global* depolarizing noise. On the other hand, [HTY+23] combines local noisy channels with τ-neighboring states, resulting in privacy guarantees that degrade exponentially fast as the number of qubits increases. This stems from the fact that in Lemma 3 in [HTY+23], the authors show quantum differential privacy with $\varepsilon = \log(1 + \tau/p^n) \simeq \tau/p^n$. In addition, both [DHL+21b] and [HTY+23] are based on ε-differential privacy while Proposition 7.3 is stated in terms of (ε,δ)-differential privacy. This is particularly useful to assess the certified accuracy of various noise regimes, including the case with no noise at all ($p = 0$).

8.9.2 Private prediction with respect to the training set S

Training a variational quantum algorithm involves finding a set of parameters $\boldsymbol{\theta}*$ that minimizes a loss function $\mathscr{L}(\boldsymbol{\theta}, S) = \frac{1}{m}\sum_{i=1}^{m} \mathrm{Tr}\{O(y_i)\rho(\boldsymbol{\theta}; \boldsymbol{x_i})\} = \frac{1}{m}\sum_{i=1}^{m} \ell(\boldsymbol{\theta}, \boldsymbol{s_i})$ with respect to a given training set $S = (\boldsymbol{s_1}, \ldots, \boldsymbol{s_m})$ where $\boldsymbol{s_i} = (\boldsymbol{x_i}, y_i)$. In this setting, we let S and S' be neighboring if $\exists i \in [m], \forall j \neq i : \boldsymbol{s_j} = \boldsymbol{s_j}$, i.e. if they differ in at most one element. Despite the existence of quantum algorithms for optimising a loss function, they're often not suitable for near-term devices. In most near-term applications, a variational quantum circuit is paired with a classical optimiser. Thus, standard techniques for differentially private (classical) optimisation can be adapted [BST14, ACG+16]. For instance, [WCY23] implements the algorithm for private stochastic gradient descent (SGD) provided in [ACG+16] to optimize the parameters of a variational quantum circuit, achieving good empirical performance. The technique provided in [ACG+16] involves a procedure known as *gradient clipping*, which consists in rescaling the gradient $\nabla_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}, \boldsymbol{s}_i)$ to ensure that its ℓ_2 norm is bounded by a suitable constant C, i.e. $\|\nabla_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}, \boldsymbol{s}_i)\|_2 \leq C$. Then, privacy is ensured by the addition of Gaussian noise with variance proportional to C^2 on each estimate of the gradient. Instead of clipping the gradient, alternative techniques such as [BST14], estimates an upper bounds UB, where

$$\forall \boldsymbol{\theta} : \|\nabla_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}, \boldsymbol{s_i})\|_2 \leq UB. \tag{8.136}$$

and add Gaussian noise proportional to UB^2 on each estimate of the gradient.

Here we show that UB can be easily estimated for some classes of variational quantum circuits. Assuming ℓ is differentiable with respect to $\boldsymbol{\theta}$ we have

$$|\ell(\boldsymbol{\theta}, \boldsymbol{s_i}) - \ell(\boldsymbol{\theta'}, \boldsymbol{s_i})| \leq UB \|\boldsymbol{\theta} - \boldsymbol{\theta'}\|_{\ell_2} \implies \|\nabla_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}, \boldsymbol{s_i})\|_2 \leq UB. \tag{8.137}$$

For $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)$, assume that each coordinate θ_j is encoded via a single gate Hamiltonian encoding, i.e. $e^{-i\theta_j H_j}$ with $\|H_i\|_2 \leq 1$. Moreover, assume that the output state is produced by a $1D$ circuit with bounded depth L (and thus the light-cone of each single qubit gate is upper bounded by $2L$). As shown in Appendix 11.2, the Hamiltonian encoding $\rho(\cdot, \boldsymbol{s}_i)$ is (Ξ, τ)-neighboring-preserving, where

$$\tau \leq \sqrt{\frac{d}{2}} \|\boldsymbol{\theta} - \boldsymbol{\theta'}\|_2 \quad \text{and} \quad \max_{\mathscr{I} \in \Xi} |\mathscr{I}| \leq 2L. \tag{8.138}$$

Hence, we have

$$|\ell(\boldsymbol{\theta}, \boldsymbol{s_i}) - \ell(\boldsymbol{\theta'}, \boldsymbol{s_i})| \leq |\mathrm{Tr}\{O(y_i)\rho(\boldsymbol{\theta}; \boldsymbol{x_i}) - \mathrm{Tr}\{O(y_i)\rho(\boldsymbol{\theta'}; \boldsymbol{x_i})\}| \tag{8.139}$$

$$\leq \|O(y_i)\|_{Lip} W_1(\rho(\boldsymbol{\theta}; \boldsymbol{x_i}), \rho(\boldsymbol{\theta'}; \boldsymbol{x_i})) \leq 3L\sqrt{\frac{d}{2}} \|O(y_i)\|_{Lip} \|\boldsymbol{\theta} - \boldsymbol{\theta'}\|_2. \tag{8.140}$$

And then

$$\forall \boldsymbol{\theta} : \|\nabla_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}, \boldsymbol{s_i})\|_2 \leq 3L\sqrt{\frac{d}{2}} \|O(y_i)\|_{Lip}. \tag{8.141}$$

CHAPTER 9

QUANTUM DIFFERENTIAL PRIVACY IN THE LOCAL MODEL

-

Locally differentially private (LDP) measurements were introduced in [AR19] and referred as *nearly trivial* measurements. Informally, the output of a LDP measurement weakly depends on the input state, and this is often ensured by the injection of noise. This comes with desirable privacy guarantees, along with an increased sample complexity for many computational tasks. Considering the detrimental impact of noise on quantum algorithms, the following question naturally arises.

Question 4. *Can we attain an exponential quantum speed-up under the stringent constraint of local differential privacy?*

Throughout this Chapter we will argue that certain computational tasks are unfeasible under this strict notion of privacy, while others can be efficiently performed. Specifically, we will demonstrate that local differential privacy is compatible with exponential quantum speed-up for specific tasks.

Our contributions. Our first set of contributions consists in several entropic inequalities for locally differentially private channels (Section 9.1). In particular, we provide a strong data processing inequality for the quantum relative entropy under locally differentially private measurements. In Section 9.2, we provide a quantum version of the equivalence between learning under local differential privacy

and statistical query learning, answering an open question posed by [AQS21]. As a corollary, we also obtain an exponential separation between learning under quantum local differential privacy and learning with separable measurements, resolving an open question posed by [AR19]. Furthermore, in Section 9.3.1, we provide an application of the aforementioned entropic inequalities to the task of asymmetric hypothesis testing with restricted measurements. Our result is a quantum analogue of the private Stein's lemma ([AAC21a], Corollary 4). Finally, in Section 9.3.2 we investigate the problem of learning from quantum data in a distributed setting under local differential privacy. We demonstrate that parity functions are efficiently learnable in this model, whereas the corresponding classical task requires exponentially many samples [KLN$^+$11a].

Related work. The task of quantum hypothesis testing under local differential privacy has also been recently explored in a simultaneous work by [HT23]. We emphasize that Theorem 9.3 provides a quadratic improvement over ([HT23], Corollary 5.14) for small values of the privacy level ε. It is also worth noting that the results in [HT23] extend beyond measurements to encompass private quantum channels.

9.1 Entropic inequalities under local privacy

A crucial fact in quantum information theory is that many physical quantities are monotone under the application of a quantum channel. For instance, the quantum relative entropy satisfies the following *data-processing inequality* (DPI), for all states ρ, σ and for every channel $\mathcal{N}$:

$$D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \leq D(\rho\|\sigma). \tag{9.1}$$

Furthermore, the same property is shared by the hockey-stick divergences, and in particular by the trace distance. When the inequality is strict, we say that a given divergence satisfies a *strong* data-processing inequality (SDPI) with respect to the channel $\mathcal{N}$. We also define the following contraction coefficients, also previously considered in [LR99, HR16, HRF22, HRF23].

$$\eta(\mathcal{N}) := \sup_{\rho,\sigma} \frac{D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))}{D(\rho\|\sigma)} \quad \text{and} \quad \eta_\gamma(\mathcal{N}) := \sup_{\rho,\sigma} \frac{E_\gamma(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))}{E_\gamma(\rho\|\sigma)}. \tag{9.2}$$

where $\gamma \geq 1$. Recall that $E_1(\rho\|\sigma) = \|\rho - \sigma\|_{\mathrm{tr}}$ and hence $\eta_1(\mathcal{N})$ is the contraction coefficient for the trace distance. If $\mathcal{N}$ satisfies (ε, δ)-LDP, then its contraction coefficient η_γ can be upper bounded as follows ([HRF23], Theorem II.2 and Corollary V.1):

$$\eta_{e^\varepsilon}(\mathcal{N}) \leq \delta \text{ and } \eta_\gamma(\mathcal{N}) \leq \varphi(\varepsilon, \delta), \tag{9.3}$$

where $\varphi(\varepsilon, \delta) := 1 - e^{-\varepsilon}(1 - \delta)$. More broadly, we can also consider inequalities involving two distinct divergences. For instance, every ε-LDP channel $\mathcal{N}$ satisfies:

$$D_M(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \leq 2\varepsilon\|\mathcal{N}(\rho) - \mathcal{N}(\sigma)\|_{\mathrm{tr}} \leq 2\varepsilon(1 - e^{-\varepsilon})\|\rho - \sigma\|_{\mathrm{tr}}, \tag{9.4}$$

where the first inequality is due to ([HRF23], Lemma III.6) and the second inequality follows from Eq. 9.3. We will now prove an analogous result, where the measured relative entropy is replaced by the quantum relative entropy.

Proposition 9.1. *For all states ρ,σ we have*

$$D\left(\rho\|\sigma\right)+D\left(\sigma\|\rho\right)\leq\left[D_{\max}(\rho\|\sigma)+D_{\max}(\sigma\|\rho)\right]\|\rho-\sigma\|_{\mathrm{tr}} \tag{9.5}$$

Proof. Recall that we can write the decomposition $\rho-\sigma=X^{+}-X^{-}$, where X^{+} and X^{-} denote respectively the positive part and the negative part of $\rho-\sigma$. We start by rearranging the expression of the quantum relative entropy as follows

$$D\left(\rho\|\sigma\right)+D\left(\sigma\|\rho\right)=\mathrm{Tr}\left[\rho\left(\log\rho-\log\sigma\right)\right]+\mathrm{Tr}\left[\sigma\left(\log\sigma-\log\rho\right)\right] \tag{9.6}$$

$$=\mathrm{Tr}\left[\left(\rho-\sigma\right)\left(\log\rho-\log\sigma\right)\right]=\mathrm{Tr}\left[\left(X^{+}-X^{-}\right)\left(\log\rho-\log\sigma\right)\right] \tag{9.7}$$

$$=\mathrm{Tr}\left[X^{+}\left(\log\rho-\log\sigma\right)\right]+\mathrm{Tr}\left[X^{-}\left(\log\sigma-\log\rho\right)\right]. \tag{9.8}$$

By definition of max-relative entropy, $\rho\leq e^{D_{\max}(\rho\|\sigma)}\sigma$. Since the logarithm is an operator monotone function, we have that $\log\rho\leq\log\left(e^{D_{\max}(\rho\|\sigma)}\sigma\right)=D_{\max}(\rho\|\sigma)I+\log\sigma$. Similarly, we also have $\log\sigma\leq D_{\max}(\sigma\|\rho)I+\log\rho$. Putting all together, we obtain

$$D\left(\rho\|\sigma\right)+D\left(\sigma\|\rho\right)\leq D_{\max}(\rho\|\sigma)\mathrm{Tr}\left[X^{+}\right]+D_{\max}(\sigma\|\rho)\mathrm{Tr}\left[X^{-}\right] \tag{9.9}$$

$$=\left[D_{\max}(\rho\|\sigma)+D_{\max}(\sigma\|\rho)\right]\|\rho-\sigma\|_{\mathrm{tr}}, \tag{9.10}$$

where the equality follows from $\mathrm{Tr}\left[X^{+}\right]=\mathrm{Tr}\left[X^{-}\right]=\|\rho-\sigma\|_{\mathrm{tr}}$. ■

We remark that an analogous result has also been recently presented in ([HT23], Eqs. 5.25-27). However, in [HT23] the sum $D\left(\rho\|\sigma\right)+D\left(\sigma\|\rho\right)$ is replaced by $D\left(\rho\|\sigma\right)$. Thus, our result is tighter of a factor 2 when the goal is to upper bound the sum $D\left(\rho\|\sigma\right)+D\left(\sigma\|\rho\right)$.

A simple application of Eq. 9.3 to Proposition 9.1 yields the following corollary, which generalizes Eq. 9.4.

Corollary 9.1. *Let $\mathcal{N}$ an ε-LDP channel. Then for all states ρ,σ we have*

$$D\left(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)\right)+D\left(\mathcal{N}(\sigma)\|\mathcal{N}(\rho)\right)\leq 2\varepsilon(1-e^{-\varepsilon})\|\rho-\sigma\|_{\mathrm{tr}}. \tag{9.11}$$

We now derive yet another improved version of Eq. 9.4, by generalizing Lemma 1 in [DJW13] to the quantum setting.

Lemma 9.1. *Let $\mathcal{M}=\{\mathcal{M}_x\}_{x\in X}$ be an ε-LDP POVM measurement. Then for all states ρ,σ we have*

$$D\left(\mathcal{M}(\rho)\|\mathcal{M}(\sigma)\right)+D\left(\mathcal{M}(\sigma)\|\mathcal{M}(\rho)\right)\leq e^{\varepsilon}(1-e^{-\varepsilon})^{2}\|\rho-\sigma\|_{\mathrm{tr}}^{2}. \tag{9.12}$$

Moreover, for every ε-LDP channel $\mathcal{N}$ and for all states ρ,σ,

$$D_{M}\left(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)\right)\leq e^{\varepsilon}(1-e^{-\varepsilon})^{2}\|\rho-\sigma\|_{\mathrm{tr}}^{2}, \tag{9.13}$$

where $D_{M}(\cdot\|\cdot)$ is the measured relative entropy.

Proof. Let $p_x = \mathrm{Tr}(\mathscr{M}_x\rho)$ and $q_x = \mathrm{Tr}(\mathscr{M}_x\sigma)$.

$$D(\mathscr{M}(\rho)\|\mathscr{M}(\sigma)) + D(\mathscr{M}(\sigma)\|\mathscr{M}(\rho)) \tag{9.14}$$

$$= \sum_x p_x \log\frac{p_x}{q_x} + \sum_x q_x \log\frac{q_x}{p_x} = \sum_x (p_x - q_x)\log\frac{p_x}{q_x}. \tag{9.15}$$

We want to upper bound $|p_x - q_x| = |\mathrm{Tr}(\mathscr{M}_x(\rho-\sigma))|$. Let $\rho-\sigma = X^+ - X^-$, where X^+ and X^- denote respectively the positive part and the negative part of $\rho-\sigma$. We can also write the spectral decompositions $X^+ = \sum_{y\in\mathscr{Y}} \lambda_y |y\rangle\langle y|$ and $X^- = \sum_{z\in\mathscr{Z}} \tau_z |z\rangle\langle z|$. First, we upper bound $p_x - q_x = \mathrm{Tr}(\mathscr{M}_x(\rho-\sigma))$

$$\mathrm{Tr}(\mathscr{M}_x(X^+ - X^-)) = \mathrm{Tr}\left(\mathscr{M}_x\left(\sum_{y\in\mathscr{Y}} \lambda_y |y\rangle\langle y|\right)\right) - \mathrm{Tr}\left(\mathscr{M}_x\left(\sum_{z\in\mathscr{Z}} \tau_z |z\rangle\langle z|\right)\right) \tag{9.16}$$

$$\leq \max_{y\in\mathscr{Y}} \mathrm{Tr}(\mathscr{M}_x |y\rangle\langle y|)\left(\sum_{y\in\mathscr{Y}} \lambda_y\right) - \min_{z\in\mathscr{Z}} \mathrm{Tr}(M_x |z\rangle\langle z|)\left(\sum_{z\in\mathscr{Z}} \tau_z\right) \tag{9.17}$$

$$= \|\rho-\sigma\|_{\mathrm{tr}}\left(\max_{y\in\mathscr{Y}} \mathrm{Tr}(\mathscr{M}_x |y\rangle\langle y|) - \min_{z\in\mathscr{Z}} \mathrm{Tr}(\mathscr{M}_x |z\rangle\langle z|)\right) \tag{9.18}$$

$$\leq \|\rho-\sigma\|_{\mathrm{tr}} \max_{y\in\mathscr{Y}} \mathrm{Tr}(\mathscr{M}_x |y\rangle\langle y|)(1-e^{-\varepsilon}), \tag{9.19}$$

where the second equality follows from the identities $\mathrm{Tr}[X^+] = \sum_{y\in\mathscr{Y}} \lambda_y = \|\rho-\sigma\|_{\mathrm{tr}}$ and $\mathrm{Tr}[X^-] = \sum_{z\in\mathscr{Z}} \tau_z = \|\rho-\sigma\|_{\mathrm{tr}}$, and the last inequality follows from ε-LDP. Proceeding in an analogous way, we derive the following lower bound.

$$\mathrm{Tr}(\mathscr{M}_x(X^+ - X^-)) \geq \min_{y\in\mathscr{Y}} \mathrm{Tr}(\mathscr{M}_x |y\rangle\langle y|)\left(\sum_{y\in\mathscr{Y}} \lambda_y\right) - \max_{z\in\mathscr{Z}} \mathrm{Tr}(\mathscr{M}_x |z\rangle\langle z|)\left(\sum_{z\in\mathscr{Z}} \tau_z\right) \tag{9.20}$$

$$= \|\rho-\sigma\|_{\mathrm{tr}}\left(\min_{y\in\mathscr{Y}} \mathrm{Tr}(\mathscr{M}_x |y\rangle\langle y|) - \max_{z\in\mathscr{Z}} \mathrm{Tr}(\mathscr{M}_x |z\rangle\langle z|)\right) \tag{9.21}$$

$$\geq \|\rho-\sigma\|_{\mathrm{tr}} \max_{z\in\mathscr{Z}} \mathrm{Tr}(\mathscr{M}_x |z\rangle\langle z|)(e^{-\varepsilon}-1), \tag{9.22}$$

where we applied again the identities $\mathrm{Tr}[X^+] = \mathrm{Tr}[X^-] = \|\rho-\sigma\|_{\mathrm{tr}}$ and ε-LDP. We can now provide an upper bound for $|\mathrm{Tr}(\mathscr{M}_x(\rho-\sigma))|$:

$$|\mathrm{Tr}(\mathscr{M}_x(\rho-\sigma))| = |\mathrm{Tr}(\mathscr{M}_x(X^+ - X^-)| = \max\{\mathrm{Tr}(\mathscr{M}_x(X^+ - X^-), \mathrm{Tr}(\mathscr{M}_x(X^- - X^+)\} \tag{9.23}$$

$$\leq \|\rho-\sigma\|_{\mathrm{tr}}(1-e^{-\varepsilon}) \max_{y\in\mathscr{Y}\cup\mathscr{Z}} \mathrm{Tr}(\mathscr{M}_x |y\rangle\langle y|), \tag{9.24}$$

Recall that, for $a, b \in \mathbb{R}_+$ ([DJW13], Lemma 4),

$$\log\frac{a}{b} \leq \frac{|a-b|}{\min\{a,b\}} \tag{9.25}$$

Thus,

$$\log\frac{p_x}{q_x} \leq \frac{|p_x - q_x|}{\min\{p_x, q_x\}} = \frac{|\mathrm{Tr}(\mathscr{M}_x(\rho-\sigma))|}{\min\{\mathrm{Tr}(\mathscr{M}_x\rho), \mathrm{Tr}(\mathscr{M}_x\sigma)\}} \tag{9.26}$$

$$\leq \frac{\|\rho-\sigma\|_{\mathrm{tr}}(1-e^{-\varepsilon}) \max_{y\in\mathscr{Y}\cup\mathscr{Z}} \mathrm{Tr}(\mathscr{M}_x |y\rangle\langle y|)}{\min\{\mathrm{Tr}(\mathscr{M}_x\rho), \mathrm{Tr}(\mathscr{M}_x\sigma)\}} \leq e^{\varepsilon}(1-e^{-\varepsilon})\|\rho-\sigma\|_{\mathrm{tr}}, \tag{9.27}$$

where we applied ε-LDP in the last inequality. Putting all together,

$$D(\mathscr{M}(\rho)\|\mathscr{M}(\sigma)) + D(\mathscr{M}(\sigma)\|\mathscr{M}(\rho)) \tag{9.28}$$

$$\leq e^{\varepsilon}(1-e^{-\varepsilon})\|\rho-\sigma\|_{\mathrm{tr}}\left(\|\rho-\sigma\|_{\mathrm{tr}}(1-e^{-\varepsilon})\max_{y\in\mathscr{Y}\cup\mathscr{Z}}\mathrm{Tr}(\mathscr{M}_x|y\rangle\langle y|)\right) \tag{9.29}$$

$$\leq e^{\varepsilon}(1-e^{-\varepsilon})^2\|\rho-\sigma\|_{\mathrm{tr}}^2, \tag{9.30}$$

where the last inequality follows from $\max_{y\in\mathscr{Y}\cup\mathscr{Z}}\mathrm{Tr}(\mathscr{M}_x|y\rangle\langle y|)\leq 1$. We proved the first part of the lemma. As for the second part, let $\widehat{\mathscr{M}}$ the POVM measurement that maximizes $D(\mathscr{N}(\widehat{\mathscr{M}}(\rho))\|\mathscr{N}(\widehat{\mathscr{M}}(\sigma)))$. Then the desired result follow from the definition of measured relative entropy.

$$D_M(\mathscr{N}(\rho)\|\mathscr{N}(\sigma)) = D(\mathscr{N}(\widehat{\mathscr{M}}(\rho))\|\mathscr{N}(\widehat{\mathscr{M}}(\sigma))) \tag{9.31}$$

$$\leq D(\mathscr{N}(\widehat{\mathscr{M}}(\rho))\|\mathscr{N}(\widehat{\mathscr{M}}(\sigma))) + D(\mathscr{N}(\widehat{\mathscr{M}}(\sigma))\|\mathscr{N}(\widehat{\mathscr{M}}(\rho))) \tag{9.32}$$

$$\leq e^{\varepsilon}(1-e^{-\varepsilon})^2\|\rho-\sigma\|_{\mathrm{tr}}^2. \tag{9.33}$$

■

We observe that, for small values of ε, Lemma 9.1 is quadratically tighter in $\|\rho-\sigma\|_{\mathrm{tr}}$ with respect to Eq. 9.4. A simple application of the "measured" Pinsker's inequality (Lemma 3.4) to Lemma 9.1 yields the following corollary.

Corollary 9.2. *Let $\mathscr{M}=\{\mathscr{M}_x\}_{x\in X}$ be an ε-LDP POVM measurement. Then for all states ρ,σ we have*

$$D\left(\mathscr{M}(\rho)\|\mathscr{M}(\sigma)\right) + D\left(\mathscr{M}(\sigma)\|\mathscr{M}(\rho)\right) \leq \frac{e^{\varepsilon}}{2}(1-e^{-\varepsilon})^2 D_M(\rho\|\sigma), \tag{9.34}$$

where $D_M(\cdot\|\cdot)$ is the measured relative entropy. Moreover, for every ε-LDP channel $\mathscr{N}$ and for all states ρ,σ,

$$D_M\left(\mathscr{N}(\rho)\|\mathscr{N}(\sigma)\right) \leq \frac{e^{\varepsilon}}{2}(1-e^{-\varepsilon})^2 D_M(\rho\|\sigma). \tag{9.35}$$

9.2 Learning under local privacy is equivalent to QSQ learning

In this section we show an equivalence between locally differentially private measurements and quantum statistical queries, answering an open question posed in ([AQS21], Question 7). In particular, we will prove that quantum statistical queries can be efficiently simulated by differentially private measurements, and vice versa, differentially private measurements can be efficiently simulated by quantum statistical queries. The latter result is less intuitive and relies on a rejection-sampling argument. The classical analog of the equivalence was proven in the seminal paper of [KLN+11a]. Interestingly, this result readily implies an exponential separation between learning under quantum local differential privacy and learning with separable measurements, answering an open question posed in ([AR19], Question 4).

9.2.1 Simulation of QStat queries with locally differentially private measurements

We first show that quantum statistical queries can be simulated efficiently with LDP measurements. The result follows by iterating the Laplace measurement defined in Section 7.4.1 and using concentration of measure.

Theorem 9.1. *If $m \geq c \cdot \frac{\log(1/\beta)k^2}{\varepsilon^2\tau^2}$ for a sufficiently large constant c, then $\mathscr{A}_{\mathscr{M},\varepsilon}$ (Algorithm 9.2.1) approximates $\mu = \mathbb{E}[\mathscr{M}(\rho)]$ within additive error $\pm\tau$ with probability at least $1-\beta$. Moreover, each measurement performed by $\mathscr{A}_{\mathscr{M},\varepsilon}$ satisfies ε-local differential privacy.*

Proof. The proof closely follows the one of Lemma 5.6 in [KLN+11a]. Algorithm 9.2.1 implements the Laplace measurement $\mathscr{M}^{\mathrm{Lap},\varepsilon}$ on each copy of ρ and then averages the results. We first show that $\frac{1}{m}\sum_i y_i$ is concentrated around $\mu := \mathbb{E}[\mathscr{M}(\rho)]$. By the Chernoff-Hoeffding bound for real-valued variables,

$$\Pr\left[\left|\frac{1}{m}\sum_{i=1}^{m} y_i - \mu\right| \geq \frac{\tau}{2}\right] \leq 2\exp\left(-\frac{\tau^2 m}{2k^2}\right). \tag{9.36}$$

The contribution of the Laplace noise can also be bounded via a standard tail inequality. By Lemma A.3 in [KLN+11a],

$$\Pr\left[\left|\frac{1}{m}\sum_{i=1}^{m} \eta_i\right| \geq \frac{\tau}{2}\right] \leq \exp\left(-\frac{\tau^2\varepsilon^2 m}{4k^2}\right) \tag{9.37}$$

And thus by union bound,

$$\Pr[|\hat{\mu}-\mu| \geq \tau] \leq 2\exp\left(-\frac{\tau^2 m}{2k^2}\right) + \exp\left(-\frac{\tau^2\varepsilon^2 m}{4k^2}\right) \leq 3\exp\left(-\frac{\tau^2\varepsilon^2 m}{4k^2}\right), \tag{9.38}$$

where $\hat{\mu} := \frac{1}{m}\sum_{i=1}^{m}(y_i+\eta_i)$. This implies that $O\left(\frac{\log(1/\beta)k^2}{\varepsilon^2\tau^2}\right)$ samples are sufficient to ensure that $\hat{\mu}$ approximates μ within additive error $\pm\tau$ with probability at least $1-\beta$. Moreover, each Laplace measurement $\mathscr{M}^{\mathrm{Lap},\varepsilon}$ satisfies ε-local differential privacy. ■

Theorem 9.1 can be easily extended to the case where an algorithm $\mathscr{B}$ makes t queries to a QSQ oracle QStat_ρ. In order to simulate $\mathscr{B}$, it's sufficient to simulate each QStat query $(\mathscr{M},\tau)$ by running $\mathscr{A}_{\mathscr{M},\varepsilon}$ with parameters $\beta' = \beta/t$ and $m' = c\cdot\frac{\log(1/\beta')k^2}{\varepsilon^2\tau^2}$ on m' (unused) copies of ρ. Then the simulation requires $m'\cdot t$ copies and produces the same output as $\mathscr{B}$ with probability at least $1-\beta$.

The above result generalizes Theorem 6.5 in [AGY20], as this previous result shows the quantum statistical queries can be simulated by (standard) differentially private measurements. Our result holds under *local* differential privacy, which provides stronger security guarantees, and thus implies the result of [AGY20]. From a practical standpoint, the two results differ as we randomize each outcome y_i, while in [AGY20] only the final average is randomized by a single injection of Laplace noise.

Combined with the upper bounds provided in [AGY20] and [AHS23], Theorem 9.1 readily implies that a wide family of concepts is learnable from quantum examples under local differential privacy, including parities, k-juntas, DNF functions and of n-qubit trivial states, i.e. the states obtained by applying an arbitrary constant depth circuit to the initial state $|0^n\rangle$.

Algorithm 3 A quantum ε-LDP algorithm $\mathscr{A}_{\mathscr{M},\varepsilon}$ that simulates QStat_ρ

Input $\rho^{\otimes m}$, a k-ary POVM $\mathscr{M}$.
Output An estimate of $\mathbb{E}[\mathscr{M}(\rho)]$ up to additive error τ.

1. Perform the (non-private) measurement $\mathscr{M}$ on each copy of ρ and let $y_1, y_2, \ldots, y_m$ be the outcomes.
2. Sample $\eta_1, \eta_2, \ldots \eta_m$ i.i.d. from the Laplace distribution centered in 0 and with scale parameter $(k-1)/\varepsilon$.
3. Return $\hat{\mu} := \frac{1}{m}\sum_{i=1}^{m}(y_i + \eta_i)$.

9.2.2 Simulation of locally differentially private measurements with QStat queries

It remains to show that locally differentially private measurements can be simulated efficiently with quantum statistical queries. We will prove it using a rejection-sampling algorithm, along the lines of [KLN+11a].

Theorem 9.2. *Let $\mathscr{M}$ be an ε-LDP measurement. Then $\mathscr{B}_\varepsilon$ (Algorithm 9.2.2) in expectation makes $O(e^\varepsilon)$ queries to QStat_ρ with accuracy $\tau = \Theta(\beta/e^{2\varepsilon})$ and the total variation distance between $\mathscr{B}_\varepsilon$'s output distribution and $\mathscr{M}(\rho)$ is at most β.*

Proof. The proof can be readily adapted from that of its classical counterpart, as presented in ([KLN+11a], Lemma 5.8). However, for the sake of thoroughness, we choose to include the entire argument. We want to sample from a distribution $\nu(\cdot)$ that is within a small total variation distance from $p := \mathscr{M}(\rho)$. To this end, we will prove a stronger statement, by ensuring that, for all $w \in [k]$, $\nu(w)$ is a multiplicative approximation of $p(w)$. In particular, we show that:

$$\nu(w) \in (1 \pm 2\beta)p(w) \tag{9.39}$$

Note that this directly implies the following:

$$|p - \nu|_{\mathrm{tv}} = \frac{1}{2}\sum_{w\in[k]} |p(w) - \nu(w)| \leq \frac{1}{2}\sum_{w\in[k]} 2\beta \cdot p(w) = \beta. \tag{9.40}$$

Let $\mathscr{M}'$, $q(w)$ and τ as in Algorithm 9.2.2. Observe the following:

$$\mathbb{E}[\mathscr{M}'(\rho)] = \frac{1-q(w)}{q(w)(e^\varepsilon - e^{-\varepsilon})}p(w) - \frac{q(w)}{q(w)(e^\varepsilon - e^{-\varepsilon})}(1-p(w)) \tag{9.41}$$

$$= \frac{(1-q(w))p(w) - q(w)(1-p(w))}{q(w)(e^\varepsilon - e^{-\varepsilon})} = \frac{p(w)-q(w)}{q(w)(e^\varepsilon - e^{-\varepsilon})}. \tag{9.42}$$

Thus,

$$\nu = \frac{p(w)-q(w)}{q(w)(e^\varepsilon - e^{-\varepsilon})} \pm \frac{2\beta}{3e^{2\varepsilon}}. \tag{9.43}$$

This allows us to upper and lower bound the probability $\widetilde{p}(w)$ defined in Step 3 of the algorithm:

$$\widetilde{p}(w) = p(w)\left(1 \pm \frac{2\beta}{3e^{2\varepsilon}} \frac{q(w)}{p(w)}(e^{\varepsilon} - e^{-\varepsilon})\right). \tag{9.44}$$

By ε-local differential privacy,

$$e^{-\varepsilon} \leq \frac{q(w)}{p(w)} \leq e^{\varepsilon}. \tag{9.45}$$

Putting all together we obtain

$$\widetilde{p}(w) = p(w)\left(1 \pm \phi\right), \tag{9.46}$$

where we set $\phi := \frac{2\beta}{3}$. Having established Eq. 9.46, we can show that the algorithm works as desired. First, we notice that the probability introduced in Step 4 of the algorithm is well defined, as Eq. 9.46 and ε-local differential privacy guarantee that $\frac{\widetilde{p}(w)}{q(w)(1+\phi)e^{\varepsilon}}$ is at most 1. In a given iteration of the algorithm, any particular element w is output with probability $q(w) \cdot \frac{\widetilde{p}(w)}{q(w)(1+\phi)e^{\varepsilon}} = \frac{\widetilde{p}(w)}{(1+\phi)e^{\varepsilon}}$ and the probability that the given iteration terminates is then $p_{terminate} = \sum_w \frac{\widetilde{p}(w)}{(1+\phi)e^{\varepsilon}}$, which is in $\frac{1\pm\phi}{(1+\phi)e^{\varepsilon}}$ by Eq. 9.46. It's easy to see that, if the algorithm terminates in the current iteration, the element w is returned with probability

$$\Pr\left[w \text{ output in the } i^{th} \text{ iteration} \,\middle|\, i^{th} \text{ iteration produces output}\right] \tag{9.47}$$

$$= \frac{\Pr\left[w \text{ output in the } i^{th} \text{ iteration}\right]}{\sum_{w' \in [k]} \Pr\left[w' \text{ output in the } i^{th} \text{ iteration}\right]} \tag{9.48}$$

$$= \frac{\widetilde{p}(w)}{(1+\phi)e^{\varepsilon} p_{terminate}} \in \frac{1 \pm \phi}{1 \pm \phi} p(w). \tag{9.49}$$

Since $\phi \leq 1/3$, we obtain

$$\Pr\left[w \text{ output in the } i^{th} \text{ iteration} \,\middle|\, i^{th} \text{ iteration produces output}\right] \in (1 \pm 3\phi)\, p(w). \tag{9.50}$$

As a consequence, if the i^{th} iteration returns output w, the total variation distance between the distribution of w and $p(\cdot)$ will be at most $\frac{3}{2}\phi = \beta$ by Eq. 9.40. It remains to upper bound the expected number of quantum statistical queries. Recall that each iteration terminates with probability at least $\frac{1-\phi}{1+\phi}e^{-\varepsilon}$, hence the expected number of iterations is at most $\frac{1+\phi}{1-\phi}e^{\varepsilon} \leq 2e^{\varepsilon}$. Since a single QStat query is performed during each iteration, the total expected QSQ query complexity is $O(e^{\varepsilon})$. ■

As for the other direction of the equivalence, also Theorem 9.2 can be extended to the case where an algorithm $\mathscr{A}$ accesses t (unused) copies of a state ρ via ε-LDP measurements $\mathscr{M}^{(1)}, \mathscr{M}^{(2)}, \ldots, \mathscr{M}^{(t)}$. In order to simulate $\mathscr{A}$, it's sufficient to simulate each ε-LDP measurement $\mathscr{M}^{(i)}$ by running $\mathscr{B}_{\varepsilon}(\mathscr{M}^{(i)})$ with parameters $\beta' = \beta/t$. Then the output distribution of the simulation and the output distribution of $\mathscr{A}$ are within a total variation distance at most β. In the classical counterpart of this result [KLN+11a], the authors provide separate proofs for the adaptive and non-adaptive cases, as they assume that the algorithm $\mathscr{A}$ might reuse some portions of the input dataset. However in our proof we don't need to treat the two cases separately, as we assumed that each measurement is performed on a new copy of the input state ρ.

Algorithm 4 A QSQ algorithm $\mathscr{B}_{\mathscr{M},\varepsilon}(\beta, \mathsf{QStat}_\rho)$ that simulates an ε-LDP measurement $\mathscr{M}$

Input Oracle access to QStat_ρ, $\varepsilon \geq 0$, $\beta \geq 0$, an ε-LDP measurement $\mathscr{M} = (\mathscr{M}_1, \mathscr{M}_2, \ldots, \mathscr{M}_k)$ with outcomes in $[k]$.
Output A number $w \sim \nu$, such that $\nu(w) \in (1 \pm 2\beta)p(w)$.

1. Apply $\mathscr{M}$ to a fixed input, for instance the all-zeros state $|\mathbf{0}\rangle := |00...0\rangle$. Let $w \sim E(|\mathbf{0}\rangle\langle\mathbf{0}|)$ be the outcome.
2. Define $q(w) := \mathrm{Tr}\{\mathscr{M}_w |\mathbf{0}\rangle\langle\mathbf{0}|\}$ and $\mathscr{M}' = (\mathscr{M}'_0, \mathscr{M}'_1)$, where $\mathscr{M}'_0 := \mathscr{M}_w$ and $\mathscr{M}'_1 := I - \mathscr{M}_w$. $\mathscr{M}'_0$ corresponds to the outcome $\frac{1-q(w)}{q(w)(e^\varepsilon - e^{-\varepsilon})}$ and $\mathscr{M}'_1$ to the outcome $-\frac{q(w)}{q(w)(e^\varepsilon - e^{-\varepsilon})}$. Let $\tau = \frac{2\beta}{3e^{2\varepsilon}}$.
3. Query the oracle $\mathsf{QStat}_\rho(\mathscr{M}', \tau)$ to compute $\nu \in \mathbb{E}[\mathscr{M}'(\rho)] \pm \tau$. Define the probability:
$$\widetilde{p}(w) = \nu q(w)(e^\varepsilon - e^{-\varepsilon}) + q(w). \tag{9.51}$$
4. Output ν with probability
$$\frac{\widetilde{p}(w)}{q(w)\left(1 + \frac{2\beta}{3}\right)e^\varepsilon}. \tag{9.52}$$
5. With the remaining probability, repeat from Step 1.

Theorem 9.2 enables the transfer of lower bounds from the QSQ model to quantum local differential privacy. In particular, ([AHS23], Theorem 17) shows that learning the following class in the QSQ model requires exponentially many samples,

$$\mathscr{C} = \left\{ |\psi_A\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x, (x^\mathsf{T} A x) \mod 2 : A \in \mathbb{F}_2^{n \times n} \rangle \right\}. \tag{9.53}$$

On the other hand, this class is efficiently learnable using separable measurements [ABDY22] and entangled measurements with classification noise [AHS23]. Specifically, this immediately implies an exponential separation between learning under quantum local differential privacy and learning with separable measurements, resolving an open question in ([AR19], Question 4).

9.3 Testing and learning quantum states under local privacy

We will now explore the effect of local differential privacy in the settings of quantum hypothesis testing and quantum multi-party learning. Intuitively, as local differential privacy is ensured by the injection of noise, this will increase the sample complexity in a testing or learning task. We confirm this intuition by providing a converse bound on the achievable rate for quantum hypothesis testing under local differential privacy. On the other hand, we also demonstrate that quantum local differential privacy is compatible with exponential quantum advantage. As a proof of principle, we prove that parity functions can be learned from quantum examples under local differential privacy.

9.3.1 Private hypothesis testing

Here we demonstrate an application of the informatic-theoretic results of Section 9.1 to the rich field of quantum hypothesis testing. We study the distinguishability of two quantum states ρ and σ using a restricted class of measurements, i.e. locally differentially private measurements performed on a single copy of the input state. In particular, we'll consider the task of *asymmetric hypothesis testing*, where one wants to minimize the rate of false positives (type-1 error) subject to a constraint on the rate of false negatives (type-2 error). We will adopt the framework developed in [BHLP14], which extends hypothesis testing to the setting of restricted measurements. Our result can also be regarded as a quantum version of the "private Chernoff-Stein lemma" provided in [AAC21b].

Let ρ and σ be two quantum states acting on some Hilbert space $\mathscr{H}$. Given either n copies of ρ or n copies of σ, we want to design a test which distinguishes the two possibilities. For an acceptance operator $\mathscr{M}^n$ (i.e. a POVM element acting on n copies of the input state), we define the error probabilities as follows

$$\begin{aligned}\alpha_n(\mathscr{M}^n) &:= \mathrm{Tr}((I-\mathscr{M}^n)\rho^{\otimes n}) \quad \text{(type-2 error)},\\ \beta_n(\mathscr{M}^n) &:= \mathrm{Tr}(\mathscr{M}^n\sigma^{\otimes n}) \quad \text{(type-1 error)}.\end{aligned}$$

Then for $0<\tau<1$, define

$$\beta_n^\tau := \inf_{\mathscr{M}^n}\{\beta_n(\mathscr{M}^n) : \alpha_n(\mathscr{M}^n)\le\tau\} \tag{9.54}$$

and the asymptotic optimal error exponent

$$E(\rho,\sigma) := \lim_{\tau\to 0}\lim_{n\to\infty} -\frac{\log\beta_n^\tau}{n}. \tag{9.55}$$

The quantum Stein's lemma [HP91] says that

$$D(\rho\|\sigma) = E(\rho,\sigma). \tag{9.56}$$

As shown by [ON05], the "strong converse" Eq. 9.56 also holds. This can be thought of as showing that Eq. 9.56 is satisfied also when the limit of $\tau\to 0$ in Eq. 9.54 is replaced by any fixed $\tau\in(0,1)$. To deal with the restricted case where only single-copy ε-LDP measurements are allowed, we'll need to define the following quantities, introduced in [BHLP14]. Consider the infinite set $\mathbf{S} = (\mathrm{S}^1,\mathrm{S}^2,\ldots,\mathrm{S}^n,\ldots)$, where each S^n is a set of measurements over $\mathscr{H}^{\otimes n}$. We define:

$$D_{\mathrm{S}^n}(\rho\|\sigma) := \sup_{\mathscr{M}\in\mathrm{S}^n}\frac{D(\mathscr{M}(\rho^{\otimes n})\|\mathscr{M}(\sigma^{\otimes n}))}{n}. \tag{9.57}$$

$$D_{\mathbf{S}}(\rho\|\sigma) := \lim_{n\to\infty} D_{\mathrm{S}^n}(\rho\|\sigma). \tag{9.58}$$

In analogy with Eq. 9.54 and Eq. 9.55, we have

$$\beta_n^\tau(\mathbf{S}) := \inf_{\mathscr{M}\in\mathrm{S}^n}\{\beta_n(\mathscr{M}) : \alpha_n\le\tau\}, \tag{9.59}$$

$$E_{\mathbf{S}}(\rho,\sigma) := \lim_{\tau\to 0}\lim_{n\to\infty} -\frac{\log\beta_n^\tau(\mathbf{S})}{n}. \tag{9.60}$$

We are now ready to upper bound $E_{\mathbf{S}}(\rho,\sigma)$ for the case of locally differentially private measurements.

Theorem 9.3 (Private quantum Stein's lemma)**.** *Let ρ and σ be two quantum states acting on some Hilbert space $\mathscr{H}$. Let S_ε be a set of ε-LDP measurements over $\mathscr{H}$. Moreover, for every $n \geq 1$, define the following convex hull*

$$\mathrm{T}^n = \mathsf{conv}\{\mathscr{T}_1 \otimes \cdots \otimes \mathscr{T}_n : \mathscr{T}_1, \ldots, \mathscr{T}_n \in \mathrm{S}_\varepsilon\}, \tag{9.61}$$

and thus let $\mathbf{T} = (\mathrm{T}^1, \mathrm{T}^2, \ldots, \mathrm{T}^n, \ldots)$. The following inequality holds:

$$E_{\mathbf{T}}(\rho, \sigma) \leq \frac{e^\varepsilon}{2}(1 - e^{-\varepsilon})^2 D_M(\rho \| \sigma), \tag{9.62}$$

where $D_M(\cdot\|\cdot)$ denotes the measured relative entropy.

Proof. The theorem follows combining the results of [BHLP14] with Corollary 9.2. In particular, ([BHLP14], Theorem 16) implies that

$$E_{\mathbf{T}}(\rho, \sigma) = D_{\mathbf{T}}(\rho \| \sigma). \tag{9.63}$$

Recall that

$$D_{\mathbf{T}}(\rho \| \sigma) = \lim_{n \to \infty} \sup_{\mathscr{M} \in \mathrm{T}^n} \frac{D(\mathscr{M}(\rho^{\otimes n}) \| \mathscr{M}(\sigma^{\otimes n}))}{n}. \tag{9.64}$$

Observe that $\mathscr{M} = \sum_{i=1}^m \lambda_i (\mathscr{M}_1^{(i)} \otimes \cdots \otimes \mathscr{M}_n^{(i)})$ for some non-negative coefficients such that $\sum_i \lambda_i = 1$ and $\mathscr{M}_1^{(i)}, \ldots, \mathscr{M}_n^{(i)} \in \mathrm{S}_\varepsilon$. Recall that the quantum relative entropy enjoys joint convexity and additivity with respect to product states. Thus,

$$\begin{aligned} D(\mathscr{M}(\rho^{\otimes n}) \| \mathscr{M}(\sigma^{\otimes n})) &= D\left(\sum_{i=1}^m \lambda_i (\mathscr{M}_1^{(i)} \otimes \cdots \otimes \mathscr{M}_n^{(i)})(\rho^{\otimes n}) \middle\| \sum_{i=1}^m \lambda_i (\mathscr{M}_1^{(i)} \otimes \cdots \otimes \mathscr{M}_n^{(i)})(\sigma^{\otimes n}) \right) \\ &\leq \sum_{i,j} \lambda_i D\left(\mathscr{M}_j^{(i)}(\rho) \middle\| \mathscr{M}_j^{(i)}(\sigma) \right) \leq n \cdot \max_{i,j} D\left(\mathscr{M}_j^{(i)}(\rho) \middle\| \mathscr{M}_j^{(i)}(\sigma) \right) \\ &\leq n \cdot \frac{e^\varepsilon}{2}(1 - e^{-\varepsilon})^2 D_M(\rho \| \sigma), \end{aligned} \tag{9.65}$$

where the last inequality follows directly from Corollary 9.2. Finally, combining Eq. 9.64 and Eq. 9.65 yields

$$D_{\mathbf{T}}(\rho \| \sigma) \leq \lim_{n \to \infty} \frac{n}{n} \cdot \frac{e^\varepsilon}{2}(1 - e^{-\varepsilon})^2 D_M(\rho \| \sigma) = \frac{e^\varepsilon}{2}(1 - e^{-\varepsilon})^2 D_M(\rho \| \sigma), \tag{9.66}$$

and hence the theorem follows.

■

9.3.2 Private multi-party learning from quantum data

We will now discuss the applications of quantum local differential privacy to the setting of multi-party computation (MPC). In many real-world scenarios, multiple parties share their data to collectively compute a function. The goal is then to achieve the best possible accuracy under some security constraints. One way to formulate the security requirement is to ask that each party learns nothing more about the other parties' data than can be learned from the output of the function computed.

This approach is adopted by the framework of *secure multi-party computation* (SMPC), both in the classical [Yao86, GMW87] and in the quantum setting [CGS02, DGJ+20]. The main shortcoming of SMPC is that the security guarantees are dependent on the auxiliary information disposed by the adversary. For instance, if k parties collectively compute an average, $k-1$ malicious parties can collaborate to infer the data of the remaining party.

To overcome these limitations, we can adopt the framework of *secure multi-party differential privacy*, defined in [KOV15]. In particular, we will consider a model where the input state $\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_k$ is distributed among k quantum parties $\mathscr{P}_1, \mathscr{P}_2, \ldots, \mathscr{P}_k$, such that the i-th party $\mathscr{P}_i$ holds the state ρ_i and disposes of a quantum computer. The parties are allowed to share classical information. In order to protect the private information contained in ρ_i, we require that the each $\mathscr{P}_i$ accesses the state ρ_i through an ε-local differentially private measurement $\mathscr{M}_i$ for some suitable $\varepsilon > 0$. Thus, for all i, for any possible output y, and for all input states ρ_i, σ_i, we have

$$\Pr[\mathscr{M}_i(\rho_i) = y] \leq e^{\varepsilon} \Pr[\mathscr{M}_i(\sigma_i) = y]. \tag{9.67}$$

One potential concern with this setting is that the injection of noise can severely limit the usefulness of the computation, hence it is no clear a priori whether a quantum speed-up can be achieved under these constraints. To address this issue, we show that that parity functions can be efficiently learned from quantum examples in a multi-party setting under local differential privacy. Classically, learning parity under local differential privacy requires exponentially many samples [KLN+11a].

For $s \in \{0,1\}^n$, the corresponding parity function $c : \{0,1\}^n \to \{-1,1\}$ is defined $c(x) = (-1)^{s \cdot x}$. Let $b^1, \ldots, b^k$ random binary strings in $\{\pm 1\}^n$, such that each b^i_x equals 1 with probability $9/10$ and -1 with probability $1/10$. Each party $\mathscr{P}_i$ holds the following quantum state:

$$|\psi_i\rangle = \sqrt{\frac{1}{2^n}} \sum_{x \in \{0,1\}^n} |x, c(x) \cdot b^i_x\rangle. \tag{9.68}$$

We remark that this definition slightly differs from the one considered in ([AGY20], Lemma 4.2), as their definition doesn't involve the random vector b_i. Instead, in our model each party holds a different input state. The vector b^i can be either regarded as classification noise or as some sensitive information regarding the i-th party. In the latter case, the adoption of local differential privacy is extremely natural, as it significantly limits the information about b^i that can be inferred by a malicious adversary, even disposing of auxiliary information.

Proposition 9.2. *Let $s \in \{0,1\}^n$ and $|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_k\rangle$ as defined above and assume that the parties $\mathscr{P}_i$'s can communicate via a classical channel. Provided that $k \geq c \cdot n\varepsilon^{-2} \log(1/\beta)$ for a sufficiently large constant c, there is an efficient quantum algorithm $\mathscr{A}$ that computes the string s with probability at least $1-\beta$. $\mathscr{A}$ consists solely in ε-LDP measurements on the states $|\psi_i\rangle$'s, classical communication and classical post-processing.*

Proof. The proof is similar to the one of ([AGY20], Lemma 4.2). It is not hard to see that $\mathrm{Inf}_j(c) = 1$ for all $j \in \mathrm{supp}(s)$ and $\mathrm{Inf}_j(c) = 0$ otherwise. As shown in [AGY20], there is a quantum measurement M_j implementable in $\mathrm{poly}(n)$ gates such that

$$\langle\psi| M_j |\psi\rangle = \mathrm{Inf}_j(c), \tag{9.69}$$

where $|\psi\rangle = \sqrt{\frac{1}{2^n}} \sum_{x\in\{0,1\}^n} |x, c(x)\rangle$. Moreover, the expected trace distance between $|\psi\rangle$ and $|\psi_i\rangle$ can be bounded as follows:

$$\mathbb{E}_{b_i} \| |\psi\rangle\langle\psi| - |\psi_i\rangle\langle\psi_i| \|_{\mathrm{tr}} = \mathbb{E}_{b_i} \left[\sqrt{1 - \langle\psi_i|\psi\rangle} \right] = \sqrt{1 - \sqrt{1 - 1/10}} < \frac{1}{4}, \tag{9.70}$$

where we took the expectation over the randomness of the string b_i. By the property of the trace distance,

$$\left| \mathbb{E}_{b_i} \langle\psi_i| M_j |\psi_i\rangle - \mathrm{Inf}_j(c) \right| < \frac{1}{4}. \tag{9.71}$$

Then the algorithm $\mathscr{A}$ estimates $\mathrm{Inf}_j(c)$ by asking $m > 64 \cdot \varepsilon^{-2} \log(3/\beta)$ parties to perform a Laplace measurement $M_j^{\mathrm{Lap},\varepsilon}$ on their state $|\psi_i\rangle$ and averaging the outcomes $\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_m$. We denote their average by $\hat{\mu} = \frac{1}{n}\sum_{i=1}^m \hat{y}_i$. We can write $\hat{y}_i = y_i + \eta$, where $y_i \sim \mathbb{E}_{b_i} \langle\psi_i| M_j |\psi_i\rangle$ and $\eta \sim \mathrm{Lap}(1/\varepsilon)$. Proceeding as in the proof of Theorem 9.1, we can show by concentration of measure that

$$\hat{\mu} = \mathbb{E}_{b_i} \langle\psi_i| M_j |\psi_i\rangle \pm 1/4, \tag{9.72}$$

with probability at least $1 - \beta$. Then the outcome $\hat{\mu}$ is in the interval $(1/2, 3/2)$ if $j \in \mathrm{supp}(s)$, otherwise is in $(-1/2, 1/2)$. Thus we can determine whether $j \in \mathrm{supp}(s)$. Repeating the procedure for all $j \in [n]$ on m unused states $|\psi_i\rangle$'s, we can determine the string s. This requires k to scale as $O(n\varepsilon^{-2}\log(1/\beta))$. ■

CHAPTER 10

CONCLUSION

There is nothing more vivifying than a hypothesis.

- Primo Levi, *The Periodic Table*

This thesis has tackled several challenges at the heart of quantum computing on noisy near-term devices. It has addressed pressing issues surrounding the impact of noise on quantum computation and the exciting prospects of quantum differential privacy in machine learning. In doing so, it has not only revealed valuable insights but has also left a roadmap for future research in the field of quantum computing and quantum information.

Now, we provide a set of takeaway messages that can inform and drive future research in the field.

- Depolarizing noise, and more broadly, local Pauli noise, serves as an idealized model that simplifies the complex features of noisy near-term devices. However, it falls short in capturing all the nuances of noisy variational quantum algorithms when confronted with real-world noise scenarios. Notably, local quantities like the projected quantum kernels exhibit distinct behaviors under the influence of unital and non-unital noise. As a result, it becomes imperative to reevaluate and extend previous research, which predominantly relies on the depolarizing noise model, such as [CCHL22a, AGL$^+$23b], to encompass the effects of non-unital perturbations.

- The quantum statistical query model provides a solid and reliable framework for designing learning algorithms tailored to quantum dynamics. Importantly, numerous algorithmic concepts originally designed for learning classical functions from quantum examples can be seamlessly adapted for the task of quantum process learning. This implies that the intrinsic

noise associated with quantum measurement implementation does not always hinder the use of quantum algorithms.

- Considering the inherent noise in near-term quantum devices, a comprehensive evaluation of their performance should take into consideration aspects like privacy and robustness against adversarial attacks. We introduced a novel framework for analyzing differential privacy within quantum algorithms, which encompasses the influence of both classical and quantum noise, along with various distinct quantum encodings. Our findings highlight that ensuring privacy necessitates a creative approach in defining an appropriate quantum neighboring relationship. Relying on conventional metrics, such as the trace distance or quantum Wasserstein distance of order 1, could yield significantly suboptimal results.

  In addition, our research has provided insight into the interplay between privacy and quantum advantage. We have shown that even when significant noise is introduced, as is the case in local differential privacy, quantum speed-ups for certain problems can still be achieved.

10.1 Future directions

We distill several open questions and conjectures concerning the topics undertaken in the present thesis.

Variational quantum algorithms under non-unital noise While we have contributed fresh insights into the trainability of quantum kernels under non-unital noise and introduced the concept of an "effective depth circuit", there are still several gaps in our understanding before we can fully comprehend the entire scenario. For the sake of clarity and convenience, we reiterate our conjecture that have yet to be proven.

Conjecture 1. *Let $\mathscr{C}$ be a noisy circuit consisting in m layers of 2-qubit gates interspersed with local noise, either of the form $\mathcal{N}_{p,q}^{(\mathrm{dep,amp}),\otimes n}$ or $\mathcal{N}_{q,p}^{(\mathrm{amp,dep}),\otimes n}$. Moreover, assume that each 2-qubit gate is sampled independently from a local 2-design. Then if $p,q=\Omega(1)$, for all $\rho,\sigma\in\mathcal{S}_n$, we have*

$$\mathbb{E}\|\mathscr{C}(\rho)-\mathscr{C}(\sigma)\|_{\mathrm{tr}}\in 2^{\Omega(-n)} \tag{10.1}$$

Quantum statistical query learning

1. The main workhorse for QSQ learning classical Boolean functions is Fourier analysis. While Fourier analysis is usually cast under the uniform distribution, the *μ-biased* Fourier analysis can be applied to every product distribution. In particular, μ-biased Fourier sampling can be used to learn linear functions [Car20] and DNFs [KRS19] under product distributions with quantum examples. Can we extend these results to the QSQ model?

2. In analogy with the previous question, we ask whether we can learn the action of unitary operators on ensembles of states that are not locally scrambled. For this, it might be fruitful to define the following generalization of the Choi-Jamiolkoski state, analogous to a quantum example under a non-uniform distribution:

$$|\nu_D(U)\rangle = (I \otimes U) \sum_{x \in \{0,1\}^n} \sqrt{D(x)}\, |x, x\rangle\,, \tag{10.2}$$

   where D is a suitable distribution over $\{0,1\}^n$. A natural approach would be to extend the μ-*biased* Fourier analysis to the unitary group. We ask this question both for quantum statistical queries and more powerful oracles.

3. Which classes of channels can be learned with quantum statistical queries?

4. What is the power of quantum statistical queries for testing properties of unitaries (and more broadly channels)? While we provided a double exponential lower bound for testing unitarity, quantum statistical queries might suffice for testing other relevant properties.

5. Following [HIN+23, NIS+23], we can restrict our model to *diagonal* measurements. Which classes of channels are learnable under this restricted model?

Quantum differential privacy

1. Our analysis focused on the privacy guarantee of unital noise channels. What are the privacy guarantees of local amplitude damping noise with respect to the generalized neighboring relationship introduced in this these?

2. Although numerous numerical assessments of the adversarial robustness of quantum algorithms have been conducted [DHL+21a, ADK23], there is a compelling need for a systematic analysis on real quantum devices. Such an analysis would delve into the robustness arising from the presence of realistic quantum noise.

3. In a recent study, it was concluded that expressive variational quantum circuits offer inherent privacy within the context of federated learning [KHL+23]. Their definition of privacy essentially aligns with preventing the server from executing a reconstruction attack on the underlying input data. This leads to the following question: can these findings be reformulated within the framework of differential privacy?

4. Can we attain a quantum advantage for a real-word learning task under local differential privacy?

CHAPTER 11

SUPPLEMENTARY MATERIALS

11.1 Improved bounds for quantum divergences

We present two technical contributions that establish tighter bounds for quantum divergences. First, we prove here a quantum version of the Bretagnolle-Huber (BH) inequality [BH78, Can22]. The proof closely follows the one of the classical BH inequality, and for this reason the quantum BH can be regarded as a folklore result. However, we include here the complete proof since, to the best of our knowledge, it does not appear in any previous reference. We remark that a different quantum generalisation of the BH inequality result was provided in [PC18] in the context of local measurements.

Lemma 11.1 (Quantum Bretagnolle-Huber inequality)**.** *For every ρ, σ we have*

$$\frac{1}{2}\|\rho-\sigma\|_1 \leq \sqrt{1-e^{-D(\rho\|\sigma)}} \tag{11.1}$$

Proof. We define the following quantity

$$U := \rho^{-1}\sigma, \tag{11.2}$$

$$V := (U-I)^{+}, \tag{11.3}$$

$$W := I+V-U = (U-I)^{-}. \tag{11.4}$$

It's well known that

$$\mathrm{Tr}(\rho V) = \mathrm{Tr}(\sigma-\rho)^+ = \frac{1}{2}\|\rho-\sigma\|_1, \tag{11.5}$$

$$\mathrm{Tr}(\rho W) = \mathrm{Tr}(\sigma-\rho)^- = \frac{1}{2}\|\rho-\sigma\|_1. \tag{11.6}$$

Moreover, remark that $(1+V)(1-W) = U$ and hence $\log U = \log(I+V) + \log(I-W)$. Applying the Jensen's inequality, we obtain

$$-D(\rho\|\sigma) \leq \mathrm{Tr}[\rho\log(\rho^{-1}\sigma)] = \mathrm{Tr}[\rho\log U] \tag{11.7}$$

$$= \mathrm{Tr}[\rho\log(I+V)] + \mathrm{Tr}[\rho\log(I-W)] \leq \log\mathrm{Tr}[\rho(I+V)] + \log\mathrm{Tr}[\rho(I-W)] \tag{11.8}$$

$$= \log(1-\mathrm{Tr}[\rho V]) + \log(1-\mathrm{Tr}[\rho W]) = \log\left(1-\frac{1}{2}\|\rho-\sigma\|_1^2\right). \tag{11.9}$$

Exponentiating both sides, rearranging and taking the square root, proves the lemma. ■

Building upon [BBG18], we prove a quantum version of the *advanced joint convexity* of the hockey-stick divergence.

Lemma 11.2 (Advanced joint convexity of the quantum hockey-stick divergence)**.** *For all states ρ_0,ρ_1,ρ_2 and $\gamma' = 1+(1-p)(\gamma-1)$, we have*

$$E_{\gamma'}(p\rho_0+(1-p)\rho_1\|p\rho_0+(1-p)\rho_2) \leq (1-p)(1-\beta)E_\gamma(\rho_1\|\rho_0) + (1-p)\beta E_\gamma(\rho_1\|\rho_2), \tag{11.10}$$

where $\beta = \gamma'/\gamma$.

Proof. Recall that

$$E_\gamma(\rho\|\sigma) := \mathrm{Tr}(\rho-\gamma\sigma)^+ = \frac{1}{2}\|\rho-\gamma\sigma\|_1 + \frac{1}{2}(1-\gamma). \tag{11.11}$$

We have

$$E_{\gamma'}(p\rho_0+(1-p)\rho_1\|p\rho_0+(1-p)\rho_2) = \mathrm{Tr}[p\rho_0+(1-p)\rho_1-\gamma'(p\rho_0+(1-p)\rho_2)]^+ \tag{11.12}$$

$$= \mathrm{Tr}[p\rho_0+(1-p)\rho_1-(1+(1-p)(\gamma-1))(p\rho_0+(1-p)\rho_2)]^+ \tag{11.13}$$

$$= (1-p)\,\mathrm{Tr}[\rho_1-\gamma(\rho_0(1-\beta)+\beta\rho_2)]^+ = (1-p)E_\gamma(\rho_1\|\rho_0(1-\beta)+\beta\rho_2) \tag{11.14}$$

$$\leq (1-p)(1-\beta)E_\gamma(\rho_1\|\rho_0) + (1-p)\beta E_\gamma(\rho_1\|\rho_2), \tag{11.15}$$

where the inequality follows from the (standard) joint-convexity of the quantum hockey-stick divergence. ■

11.2 Quantum encodings

Quantum encodings, also known as quantum feature maps or quantum embedding, are classical-to-quantum functions mapping vectors to quantum states. In this section, we review some popular encodings and highlight their connection with various quantum distances and neighbouring relationships. We refer to [Sch21] for more details about the encodings and their corresponding kernel (i.e. the value of $|\langle\psi_{\boldsymbol{x}}|\psi_{\boldsymbol{x}'}\rangle|^2$ for two vectors $\boldsymbol{x},\boldsymbol{x}'$). Throughout this section, we will show that encoding vectors close in various p-distance leads to states that are either close in trace distance or that can be mapped one into the other by a local operation.

Amplitude encoding. A normalised vector $\boldsymbol{x} = (x_1,\dots,x_{2^n}) \in \mathbb{C}^{2^n}$, $\|\boldsymbol{x}\|_2 = 1$ can be represented by the amplitudes of a quantum state $|\psi_{\boldsymbol{x}}\rangle$ via

$$\boldsymbol{x} \mapsto |\psi_{\boldsymbol{x}}\rangle = \sum_{j=1}^{2^n} x_j |j\rangle . \tag{11.16}$$

For two normalised vectors $\boldsymbol{x},\boldsymbol{x}'$ we have

$$|\langle\psi_{\boldsymbol{x}}|\psi_{\boldsymbol{x}'}\rangle| = |\boldsymbol{x}^\dagger \boldsymbol{x}'| = \left|1 - \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{x}'\|_2^2\right|, \tag{11.17}$$

where the second identity holds for any pair of normalised vectors. Hence,

$$\frac{1}{2}\| |\psi_{\boldsymbol{x}}\rangle\langle\psi_{\boldsymbol{x}}| - |\psi_{\boldsymbol{x}'}\rangle\langle\psi_{\boldsymbol{x}'}| \|_1 = \sqrt{1-|\langle\psi_{\boldsymbol{x}}|\psi_{\boldsymbol{x}'}\rangle|^2} \tag{11.18}$$

$$= \sqrt{1-|\boldsymbol{x}^\dagger\boldsymbol{x}'|^2} = \sqrt{1-\left(1-\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{x}'\|_2^2\right)^2} \tag{11.19}$$

$$\leq \|\boldsymbol{x}-\boldsymbol{x}'\|_2. \tag{11.20}$$

Rotation encoding. Rotation encoding is a qubit-based embedding without any normalisation condition. Given a vector $\boldsymbol{x}$ in the hypercube $[0,2\pi]^{\otimes n}$, the i^{th} feature x_i is encoded into the i^{th} qubit via a Pauli rotation. For example, a Pauli-Y rotation puts the qubit into state $|q_i(x_i)\rangle = \cos(x_i)|0\rangle + \sin(x_i)|1\rangle$. The data-encoding feature map is therefore given by

$$\phi : \boldsymbol{x} \to \rho(\boldsymbol{x}) := |\phi(\boldsymbol{x})\rangle\langle\phi(\boldsymbol{x})| \text{ with } |\phi(\boldsymbol{x})\rangle = \sum_{q_1,\dots,q_n=0}^{1} \prod_{k=1}^{n} \cos(x_k)^{q_k}\sin(x_k)^{1-q_k}|q_1,\dots,q_n\rangle . \tag{11.21}$$

Let $\mathscr{I} = \{i : x_i \neq x_i'\}$. We have that $|\mathscr{I}| = \|\boldsymbol{x}-\boldsymbol{x}'\|_0$. We immediately see that

$$\mathrm{Tr}_{\mathscr{I}}\,\rho(\boldsymbol{x}) = \mathrm{Tr}_{\mathscr{I}}\,\rho(\boldsymbol{x}'). \tag{11.22}$$

Hamiltonian encoding. Let $\boldsymbol{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ be a vector. Following [BFH23], consider the following parameterised quantum circuit

$$|\psi(\boldsymbol{x})\rangle = U_1(x_1) \cdots U_N(x_N) |\psi_0\rangle, \tag{11.23}$$

consisting of N parametric unitary operators $U_i(x_i) \in \mathscr{U}_n$ acting on the initial state $|\psi_0\rangle$. Let $\rho(\boldsymbol{x}) := |\psi(\boldsymbol{x})\rangle\langle\psi(\boldsymbol{x})|$. These unitaries can also be written as $U_j(x_j) = e^{-ix_j H_j}$, where the Hamiltonian $H_i = H_i^\dagger$ generates the gate U_i. The following result shows that quantum circuits are robust to slight perturbation of the classical parameters.

Lemma 11.3 (Adapted from Theorem 2.2, [BFH23])**.** *Let $\boldsymbol{x}, \boldsymbol{x'} \in \mathbb{R}^N$. $U(\theta) = e^{-i\theta H}$. For any initial state $|\psi_0\rangle$ we have*

$$\| |\psi(\boldsymbol{x})\rangle\langle\psi(\boldsymbol{x})| - |\psi(\boldsymbol{x'})\rangle\langle\psi(\boldsymbol{x'})| \|_2 \leq \sum_{i=1}^{N} \|H_i\|_2 |x_i - x_i'| \leq \|\boldsymbol{x} - \boldsymbol{x'}\|_1 \max_i \|H_i\|_2. \tag{11.24}$$

Remark also that for ρ, σ pure states we have $\|\rho - \sigma\|_1 = \sqrt{2}\|\rho - \sigma\|_2$ and for any vectors $x, x' \in \mathbb{R}^N$ we have $\|\boldsymbol{x} - \boldsymbol{x'}\|_1 \leq \sqrt{N}\|\boldsymbol{x} - \boldsymbol{x'}\|_2$. Then we have:

$$\frac{1}{2}\|\rho(\boldsymbol{x}) - \rho(\boldsymbol{x'})\|_1 \leq \sqrt{\frac{1}{2}}\|\boldsymbol{x} - \boldsymbol{x'}\|_1 \max_i \|H_i\|_2 \tag{11.25}$$

$$\leq \sqrt{\frac{N}{2}}\|\boldsymbol{x} - \boldsymbol{x'}\|_2 \max_i \|H_i\|_2. \tag{11.26}$$

It's easy to see that the circuits $U(\boldsymbol{x})$ and $U(\boldsymbol{x'})$ coincides excepts for $\|\boldsymbol{x} - \boldsymbol{x'}\|_0$ gates. In order to investigate the local structure of the output, we need to introduce some assumptions on the circuit architecture. For instance, assuming that the circuit has 1-dimensional connectivity and depth L, there exists $\mathscr{I} \subseteq [n]$, $|\mathscr{I}| \leq 2L\|\boldsymbol{x} - \boldsymbol{x'}\|_0$, such that

$$\mathrm{Tr}_{\mathscr{I}}\, \rho(\boldsymbol{x}) = \mathrm{Tr}_{\mathscr{I}}\, \rho(\boldsymbol{x'}). \tag{11.27}$$

11.2.1 Noisy encodings

A case of interest is when the circuit $U(\boldsymbol{x})$ is interspersed of L layers of local Pauli noise $\mathscr{P}_q$. Let $\mathscr{C}_{\boldsymbol{x}}$ be the channel describing the composition of unitaries and noise:

$$\mathscr{C}_{\boldsymbol{x}}(\rho_0) = \mathscr{P}_q^{\otimes n} \circ U_N(x_N)(\cdot)U_N(x_N)^\dagger \circ \mathscr{P}_q^{\otimes n} \circ \cdots \circ \mathscr{P}_q^{\otimes n} \circ U_1(x_1)(\rho_0)U_1^\dagger(x_1), \tag{11.28}$$

where $\mathscr{P}_q$ represent a local Pauli noise channel with noise strength q. Then by Lemma 4.2, we get:

$$D_2(\mathscr{C}_{\boldsymbol{x}}(\rho_0)\|I/2^n) \leq q^{2L}n. \tag{11.29}$$

and by Pinsker's inequality,

$$\frac{1}{2}\left\|\mathscr{C}_{\boldsymbol{x}}(\rho_0) - \frac{I}{2^n}\right\|_1 \leq \sqrt{\frac{q^{2L}n}{2}}. \tag{11.30}$$

Alternatively, by the quantum Bretagnolle-Huber inequality (Lemma 11.1),

$$\frac{1}{2}\left\|\mathscr{C}_{\boldsymbol{x}}(\rho_0)-\frac{I}{2^n}\right\|_1 \leq \sqrt{1-\exp\left(-q^{2L}n\right)}. \tag{11.31}$$

And by the triangle inequality

$$\frac{1}{2}\|\mathscr{C}_{\boldsymbol{x}}(\rho_0)-\mathscr{C}_{\boldsymbol{x}'}(\rho_0)\|_1 \leq 2\min\left\{\sqrt{\frac{q^{2L}n}{2}}, \sqrt{1-\exp\left(-q^{2L}n\right)}\right\}. \tag{11.32}$$

High noise regime Now, assume that $\rho(\cdot)$ is an encoding post-processed by a channel $\mathscr{A}$, consisting in L layers such that each of them has light-cone $\mathscr{I}$ and its followed by local depolarising noise with noise parameter p. If p satisfies $2|\mathscr{I}|(1-p)<1$, we have from ([HRF23], Proposition IV.8),

$$\frac{1}{2}\|\mathscr{A}(\rho(\boldsymbol{x}))-\mathscr{A}(\rho(\boldsymbol{x'}))\|_1 \tag{11.33}$$

$$\leq (2|\mathscr{I}|(1-p))^L W_1(\rho(\boldsymbol{x}),\rho(\boldsymbol{x'})) \tag{11.34}$$

For $\rho(\boldsymbol{x}) \overset{(\Xi,\tau)}{\sim} \rho(\boldsymbol{x'})$ we have

$$\frac{1}{2}\|\rho(\boldsymbol{x})-\rho(\boldsymbol{x'})\|_1 \leq W_1(\rho(\boldsymbol{x}),\rho(\boldsymbol{x'})) \leq \min\left\{\max_{\mathscr{I}\in\Xi}|\mathscr{I}|\frac{3}{2}\tau, n\tau\right\}. \tag{11.35}$$

11.3 Private quantum-inspired sampling

Our argument is similar to the one of (Problem 1.b, [Ull17]) for uniform subsampling, but we include the complete proof here for clarity. Given a normalised vector $x=(x_1,\ldots,x_n)\in\mathbb{C}^n$, let $|x\rangle := \sum_{i=1}^n x_i|i\rangle$ be the amplitude encoding defined in the previous section.

Theorem 11.1 (DP amplification by quantum-inspired sampling)**.** *For any $x\in\mathbb{C}^n$, let $s=(s_1,\ldots,s_m)$ be the measurement outcomes in the computational basis of $|x\rangle^{\otimes m}$. Denote $\mathscr{S}$ as the sampling mechanism that maps x into s. Let $\mathscr{A}$ be a (ε,δ)-DP algorithm that takes only s as input. Then $\mathscr{A}'=\mathscr{A}\circ\mathscr{S}$ is (ε',δ')-DP, with $\varepsilon'=\log(1+(e^\varepsilon-1)m(\alpha+\beta))$ and $\delta'=\delta m(\alpha+\beta)$.*

Proof. We will use $T\subseteq\{1,\ldots,n\}$ to denote the identities of the m-subsampled elements $s_1,\ldots,s_m$ (i.e. their index, not their actual value). Note that T is a random variable and that the randomness of $\mathscr{A}':=\mathscr{A}\circ\mathscr{S}$ includes both the randomness of the sample T and the random coins of $\mathscr{A}$. Let $x\sim x'$ be adjacent datasets and assume that x and x' differ only on some row t. Let s (or s') be a subsample from x (or x') containing the rows in T . Let F be an arbitrary subset of the range of $\mathscr{A}$). For convenience, define $p=(\alpha+\beta)m$. Note that, by definition of quantum amplitude encoding and by union bound,

$$\Pr[i\in T] \leq m\times\Pr[|x\rangle \text{ collapses to state } |i\rangle] \leq m(\alpha+\beta) := p \tag{11.36}$$

To show $(\log(1+p(e^{\varepsilon}-1)), p\delta)$-DP, we have to bound the ratio

$$\frac{\Pr[\mathscr{A}'(x)\in F]-p\delta}{\Pr[\mathscr{A}'(x')\in F]} \leq \frac{p\Pr[\mathscr{A}(s)\in F|i\in T]+(1-p)\Pr[\mathscr{A}(s)\in F|i\notin T]-p\delta}{p\Pr[\mathscr{A}(s')\in F|i\in T]+(1-p)\Pr[\mathscr{A}(s')\in F|i\notin T]} \tag{11.37}$$

by $p(1+(e^{\varepsilon}-1))$. For simplicity, define the quantities

$$C = \Pr[\mathscr{A}(s)\in F|i\in T] \tag{11.38}$$

$$C' = \Pr[\mathscr{A}(s')\in F|i\in T] \tag{11.39}$$

$$D = \Pr[\mathscr{A}(s)\in F|i\notin T] = \Pr[\mathscr{A}(s')\in F|i\notin T]. \tag{11.40}$$

We can rewrite the ratio as

$$\frac{\Pr[\mathscr{A}'(x)\in F]-p\delta}{\Pr[\mathscr{A}'(x')\in F]} = \frac{pC+(1-p)D-p\delta}{pC'+(1-p)D}. \tag{11.41}$$

Now we use the fact that, by (ε,δ)-DP, $C\leq \min\{C',D\}+\delta$. Plugging all together, we get

$$pC+(1-p)D-p\delta \leq p(e^{\varepsilon}\min\{C',D\})+(1-p)D \tag{11.42}$$

$$\leq p(\min\{C',D\}+(e^{\varepsilon}-1)\min\{C',D\})+(1-p)D \tag{11.43}$$

$$\leq p(C'+(e^{\varepsilon}-1)(pC'+(1-p)D))+(1-p)D \tag{11.44}$$

$$\leq (pC'+(1-p)D)+p(e^{\varepsilon}-1))(pC'+(1-p)D) \leq (1+p(e^{\varepsilon}-1))(pC'+(1-p)D), \tag{11.45}$$

where the third-to-last line follow from $\min\{x,y\}\leq \alpha x+(1-\alpha)y$ for every $0\leq\alpha\leq 1$. To conclude the proof, we rewrite the ratio and get the desired bound.

$$\frac{\Pr[\mathscr{A}'(x)\in F]-p\delta}{\Pr[\mathscr{A}'(x')\in F]} \leq 1+p(e^{\varepsilon}-1). \tag{11.46}$$

■

BIBLIOGRAPHY

[AA23] Anurag Anshu and Srinivasan Arunachalam.
A survey on the complexity of learning quantum states.
arXiv preprint arXiv:2305.20069, 2023.

[AAB+19] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al.
Quantum supremacy using a programmable superconducting processor.
Nature, 574(7779):505–510, 2019.

[AAC21a] Shahab Asoodeh, Maryam Aliakbarpour, and Flavio P Calmon.
Local differential privacy is equivalent to contraction of an f-divergence.
In *2021 IEEE International Symposium on Information Theory (ISIT)*, pages 545–550. IEEE, 2021.

[AAC21b] Shahab Asoodeh, Maryam Aliakbarpour, and Flavio P. Calmon.
Local differential privacy is equivalent to contraction of an f-divergence.
In *2021 IEEE International Symposium on Information Theory (ISIT)*, page 545–550. IEEE Press, 2021.

[AACM+22] John M Abowd, Robert Ashmead, Ryan Cumings-Menon, Simson Garfinkel, Micah Heineck, Christine Heiss, Robert Johns, Daniel Kifer, Philip Leclerc, Ashwin Machanavajjhala, et al.
The 2020 census disclosure avoidance system topdown algorithm.
Harvard Data Science Review, (Special Issue 2), 2022.

[Aar18] Scott Aaronson.
Shadow tomography of quantum states.
In *Proceedings of the 50th annual ACM SIGACT symposium on theory of computing*, pages 325–338, 2018.

[ABDY22] Srinivasan Arunachalam, Sergey Bravyi, Arkopal Dutt, and Theodore J Yoder.
Optimal algorithms for learning quantum phase states.
arXiv preprint arXiv:2208.07851, 2022.

[Abo18] John M Abowd.
The us census bureau adopts differential privacy.

In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2867–2867, 2018.

[ACC+21] Andrew Arrasmith, Marco Cerezo, Piotr Czarnik, Lukasz Cincio, and Patrick J Coles.
Effect of barren plateaus on gradient-free optimization.
Quantum, 5:558, 2021.

[ACG+16] Martin Abadi, Andy Chu, Ian Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang.
Deep learning with differential privacy.
Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Oct 2016.

[ACK21] Armando Angrisani, Brian Coyle, and Elham Kashefi.
Probably approximately correct quantum source coding.
arXiv preprint arXiv:2112.06841, 2021.

[ADK22] Armando Angrisani, Mina Doosti, and Elham Kashefi.
Differential privacy amplification in quantum and quantum-inspired algorithms.
arXiv preprint arXiv:2203.03604, 2022.

[ADK23] Armando Angrisani, Mina Doosti, and Elham Kashefi.
A unifying framework for differentially private quantum algorithms.
arXiv preprint arXiv:2307.04733, 2023.

[AGL+23a] Dorit Aharonov, Xun Gao, Zeph Landau, Yunchao Liu, and Umesh Vazirani.
A polynomial-time classical algorithm for noisy random circuit sampling.
In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 945–957, 2023.

[AGL+23b] Dorit Aharonov, Xun Gao, Zeph Landau, Yunchao Liu, and Umesh Vazirani.
A polynomial-time classical algorithm for noisy random circuit sampling.
In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 945–957, 2023.

[AGY20] Srinivasan Arunachalam, Alex B Grilo, and Henry Yuen.
Quantum statistical query learning.
arXiv preprint arXiv:2002.08240, 2020.

[AHS23] Srinivasan Arunachalam, Vojtech Havlicek, and Louis Schatzki.
On the role of entanglement and statistics in learning.
arXiv preprint arXiv:2306.03161, 2023.

[AK22a] Armando Angrisani and Elham Kashefi.
Quantum local differential privacy and quantum statistical query model.
arXiv preprint arXiv:2203.03591, 2022.

[AK22b] Armando Angrisani and Elham Kashefi.
Quantum local differential privacy and quantum statistical query model.
ArXiv, abs/2203.03591, 2022.

[Ang23] Armando Angrisani.
Learning unitaries with quantum statistical queries, 2023.

[AQS21] Srinivasan Arunachalam, Yihui Quek, and John Smolin.
Private learning implies quantum stability.
In *Advances in Neural Information Processing Systems 34 pre-proceedings (NeurIPS 2021)*, NIPS'21, 2021.

[AR19] Scott Aaronson and Guy N Rothblum.
Gentle measurement of quantum states and differential privacy.
In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 322–333, 2019.

[AS07] Alp Atıcı and Rocco A Servedio.
Quantum algorithms for learning and testing juntas.
Quantum Information Processing, 6(5):323–348, 2007.

[BBG18] Borja Balle, Gilles Barthe, and Marco Gaboardi.
Privacy amplification by subsampling: Tight analyses via couplings and divergences.
In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 6280–6290, Red Hook, NY, USA, 2018. Curran Associates Inc.

[BC23] Tomislav Begušić and Garnet Kin Chan.
Fast classical simulation of evidence for the utility of quantum computing before fault tolerance.
arXiv preprint arXiv:2306.16372, 2023.

[BF12] Winton Brown and Omar Fawzi.
Scrambling speed of random quantum circuits.
arXiv preprint arXiv:1210.6644, 2012.

[BF15] Winton Brown and Omar Fawzi.
Decoupling with random quantum circuits.
Communications in mathematical physics, 340:867–900, 2015.

[BFH23] J. Berberich, D. Fink, and C. Holm.
Robustness of quantum algorithms against coherent control errors, 2023.

[BH78] Jean Bretagnolle and Catherine Huber.
Estimation des densités: risque minimax.
Séminaire de probabilités de Strasbourg, 12:342–363, 1978.

[BHLP14] Fernando GSL Brandao, Aram W Harrow, James R Lee, and Yuval Peres.
Adversarial hypothesis testing and a quantum stein's lemma for restricted measurements.
In *Proceedings of the 5th conference on Innovations in theoretical computer science*, pages 183–194, 2014.

[BJ95] Nader H Bshouty and Jeffrey C Jackson.
Learning dnf over the uniform distribution using a quantum example oracle.
In *Proceedings of the eighth annual conference on Computational learning theory*, pages 118–127, 1995.

[BLM20] M. Bun, R. Livni, and S. Moran.
An equivalence between private classification and online prediction.
In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 389–402, Los Alamitos, CA, USA, nov 2020. IEEE Computer Society.

[BNS+21] Raef Bassily, Kobbi Nissim, Adam Smith, Thomas Steinke, Uri Stemmer, and Jonathan Ullman.
Algorithmic stability for adaptive data analysis.
SIAM Journal on Computing, 50(3):STOC16–377, 2021.

[BO21] Costin Bădescu and Ryan O'Donnell.
Improved quantum data analysis.
In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1398–1411, 2021.

[BS16] Mark Bun and Thomas Steinke.
Concentrated differential privacy: Simplifications, extensions, and lower bounds.
In *Theory of Cryptography: 14th International Conference, TCC 2016-B, Beijing, China, October 31-November 3, 2016, Proceedings, Part I*, pages 635–658. Springer, 2016.

[BST14] Raef Bassily, Adam Smith, and Abhradeep Thakurta.
Private empirical risk minimization: Efficient algorithms and tight error bounds.
In *2014 IEEE 55th annual symposium on foundations of computer science*, pages 464–473. IEEE, 2014.

[BTT18] Raef Bassily, Om Thakkar, and Abhradeep Thakurta.
Model-agnostic private learning.
In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 7102–7112, Red Hook, NY, USA, 2018. Curran Associates Inc.

[BY23] Zongbo Bao and Penghui Yao.
Nearly optimal algorithms for testing and learning quantum junta channels.
arXiv preprint arXiv:2305.12097, 2023.

[Can22] Clément L. Canonne.
A short note on an inequality between kl and tv, 2022.

[Car20] Matthias C Caro.
Quantum learning boolean linear functions wrt product distributions.
Quantum Information Processing, 19(6):172, 2020.

[Car22] Matthias C Caro.
Learning quantum processes and hamiltonians via the pauli transfer matrix.
arXiv preprint arXiv:2212.04471, 2022.

[CCC19] Patrick J Coles, M Cerezo, and Lukasz Cincio.
Strong bound between trace distance and hilbert-schmidt distance for low-rank states.
Physical Review A, 100(2):022103, 2019.

[CCHL22a] Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, and Jerry Li.
The complexity of nisq.
arXiv preprint arXiv:2210.07234, 2022.

[CCHL22b] Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, and Jerry Li.
Exponential separations between learning with and without quantum memory.
In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 574–585. IEEE, 2022.

[CDE+23] Rachel Cummings, Damien Desfontaines, David Evans, Roxana Geambasu, Matthew Jagielski, Yangsibo Huang, Peter Kairouz, Gautam Kamath, Sewoong Oh, Olga Ohrimenko, et al.
Challenges towards the next frontier in privacy.
arXiv preprint arXiv:2304.06929, 2023.

[CDWE19] Arnaud Carignan-Dugas, Joel J Wallman, and Joseph Emerson.
Bounding the average gate fidelity of composite channels using the unitarity.
New Journal of Physics, 21(5):053016, 2019.

[CGL+20] Nai-Hui Chia, András Gilyén, Tongyang Li, Han-Hsuan Lin, Ewin Tang, and Chunhao Wang.
Sampling-Based Sublinear Low-Rank Matrix Arithmetic Framework for Dequantizing Quantum Machine Learning, page 387–400.
Association for Computing Machinery, New York, NY, USA, 2020.

[CGS02] Claude Crépeau, Daniel Gottesman, and Adam Smith.
Secure multi-party quantum computation.
In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 643–652, 2002.

[CHE+23] Matthias C Caro, Hsin-Yuan Huang, Nicholas Ezzell, Joe Gibbs, Andrew T Sornborger, Lukasz Cincio, Patrick J Coles, and Zoë Holmes.
Out-of-distribution generalization for learning quantum dynamics.
Nature Communications, 14(1):3751, 2023.

[CHI+23] Matthias C Caro, Marcel Hinsche, Marios Ioannou, Alexander Nietner, and Ryan Sweke.
Classical verification of quantum learning.
arXiv preprint arXiv:2306.04843, 2023.

[Cho75] Man-Duen Choi.
Completely positive linear maps on complex matrices.
Linear Algebra and its Applications, 10(3):285–290, 1975.

[CHTW04] Richard Cleve, Peter Hoyer, Benjamin Toner, and John Watrous.
Consequences and limits of nonlocal strategies.
In *Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.*, pages 236–249. IEEE, 2004.

[Chu32] Alonzo Church.
A set of postulates for the foundation of logic.
Annals of mathematics, pages 346–366, 1932.

[CJK+18] Graham Cormode, Somesh Jha, Tejas Kulkarni, Ninghui Li, Divesh Srivastava, and Tianhao Wang.
Privacy at scale: Local differential privacy in practice.
In *Proceedings of the 2018 International Conference on Management of Data*, pages 1655–1658, 2018.

[CLSZ95] Isaac L Chuang, Raymond Laflamme, Peter W Shor, and Wojciech H Zurek.
Quantum computers, factoring, and decoherence.
Science, 270(5242):1633–1635, 1995.

[CLW18] Nai-Hui Chia, Han-Hsuan Lin, and Chunhao Wang.
Quantum-inspired sublinear classical algorithms for solving low-rank linear systems, 2018.

[CMH17] Matthias Christandl and Alexander Müller-Hermes.
Relative entropy bounds on quantum, private and repeater capacities.
Commun. Math. Phys., 353(2):821–852, may 2017.

[CMS11a] Kamalika Chaudhuri, Claire Monteleoni, and Anand D Sarwate.
Differentially private empirical risk minimization.
Journal of Machine Learning Research, 12(3), 2011.

[CMS11b] Kamalika Chaudhuri, Claire Monteleoni, and Anand D. Sarwate.
Differentially private empirical risk minimization.

Journal of Machine Learning Research, 12(29):1069–1109, 2011.

[CN97] Isaac L Chuang and Michael A Nielsen.
Prescription for experimental determination of the dynamics of a quantum black box.
Journal of Modern Optics, 44(11-12):2455–2467, 1997.

[CNY23] Thomas Chen, Shivam Nadimpalli, and Henry Yuen.
Testing and learning quantum juntas nearly optimally.
In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1163–1185. SIAM, 2023.

[Cob64] Alan Cobham.
The intrinsic computational difficulty of functions.
In *Proc. 1964 Congress for Logic, Methodology, and the Philosophy of Science*, pages 24–30. North-Holland, 1964.

[Cov99] Thomas M Cover.
Elements of information theory.
John Wiley & Sons, 1999.

[CRK19] Jeremy Cohen, Elan Rosenfeld, and Zico Kolter.
Certified adversarial robustness via randomized smoothing.
In *international conference on machine learning*, pages 1310–1320. PMLR, 2019.

[CS96] A Robert Calderbank and Peter W Shor.
Good quantum error-correcting codes exist.
Physical Review A, 54(2):1098, 1996.

[Csi64] Imre Csiszár.
Eine informationstheoretische ungleichung und ihre anwendung auf beweis der ergodizitaet von markoffschen ketten.
Magyer Tud. Akad. Mat. Kutato Int. Koezl., 8:85–108, 1964.

[CSV+21] Marco Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J Coles.
Cost function dependent barren plateaus in shallow parametrized quantum circuits.
Nature communications, 12(1):1791, 2021.

[DCEL09] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine.
Exact and approximate unitary 2-designs and their application to fidelity estimation.
Physical Review A, 80(1):012304, 2009.

[DF18] Cynthia Dwork and Vitaly Feldman.
Privacy-preserving prediction.
In *Conference On Learning Theory*, pages 1693–1702. PMLR, 2018.

[DFH+15] Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Aaron Leon Roth.

Preserving statistical validity in adaptive data analysis.
In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 117–126, 2015.

[DGJ+20] Yfke Dulek, Alex B. Grilo, Stacey Jeffery, Christian Majenz, and Christian Schaffner.
Secure multi-party quantum computation with a dishonest majority.
In *Advances in Cryptology – EUROCRYPT 2020*, pages 729–758. Springer International Publishing, 2020.

[DHJB21] Alexander M Dalzell, Nicholas Hunter-Jones, and Fernando GSL Brandão.
Random quantum circuits transform local noise into global white noise.
arXiv preprint arXiv:2111.14907, 2021.

[DHJB22] Alexander M Dalzell, Nicholas Hunter-Jones, and Fernando GSL Brandão.
Random quantum circuits anticoncentrate in log depth.
PRX Quantum, 3(1):010333, 2022.

[DHL+21a] Yuxuan Du, Min-Hsiu Hsieh, Tongliang Liu, Dacheng Tao, and Nana Liu.
Quantum noise protects quantum classifiers against adversaries.
Physical Review Research, 3(2), may 2021.

[DHL+21b] Yuxuan Du, Min-Hsiu Hsieh, Tongliang Liu, Dacheng Tao, and Nana Liu.
Quantum noise protects quantum classifiers against adversaries.
Physical Review Research, 3(2), May 2021.

[DHL+22] Yuxuan Du, Min-Hsiu Hsieh, Tongliang Liu, Shan You, and Dacheng Tao.
Quantum differentially private sparse regression learning.
IEEE Transactions on Information Theory, 68(8):5217–5233, aug 2022.

[DJW13] John C. Duchi, Michael I. Jordan, and Martin J. Wainwright.
Local privacy, data processing inequalities, and statistical minimax rates, 2013.

[DMNS06] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith.
Calibrating noise to sensitivity in private data analysis.
In *Proceedings of the Third Conference on Theory of Cryptography*, TCC'06, page 265–284, Berlin, Heidelberg, 2006. Springer-Verlag.

[DPMRF23] Giacomo De Palma, Milad Marvian, Cambyse Rouzé, and Daniel Stilck França.
Limitations of variational quantum algorithms: a quantum optimal transport approach.
PRX Quantum, 4(1):010309, 2023.

[DPMTL21a] Giacomo De Palma, Milad Marvian, Dario Trevisan, and Seth Lloyd.
The quantum wasserstein distance of order 1.
IEEE Transactions on Information Theory, 67(10):6627–6643, 2021.

[DPMTL21b] Giacomo De Palma, Milad Marvian, Dario Trevisan, and Seth Lloyd.

The quantum wasserstein distance of order 1.
IEEE Transactions on Information Theory, 67(10):6627–6643, Oct 2021.

[DR14] Cynthia Dwork and Aaron Roth.
The algorithmic foundations of differential privacy.
9(3–4):211–407, August 2014.

[F+18] Richard P Feynman et al.
Simulating physics with computers.
Int. j. Theor. phys, 21(6/7), 2018.

[Far23] Farhad Farokhi.
Privacy against hypothesis-testing adversaries for quantum computing, 2023.

[FGG+23] Bill Fefferman, Soumik Ghosh, Michael Gullans, Kohdai Kuroiwa, and Kunal Sharma.
Effect of non-unital noise on random circuit sampling.
arXiv preprint arXiv:2306.16659, 2023.

[FGP21] Daniel Stilck França and Raul García-Patrón.
Limitations of optimization algorithms on noisy quantum devices.
Nature Physics, 17(11):1221–1227, oct 2021.

[FMHS22] Omar Fawzi, Alexander Müller-Hermes, and Ala Shayeghi.
A lower bound on the space overhead of fault-tolerant quantum computation.
2022.

[FQR22] Marco Fanizza, Yihui Quek, and Matteo Rosati.
Learning quantum processes without input control.
arXiv preprint arXiv:2211.05005, 2022.

[FS17] Vitaly Feldman and Thomas Steinke.
Generalization for adaptively-chosen estimators via stable median.
In *Conference on Learning Theory*, pages 728–757. PMLR, 2017.

[gdp] General Data Protection Regulation. Regulation (EU) 2016/679 of the European Parliament and of the European Council of 27 April 2016.

[GJ14] Gus Gutoski and Nathaniel Johnston.
Process tomography for unitary quantum channels.
Journal of Mathematical Physics, 55(3), 2014.

[GL22] Aravind Gollakota and Daniel Liang.
On the hardness of pac-learning stabilizer states with noise.
Quantum, 6:640, 2022.

[GLT18] András Gilyén, Seth Lloyd, and Ewin Tang.
Quantum-inspired low-rank stochastic regression with logarithmic dependence on the dimension, 2018.

[GMW87] Oded Goldreich, Silvio Micali, and Avi Wigderson.
How to solve any protocol problem.
In *Proc. of STOC*, 1987.

[GWOB19] Edward Grant, Leonard Wossnig, Mateusz Ostaszewski, and Marcello Benedetti.
An initialization strategy for addressing barren plateaus in parametrized quantum circuits.
Quantum, 3:214, 2019.

[HBM+21] Hsin-Yuan Huang, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, and Jarrod R. McClean.
Power of data in quantum machine learning.
Nature Communications, 12(1), may 2021.

[HCP22] Hsin-Yuan Huang, Sitan Chen, and John Preskill.
Learning to predict arbitrary quantum processes.
arXiv preprint arXiv:2210.14894, 2022.

[HCT+19] Vojtěch Havlíček, Antonio D Córcoles, Kristan Temme, Aram W Harrow, Abhinav Kandala, Jerry M Chow, and Jay M Gambetta.
Supervised learning with quantum-enhanced feature spaces.
Nature, 567(7747):209–212, 2019.

[HIN+23] M Hinsche, M Ioannou, A Nietner, J Haferkamp, Y Quek, D Hangleiter, J-P Seifert, J Eisert, and R Sweke.
One t gate makes distribution learning hard.
Physical Review Letters, 130(24):240602, 2023.

[Hir23] Christoph Hirche.
Benefits and detriments of noise in quantum classification.
2023.

[HKP20] Hsin-Yuan Huang, Richard Kueng, and John Preskill.
Predicting many properties of a quantum system from very few measurements.
Nature Physics, 16(10):1050–1057, 2020.

[HM23] Aram W Harrow and Saeed Mehraban.
Approximate unitary t-designs by short random quantum circuits using nearest-neighbor and long-range gates.
Communications in Mathematical Physics, pages 1–96, 2023.

[HOT81] Fumio Hiai, Masanori Ohya, and Makoto Tsukada.
Sufficiency, kms condition and relative entropy in von neumann algebras.
Pacific Journal of Mathematics, 96(1):99–109, 1981.

[HP91] Fumio Hiai and Dénes Petz.

The proper formula for relative entropy and its asymptotics in quantum probability.
Communications in mathematical physics, 143(1):99–114, 1991.

[HR16] Fumio Hiai and Mary Beth Ruskai.
Contraction coefficients for noisy quantum channels.
Journal of Mathematical Physics, 57(1), 2016.

[HRF22] Christoph Hirche, Cambyse Rouzé, and Daniel Stilck França.
On contraction coefficients, partial orders and approximation of capacities for quantum channels.
Quantum, 6:862, 2022.

[HRF23] Christoph Hirche, Cambyse Rouzé, and Daniel Stilck França.
Quantum differential privacy: An information theory perspective.
IEEE Transactions on Information Theory, 2023.

[HSCC22] Zoë Holmes, Kunal Sharma, Marco Cerezo, and Patrick J Coles.
Connecting ansatz expressibility to gradient magnitudes and barren plateaus.
PRX Quantum, 3(1):010313, 2022.

[HT23] Christoph Hirche and Marco Tomamichel.
Quantum renyi and f-divergences from integral representations.
arXiv preprint arXiv:2306.12343, 2023.

[HTY+23] Jhih-Cing Huang, Yu-Lin Tsai, Chao-Han Huck Yang, Cheng-Fang Su, Chia-Mu Yu, Pin-Yu Chen, and Sy-Yen Kuo.
Certified robustness of quantum classifiers against adversarial examples through quantum noise.
In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2023.

[Jam72] A. Jamiołkowski.
Linear transformations which preserve trace and positive semidefiniteness of operators.
Reports on Mathematical Physics, 3(4):275–278, 1972.

[JGM+23] Sofiene Jerbi, Casper Gyurik, Simon C Marshall, Riccardo Molteni, and Vedran Dunjko.
Shadows of quantum machine learning.
arXiv preprint arXiv:2306.00061, 2023.

[KB14] Diederik P Kingma and Jimmy Ba.
Adam: A method for stochastic optimization.
arXiv preprint arXiv:1412.6980, 2014.

[KBR16] Peter Kairouz, Keith Bonawitz, and Daniel Ramage.
Discrete distribution estimation under local privacy.

In *International Conference on Machine Learning*, pages 2436–2444. PMLR, 2016.

[Kea98a] Michael Kearns.
Efficient noise-tolerant learning from statistical queries.
Journal of the ACM (JACM), 45(6):983–1006, 1998.

[Kea98b] Michael Kearns.
Efficient noise-tolerant learning from statistical queries.
J. ACM, 45(6):983–1006, nov 1998.

[KEA+23] Youngseok Kim, Andrew Eddins, Sajant Anand, Ken Xuan Wei, Ewout Van Den Berg, Sami Rosenblatt, Hasan Nayfeh, Yantao Wu, Michael Zaletel, Kristan Temme, et al.
Evidence for the utility of quantum computing before fault tolerance.
Nature, 618(7965):500–505, 2023.

[KHL+23] Niraj Kumar, Jamie Heredge, Changhao Li, Shaltiel Eloul, Shree Hari Sureshbabu, and Marco Pistoia.
Expressive variational quantum circuits provide inherent privacy in federated learning.
arXiv preprint arXiv:2309.13002, 2023.

[KLN+11a] Shiva Prasad Kasiviswanathan, Homin K. Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith.
What can we learn privately?
SIAM J. Comput., 40(3):793–826, June 2011.

[KLN+11b] Shiva Prasad Kasiviswanathan, Homin K Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith.
What can we learn privately?
SIAM Journal on Computing, 40(3):793–826, 2011.

[KOV14] Peter Kairouz, Sewoong Oh, and Pramod Viswanath.
Extremal mechanisms for local differential privacy.
Advances in neural information processing systems, 27, 2014.

[KOV15] Peter Kairouz, Sewoong Oh, and Pramod Viswanath.
Secure multi-party differential privacy.
Advances in neural information processing systems, 28, 2015.

[KRS19] V Kanade, A Rocchetto, and S Severini.
Learning dnfs under product distributions via μ-biased quantum fourier sampling.
Quantum Information and Computation, 19(15&16), 2019.

[KRUDW08] Julia Kempe, Oded Regev, Falk Unger, and Ronald De Wolf.
Upper bounds on the noise threshold for fault-tolerant quantum computing.
In *International Colloquium on Automata, Languages, and Programming*, pages 845–856. Springer, 2008.

[KSW20] Sumeet Khatri, Kunal Sharma, and Mark M Wilde.
Information-theoretic aspects of the generalized amplitude-damping channel.
Physical Review A, 102(1):012401, 2020.

[LAG+19] Mathias Lecuyer, Vaggelis Atlidakis, Roxana Geambasu, Daniel Hsu, and Suman Jana.
Certified robustness to adversarial examples with differential privacy, 2019.

[LAT21] Yunchao Liu, Srinivasan Arunachalam, and Kristan Temme.
A rigorous and robust quantum speed-up in supervised machine learning.
Nature Physics, 17(9):1013–1017, 2021.

[LLD21] Weikang Li, Sirui Lu, and Dong-Ling Deng.
Quantum federated learning through blind quantum computing.
Science China Physics, Mechanics & Astronomy, 64(10), sep 2021.

[LR99] Andrew Lesniewski and Mary Beth Ruskai.
Monotone riemannian metrics and relative entropy on noncommutative probability spaces.
Journal of Mathematical Physics, 40(11):5702–5724, 1999.

[LSH+13] Nima Lashkari, Douglas Stanford, Matthew Hastings, Tobias Osborne, and Patrick Hayden.
Towards the fast scrambling conjecture.
Journal of High Energy Physics, 2013(4):1–33, 2013.

[LTD+22] Jonas Landman, Slimane Thabet, Constantin Dalyac, Hela Mhiri, and Elham Kashefi.
Classically approximating variational quantum machine learning with random fourier features.
In *The Eleventh International Conference on Learning Representations*, 2022.

[MAE+23] Antonio Anna Mele, Armando Angrisani, Jens Eisert, Soumik Ghosh, Sumeet Khatri, Yihui Quek, and Daniel Stilck França.
Noise-induced absence of barren plateaus: Non-unital noise can be a friendly foe.
2023.

[MBS+18] Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven.
Barren plateaus in quantum neural network training landscapes.
Nature communications, 9(1):4812, 2018.

[MdW13] Ashley Montanaro and Ronald de Wolf.
A survey of quantum property testing.
arXiv preprint arXiv:1310.2035, 2013.

[Mei18] Sebastian Meiser.
Approximate and probabilistic differential privacy definitions.

Cryptology ePrint Archive, 2018.

[Mel23] Antonio Anna Mele.
Introduction to haar measure tools in quantum information: A beginner's tutorial.
arXiv preprint arXiv:2307.08956, 2023.

[MGN20] Abhijeet Melkani, Clemens Gneiting, and Franco Nori.
Eigenstate extraction with neural-network tomography.
Phys. Rev. A, 102:022412, Aug 2020.

[MH11] Milán Mosonyi and Fumio Hiai.
On the quantum rényi relative entropies and related capacity formulas.
IEEE Transactions on Information Theory, 57(4):2474–2487, 2011.

[Mir17] Ilya Mironov.
Rényi differential privacy.
In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*. IEEE, aug 2017.

[MKW21] Carlos Ortiz Marrero, Mária Kieferová, and Nathan Wiebe.
Entanglement-induced barren plateaus.
PRX Quantum, 2(4):040316, 2021.

[MLDS+13] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel.
On quantum rényi entropies: A new generalization and some properties.
Journal of Mathematical Physics, 54(12):122203, dec 2013.

[MO10] Ashley Montanaro and Tobias J Osborne.
Quantum boolean functions.
Chicago Journal OF Theoretical Computer Science, 1:1–45, 2010.

[MT07] Frank McSherry and Kunal Talwar.
Mechanism design via differential privacy.
In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 94–103, 2007.

[Nap22] John Napp.
Quantifying the barren plateau phenomenon for a model of unstructured variational ans\"{a} tze.
arXiv preprint arXiv:2203.06174, 2022.

[NC10] Michael A Nielsen and Isaac L Chuang.
Quantum computation and quantum information.
Cambridge university press, 2010.

[NGW23] Theshani Nuradha, Ziv Goldfeld, and Mark M. Wilde.
Quantum pufferfish privacy: A flexible privacy framework for quantum systems, 2023.

[NIS+23] Alexander Nietner, Marios Ioannou, Ryan Sweke, Richard Kueng, Jens Eisert, Marcel Hinsche, and Jonas Haferkamp.
On the average-case complexity of learning output distributions of quantum circuits.
arXiv preprint arXiv:2305.05765, 2023.

[NS06] Arvind Narayanan and Vitaly Shmatikov.
How to break anonymity of the netflix prize dataset.
arXiv preprint cs/0610105, 2006.

[NS07] Arvind Narayanan and Vitaly Shmatikov.
How to break anonymity of the netflix prize dataset, 2007.

[O'D21] Ryan O'Donnell.
Analysis of boolean functions.
arXiv preprint arXiv:2105.10386, 2021.

[Ohm09] Paul Ohm.
Broken promises of privacy: Responding to the surprising failure of anonymization.
UCLA l. Rev., 57:1701, 2009.

[ON05] Tomohiro Ogawa and Hiroshi Nagaoka.
Strong converse and stein's lemma in quantum hypothesis testing.
In *Asymptotic Theory of Quantum Statistical Inference*, pages 28–42. WORLD SCIENTIFIC, feb 2005.

[PAE+16] Nicolas Papernot, Martín Abadi, Úlfar Erlingsson, Ian Goodfellow, and Kunal Talwar.
Semi-supervised knowledge transfer for deep learning from private training data.
In *International Conference on Learning Representations*, 2016.

[PC18] Chae-Yeun Park and Jaeyoon Cho.
Correlations in local measurements and entanglement in many-body systems.
Physical Review A, 98(1), jul 2018.

[PJSPP21] Aidan Pellow-Jarman, Ilya Sinayskiy, Anban Pillay, and Francesco Petruccione.
A comparison of various classical optimizers for a variational quantum linear solver.
Quantum Information Processing, 20(6):202, 2021.

[PP21] Jason L Pereira and Stefano Pirandola.
Bounds on amplitude-damping-channel discrimination.
Phys. Rev. A, 103(2):022610, 2021.

[PPV10] Yury Polyanskiy, H. Vincent Poor, and Sergio Verdu.
Channel coding rate in the finite blocklength regime.
IEEE Transactions on Information Theory, 56(5):2307–2359, 2010.

[Pre18a] John Preskill.
Quantum computing in the nisq era and beyond.

Quantum, 2:79, 2018.

[Pre18b] John Preskill.
Quantum Computing in the NISQ era and beyond.
Quantum, 2:79, August 2018.
Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften.

[Pre23] John Preskill.
Quantum computing 40 years later.
In *Feynman Lectures on Computation*, pages 193–244. CRC Press, 2023.

[PY10] Sinno Jialin Pan and Qiang Yang.
A survey on transfer learning.
IEEE Transactions on Knowledge and Data Engineering, 22(10):1345–1359, 2010.

[PZ22] Feng Pan and Pan Zhang.
Simulation of quantum circuits using the big-batch tensor network method.
Physical Review Letters, 128(3):030501, 2022.

[QFK+22] Yihui Quek, Daniel Stilck França, Sumeet Khatri, Johannes Jakob Meyer, and Jens Eisert.
Exponentially tighter bounds on limitations of quantum error mitigation, 2022.

[Rag03] Maxim Raginsky.
Scaling and renormalization in fault-tolerant quantum computers.
Quant. Inf. Proc., 2:249–258, 2003.

[Raz03] A. A. Razborov.
An upper bound on the threshold quantum decoherence rate.
2003.

[RCA+22] Daniil Rabinovich, Ernesto Campos, Soumik Adhikary, Ekaterina Pankovets, Dmitry Vinichenko, and Jacob Biamonte.
On the gate-error robustness of variational quantum algorithms.
arXiv preprint arXiv:2301.00048, 2022.

[RF21] Cambyse Rouzé and Daniel Stilck França.
Learning quantum many-body systems from a few copies.
arXiv preprint arXiv:2107.03333, 2021.

[RSP+21] Ryan Rogers, Subbu Subramaniam, Sean Peng, David Durfee, Seunghyun Lee, Santosh Kumar Kancha, Shraddha Sahay, and Parvez Ahammad.
Linkedin's audience engagements api: A privacy preserving data analytics system at scale.
Journal of Privacy and Confidentiality, 11(3), 2021.

[RWZ22] Cambyse Rouzé, Melchior Wirth, and Haonan Zhang.

Quantum talagrand, kkl and friedgut's theorems and the learnability of quantum boolean functions, 2022.

[Sch21] Maria Schuld.
Supervised quantum machine learning models are kernel methods.
arXiv preprint arXiv:2101.11020, 2021.

[SEM23] Franz J Schreiber, Jens Eisert, and Johannes Jakob Meyer.
Classical surrogates for quantum learning models.
Physical Review Letters, 131(10):100803, 2023.

[SFGP21] Daniel Stilck França and Raul Garcia-Patron.
Limitations of optimization algorithms on noisy quantum devices.
Nature Physics, 17(11):1221–1227, 2021.

[Sho97] Peter W. Shor.
Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer.
SIAM Journal on Computing, 26(5):1484–1509, Oct 1997.

[SK19] Maria Schuld and Nathan Killoran.
Quantum machine learning in feature hilbert spaces.
Physical review letters, 122(4):040504, 2019.

[SKCC20] Kunal Sharma, Sumeet Khatri, Marco Cerezo, and Patrick J Coles.
Noise resilience of variational quantum compiling.
New Journal of Physics, 22(4):043006, 2020.

[SMT17] Makhamisa Senekane, Mhlambululi Mafu, and Benedict Molibeli Taele.
Privacy-preserving quantum machine learning using differential privacy.
In *2017 IEEE AFRICON*, pages 1432–1435. IEEE, 2017.

[SSBD14] Shai Shalev-Shwartz and Shai Ben-David.
Understanding machine learning: From theory to algorithms.
Cambridge university press, 2014.

[SV16] Igal Sason and Sergio Verdú.
f-divergence inequalities.
IEEE Transactions on Information Theory, 62(11):5973–6006, 2016.

[SW12] Naresh Sharma and Naqueeb Ahmad Warsi.
On the strong converses for the quantum channel capacity theorems.
arXiv preprint arXiv:1205.1712, 2012.

[Tan19] Ewin Tang.
A quantum-inspired classical algorithm for recommendation systems.

In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 217–228, New York, NY, USA, 2019. Association for Computing Machinery.

[Tan21] Ewin Tang.
Quantum principal component analysis only achieves an exponential speedup because of its state preparation assumptions.
Physical Review Letters, 127(6), Aug 2021.

[Tom15] Marco Tomamichel.
Quantum information processing with finite resources: mathematical foundations, volume 5.
Springer, 2015.

[Tur36] Alan Turing.
On computable numbers, with an application to the entscheidungsproblem.
J. of Math, 58(345-363):5, 1936.

[Tur52] AM Turing.
The chemical basis of morphogenesis.
Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, pages 37–72, 1952.

[TWH22] Supanut Thanasilp, Samson Wang, and Zoë Holmes.
Exponential concentration and untrainability in quantum kernel methods.
arXiv preprint arXiv:2208.11060, 2022.

[Ull17] Jonathan Ullman.
Cs7880: Rigorous approaches to data privacy, 2017.

[Vad17] Salil Vadhan.
The Complexity of Differential Privacy, pages 347–450.
Springer, Yehuda Lindell, ed., 2017.

[vEH14] Tim van Erven and Peter Harremoes.
Ré nyi divergence and kullback-leibler divergence.
IEEE Transactions on Information Theory, 60(7):3797–3820, jul 2014.

[VNB+66] John Von Neumann, Arthur W Burks, et al.
Theory of self-reproducing automata.
IEEE Transactions on Neural Networks, 5(1):3–14, 1966.

[War65] Stanley L. Warner.
Randomized response: A survey technique for eliminating evasive answer bias.
Journal of the American Statistical Association, 60(309):63–69, 1965.

[Wat18] John Watrous.

The theory of quantum information.
Cambridge university press, 2018.

[WCY23] William M Watkins, Samuel Yen-Chi Chen, and Shinjae Yoo.
Quantum machine learning with differential privacy.
Scientific Reports, 13(1):2453, 2023.

[WD23] Chirag Wadhwa and Mina Doosti.
Learning quantum processes with quantum statistical queries, 2023.

[WFC+21] Samson Wang, Enrico Fontana, Marco Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J Coles.
Noise-induced barren plateaus in variational quantum algorithms.
Nature communications, 12(1):6961, 2021.

[WGHF15] Joel Wallman, Chris Granade, Robin Harper, and Steven T Flammia.
Estimating the coherence of noise.
New Journal of Physics, 17(11):113020, 2015.

[Wil13] Mark M Wilde.
Quantum information theory.
Cambridge university press, 2013.

[WLF16] Yu-Xiang Wang, Jing Lei, and Stephen E. Fienberg.
Learning with differential privacy: Stability, learnability and the sufficiency and necessity of erm principle.
Journal of Machine Learning Research, 17(183):1–40, 2016.

[XZA+23] Zheng Xu, Yanxiang Zhang, Galen Andrew, Christopher A Choquette-Choo, Peter Kairouz, H Brendan McMahan, Jesse Rosenstock, and Yuanbo Zhang.
Federated learning of gboard language models with differential privacy.
arXiv preprint arXiv:2305.18465, 2023.

[Yao86] Andrew Chi-Chih Yao.
How to generate and exchange secrets (extended abstract).
In *27th Annual Symposium on Foundations of Computer Science, Toronto, Canada, 27-29 October 1986*, pages 162–167. IEEE Computer Society, 1986.

[YH20] Yuuya Yoshida and Masahito Hayashi.
Classical mechanism is optimal in classical-quantum differentially private mechanisms.
In *2020 IEEE International Symposium on Information Theory (ISIT)*, pages 1973–1977. IEEE, 2020.

[YLLP23] Hang Yang, Xunbo Li, Zhigui Liu, and Witold Pedrycz.
Improved differential privacy noise mechanism in quantum machine learning.

IEEE Access, 11:50157–50164, 2023.

[Yos21] Yuuya Yoshida.
Mathematical comparison of classical and quantum mechanisms in optimization under local differential privacy, 2021.

[YW23] Nengkun Yu and Tzu-Chieh Wei.
Learning marginals suffices!
arXiv preprint arXiv:2303.08938, 2023.

[ZY17a] Li Zhou and Mingsheng Ying.
Differential privacy in quantum computation.
In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pages 249–262, 2017.

[ZY17b] Li Zhou and Mingsheng Ying.
Differential privacy in quantum computation.
In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pages 249–262, 2017.