THEORY OF PRIVACY AND TESTING IN A QUANTUM WORLD

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Theshani Nuradha Piliththuwasam Gallage

August 2025

THEORY OF PRIVACY AND TESTING IN A QUANTUM WORLD

Theshani Nuradha Piliththuwasam Gallage, Ph.D.

Cornell University 2025

With the surging interest in quantum and hybrid classical-quantum systems, ensuring the privacy of generated quantum data is essential. Quantum differential privacy (QDP) has been introduced to ensure privacy for quantum states. However, the versatility of QDP is limited. We introduce a flexible privacy framework for quantum systems termed as Quantum PP (QPP). We show that QPP is captured exactly by an information-spectrum divergence, endowing the latter with its first operational interpretation. Then this divergence is used to study properties of QPP mechanisms, showcasing distinctions that arise in the quantum setting compared to the classical setting. We introduce the first algorithms that can be used to audit for privacy given a mechanism, which facilitates new approaches with the access to quantum devices. We analyse how the QPP framework provides better privacy-utility tradeoffs with the flexibility to incorporate application-specific criteria. To provide further insights on QPP, we comprehensively study measured hockey-stick divergences that find operational interpretation in the QPP framework as optimal privacy parameters.

Studying statistical problems under privacy constraints is vital for understanding the price that we have to pay to ensure privacy. To this end, the contraction of statistical measures and divergences under privacy constraints is an important technical tool. However, in the quantum setting, this area of research is largely unexplored even for fundamental statistical tasks. We characterize the contraction of quantum divergences under a local variant of quantum privacy,

which enables this field of study. We completely characterize the privatized contraction coefficient of the trace distance, among others. Next, we utilize the information-theoretic tools developed to study statistical tasks under privacy constraints. To this end, we characterize the cost of privacy in quantum hypothesis testing and learning expectation of observables while ensuring privacy of quantum states.

In sum, this thesis lays a theoretical foundation for ensuring the privacy of quantum data by introducing flexible privacy frameworks tailored to quantum systems. It also evaluates the tradeoffs involved in maintaining privacy during the testing and learning of properties of quantum data, offering information-theoretic tools to study statistical tasks under privacy constraints.

BIOGRAPHICAL SKETCH

Theshani Nuradha Piliththuwasam Gallage completed her primary and high school studies from Olcott Model School and Sujatha Vidyalaya, Matara, Sri Lanka. She joined the University of Moratuwa, Sri Lanka, for her undergraduate studies in 2014, after the nation-wide university entrance examination. She graduated as the Gold Medalist of the Department of Electronic and Telecommunication Engineering in 2018.

Her research journey began with a research internship in the Robotics and Innovation Laboratory at Singapore University of Technology and Design in 2017. After her undergraduate graduation, she joined the Department of Electronic and Telecommunication Engineering, University of Moratuwa, as a Lecturer on Contract in 2019. In 2021, she started her Ph.D. journey in the School of Electrical and Computer Engineering at Cornell University.

Her research interests include classical and quantum information theory, privacy, and statistical learning. During her Ph.D., she is advised by Prof. Mark M. Wilde and Prof. Ziv Goldfeld. She was recognized as a Trailblazing Young Researcher by the Division of Engineering and Applied Sciences at Caltech in 2024. She was also the recipient of the Robert Mozia Graduate Distinguished Service Award by the College of Engineering at Cornell University in 2025.

To all the beautiful souls who inspired me and believed in me!

ACKNOWLEDGEMENTS

My Ph.D. journey was quite non-linear, with several twists and turns that made me who I am today. Many wonderful human beings supported, inspired, and encouraged me while playing a huge role in my life during these four and a half years away from my comfort zone.

I'm forever thankful to the free education that I received from the Ministry of Education, Sri Lanka, for 17 continuous years (13 school +4 undergrad), for creating the foundation for my academic journey. For Cornell University, for offering me the opportunity to meet and learn from wonderful people whom I would not have met otherwise. I would like to thank all my teachers, mentors, instructors, and lecturers who have shaped me for all these years, since this is a culmination of all of your efforts in not only teaching me the curriculum, most importantly, the growth mindset, and the overall journey of LIFE.

First, I would like to thank my Ph.D. advisors and committee with whom I got to work closely and draw inspiration to craft my academic journey during my PhD and beyond. To this end, I thank Ziv Goldfeld for introducing me to research projects in the space of information theory and privacy, where I worked extensively during my PhD. I also thank him for providing me with constructive feedback on my scientific writing and research methodologies. After one and a half years into my PhD, I met Mark Wilde, which was a turning point in my academic journey, where my research expanded to quantum information theory. I'm grateful to him for believing in me and supporting me even when I was completely new to the field of quantum information science. To my Ph.D. advisors: Thank you for all the academic advice and feedback that you gave me to develop me into who I am today. I will definitely look up to you when I progress through this journey, to a distance that it will take me. I'm also grateful

to Jayadev Acharya and Aaron B. Wagner for being in my PhD committee and providing me with feedback on my research projects and for their continued support during my academic journey at Cornell.

I am deeply grateful to all my collaborators with whom I had the privilege to work and tackle mathematical problems together. Each of them brought unique perspectives, experiences, and insights that enriched my understanding and introduced me to new research directions and concepts. Collaborating on research problems, particularly those I grew to enjoy the most, has been a highlight of my academic journey, as it fostered enlightening discussions and exchange of ideas. In this regard, my heartfelt thanks go to Hao-Chung Cheng, Jenny Chen, Nilanjana Datta, Ian George, Kristjan Greenewald, Christoph Hirche, Zoë Holmes, Felix Leditzky, Nana Liu, Ivy Luo, Hemant Mishra, Dhrumil Patel, Galen Reeves, Soorya Rethinasamy, Robert Salzmann, Vishal Singh, Kathie Wang, and Hanna Westerheim. Additionally, I am especially thankful to Mark and Ziv for connecting me with these incredible collaborators and for their continued support in these collaborations.

My family away from home played a huge role in my emotional stability and in creating memories that I will cherish for my entire lifetime. My roommates, the first human beings that I met in person in Ithaca, Manushi Trivedi and Sonali Uppal, thank you a lot for being there for me and supporting me in being the usual me in a completely strange environment after Flora and Fauna on the way. Also, for my current roommate, Hansadi Jayamaha, for continuing that emotional support and being there for me to talk about anything and everything. Tharushi Jayasekara for our Hasbrouck memories and always cheering us to do our best. I'm forever grateful to all four of you for being sisters to me when I was far away from my comfort zone, where even my biological

sister was a little bit jealous of. I also like to remind my first friend from ECE, whom I first met on social media, before even coming to Cornell- Kiran Rokade- for being there to share experiences as an ECE Ph.D. student, and also for celebrating our birthdays together. I would also like to thank the extended Sri Lankan community at Cornell (Hansadi, Tharushi, Mihili, Upekha, Danushka, Chathura ayya, Pubudu ayya, Hasindu, Vishwa, Faheem uncle, Soosan aunty, Prasanna ayya, Nidharshi akka, Gajaba, Sachille ayya), Ithaca gang (Manushi, Kiran, Parag, Anurag, Apurva, Faahar) for the lovely moments that we cherished together in cultural celebrations, board-game nights, to in-state and out-of-state trips. I cannot forget my labmates and colleagues at work who choose to be friends. So a huge shout out to the Quantum Information Theory research group (Aby, Dhrumil, Hami, Kaiyuan, Michele, Soorya, Vishal) and Information Theory gang (Yang, Sharang, Haiyun, Carrie), for making the lab/office time joyful with both technical and philosophical discussions from work to life. I would also like to thank my friends from the University of Moratuwa and the ENTC family for cheering me up whenever we talk and meet.

Finally, my deepest thanks and heartfelt love go to my family for their unwavering support throughout my life and at every step of my journey. Amma (Deepika Jinasena) and Thaththa (P. G. Chandrarathna), for always believing in me and granting me the freedom to explore and discover who I am and what I am passionate about. Your constant love and care have been my greatest strength, helping me rise again whenever I fall. My sister (Charani Hasara), for always being there for me, encouraging me, and offering advice from her maturity as my younger sister. My grandmother, for always keeping me in her prayers. My husband (Pubudu Milan), for his love and patience with me during our mostly long-distance relationship, and for supporting me wholeheartedly in

my decisions.

This journey would not have been as joyful and meaningful without these incredible individuals who have stood by me, both physically and from miles away. To everyone who encouraged me and even shared a smile with me when I needed it the most, thank you for being part of my Ph.D. journey at Cornell from 2021-25.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

With the exponential increase in personal data shared online and recent advancements in data mining techniques, privacy concerns have become more pressing than ever. Statistical privacy frameworks seek to address these threats in a principled manner, subject to formal guarantees [56]. Differential privacy (DP) [40] is a popular framework that preserves the privacy of individual records while enabling aggregate queries about a database (see Definition 3). However, DP only deals with one type of private information (individual records modeled by rows of the database) and does not allow to incorporate domain knowledge into the framework. To address these limitations, a versatile generalization of DP called Pufferfish Privacy (PP) was proposed in [77].

PP allows for customizing which information is regarded as private and explicitly integrates distributional assumptions into the definition [77, 121, 143] (see Definition 4). PP has found use in several applications, including smart metering [17, 74] and trajectory monitoring with location tracking [81, 141]. Information-theoretic formulations of classical DP and PP have been proposed in [27] and [92], respectively.

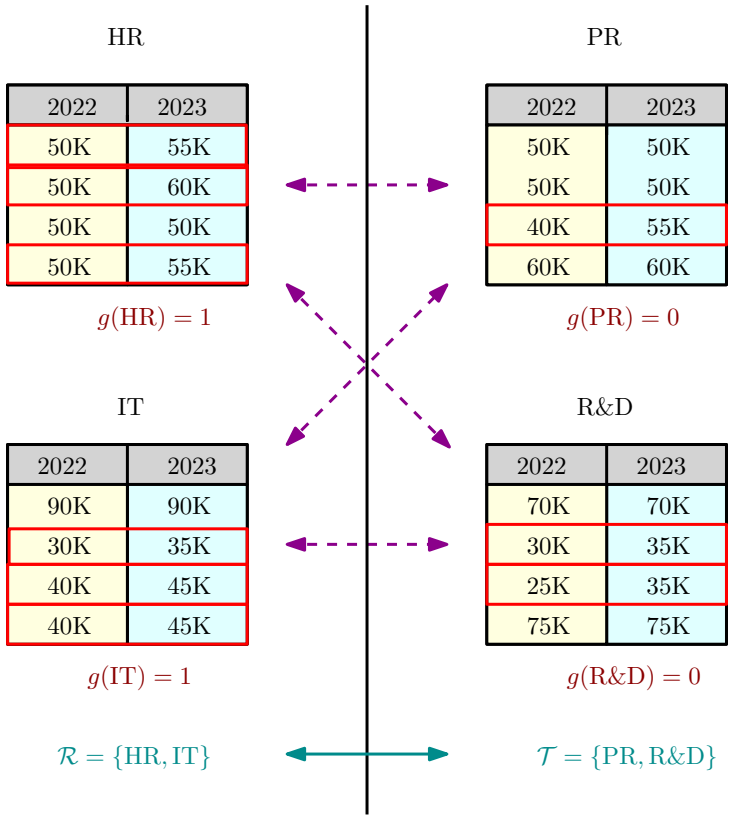


Figure 1.1: 2022-2033 salary data in four departments: HR, IT, PR, R&D. The goal is to publish the average 2023 salary in each department (the average of the blue cells) while hiding whether the number raises (marked by red frames) is ≤ 2 corresponding to $g(\cdot) = 0$ or > 2 corresponding to $g(\cdot) = 1$. The average 2022 salaries (yellow cells) are public knowledge [92, Figure 1]

.

1.1.1 Pufferfish Privacy

Let us study the following example (with classical data) to identify the distinctions between DP and PP frameworks given in [92, Figure 1]. Consider salary data from 2022-2023 at a company with four departments: HR, IT, PR, and R&D. The company wants to publish the average 2023 salary in each department while concealing whether more or less than m employees got a raise. The average salaries from 2022 are publicly available. More formally, the goal is to publish $f(x) = \frac{1}{n}\sum_{i=1}^{n} x(i, 2023)$ while privatizing whether $g(x) = \mathbb{1}_{\mathcal{A}_m}$, where

$\mathcal{A}_m = \{|\{i : x(i, 2022) < x(i, 2023)\}| > m\}$, $x \in \mathcal{X} := \{\text{HR}, \text{IT}, \text{PR}, \text{R\&D}\}$, and $x(i, j)$ is the salary of the ith employee during year j in department x. The average salary from 2022, i.e., $w(x) = \frac{1}{n} \sum_{i=1}^{n} x(i, 2022)$), is public knowledge. See Fig. 1.1 for an instance of the described scenario ($n = 4$ and $m = 2$).

DP operates by making any pair of neighboring databases indistinguishable, with the definition of neighbors being up to the privacy mechanism designer. In the scenario above, one may apply a DP-based approach by pairing as neighbors every two departments between which there is a difference in the function value g (whether the number of employees getting a raise is more than m). For the example from Fig. 1.1, this amounts to the set of pairs $\{(\text{HR}, \text{PR}), (\text{HR}, \text{R\&D}), (\text{IT}, \text{PR}), (\text{IT}, \text{R\&D})\}$ (marked by the dashed purple arrows in the figure). However, by following this approach, we guarantee a stricter privacy requirement than necessary. Indeed, upon observing the privatized version of the published query f and assuming w is publicly known, we only need to make the sets $g^{-1}(0) = \{\text{PR}, \text{R\&D}\}$ and $g^{-1}(1) = \{\text{HR}, \text{IT}\}$ indistinguishable (marked by the solid dark cyan arrow in Fig. 1.1). The benefit of targeting this relaxed notion of privacy is that it enables us to achieve improved accuracy and utility. The PP framework is designed to do just that, by enabling full customization of the events that are regarded as private. In addition, PP allows integrating into the framework domain knowledge on the distribution over databases; by considering the set of all possible distributions, this reduces to the worst-case requirement of DP.

1.1.2 Privacy for Quantum Data

The surging interest in quantum and hybrid classical-quantum systems has sparked a parallel interest in their application for decision-making tasks. With that, it is essential to ensure the privacy of quantum states in a similar vein to classical data. Quantum differential privacy (QDP) has been recently introduced to ensure privacy for quantum states [1, 62, 144]. However, as explained in the classical setting, the versatility of QDP is limited.

In this thesis, we work towards designing flexible privacy frameworks for quantum systems while incorporating features of quantum systems and identifying whether these privacy frameworks satisfy desired properties.

In Chapter 3, we propose a versatile privacy framework for quantum systems, termed quantum pufferfish privacy (QPP). Inspired by classical pufferfish privacy, our formulation generalizes and addresses limitations of quantum differential privacy by offering flexibility in specifying private information, feasible measurements, and domain knowledge. We show that QPP can be equivalently formulated in terms of the Datta-Leditzky information spectrum divergence, thus providing the first operational interpretation thereof. We reformulate this divergence as a semi-definite program and derive several properties of it, which are then used to prove convexity, composability, and post-processing of QPP mechanisms. Parameters that guarantee QPP of the depolarization mechanism are also derived. We analyze the privacy-utility tradeoff of general QPP mechanisms and, again, study the depolarization mechanism as an explicit instance. The QPP framework is then applied to privacy auditing for identifying privacy violations via a hypothesis testing pipeline that leverages quantum algorithms.

Chapter 3 is based on [95]:

T. Nuradha, Z. Goldfeld, and M. M. Wilde, "Quantum Pufferfish Privacy-A Flexible Privacy Framework for Quantum Systems", in *IEEE Transactions on Information Theory*, August 2024, doi: 10.1109/TIT.2024.3404927.

1.1.3 Information-Theoretic Quantities to Analyze Privacy

In Chapter 3, we show that a special class of QPP framework, when the adversary is allowed to use all possible measurements, can be equivalently formulated in terms of the Datta-Leditzky information spectrum divergence. This connection provides approaches to prove that the QPP framework satisfies the desired properties. This approach cross-fertilizes two fields, namely quantum information theory and privacy. In particular, this enables identifying operational interpretations of quantum information-theoretic quantities (the first operational interpretation of Datta-Leditzky divergence) and, in return, provides interesting behaviours of operational tasks such as privacy (e.g., how privacy degrades when composing multiple quantum private mechanisms).

In Chapter 4, we expand the connection between quantum information-theoretic tools and QPP further. The hockey-stick divergence is a fundamental quantity characterizing several statistical privacy frameworks that ensure privacy for classical and quantum data. In such quantum privacy frameworks, the adversary is allowed to perform all possible measurements. However, in practice, there are typically limitations to the set of measurements that can be performed. To this end, we comprehensively analyze the measured hockey-stick divergence under several classes of practically relevant measurement classes.

We prove several of its properties, including data processing and convexity. We show that it is efficiently computable by semi-definite programming for some classes of measurements and can be analytically evaluated for Werner and isotropic states. Notably, we show that the measured hockey-stick divergence characterizes optimal privacy parameters in the QPP framework. With this connection and the developed technical tools, we enable methods to quantify and audit privacy for several practically relevant settings. Lastly, we introduce the measured hockey-stick divergence of channels and explore its applications in ensuring privacy for channels.

Chapter 4 is based on [97]:

> **T. Nuradha**, V. Singh, M. M. Wilde. "Measured Hockey-Stick Divergence and its Applications to Quantum Pufferfish Privacy". arXiv:2501.12359 (Accepted to International Symposium of Information Theory- ISIT 2025).

1.1.4 Statistical Tasks under Privacy Constraints

A quantum generalized divergence by definition satisfies the data-processing inequality; as such, the relative decrease in such a divergence under the action of a quantum channel is at most one. This relative decrease is formally known as the contraction coefficient of the channel and the divergence. Interestingly, there exist combinations of channels and divergences for which the contraction coefficient is strictly less than one. Furthermore, understanding the contraction coefficient is fundamental for the study of statistical tasks under privacy constraints. However, finding exact characterizations of contraction coefficients of quantum divergences and studying statistical tasks under quantum privacy

constraints are still largely unexplored.

To this end, in Chapter 5, we establish upper bounds on contraction coefficients for the hockey-stick divergence under privacy constraints, where privacy is quantified with respect to the quantum local differential privacy (QLDP) framework, and we fully characterize the contraction coefficient for the trace distance under privacy constraints. Next, in Chapter 6, we apply our findings to establish bounds on the sample complexity of quantum hypothesis testing under privacy constraints and learning expectation values with privatized quantum states. Furthermore, we study various scenarios in which the sample complexity bounds are tight, while providing order-optimal quantum channels that achieve those bounds. With this, we characterize the cost of privacy in the setting of hypothesis testing and learning.

Chapter 5 and Chapter 6 are primarily based on [99]:

> **T. Nuradha** and M. M. Wilde, "Contraction of Private Quantum Channels and Private Quantum Hypothesis Testing", in *IEEE Transactions on Information Theory*, March 2025, doi: 10.1109/TIT.2025.3527859.

1.2 Other Ph.D. Research Projects

In addition to the research publications featured in this thesis in detail, I also co-authored the following papers during my doctoral studies.

1. **T. Nuradha**, and Z. Goldfeld, "Pufferfish Privacy: An Information-Theoretic Study", in *IEEE Transactions on Information Theory, Nov. 2023*, doi: 10.1109/TIT.2023.3296288. [93]

   To address the limitations of differential privacy, Pufferfish Privacy (PP) was proposed, yet its flexibility comes with additional challenges in analysis and deriving mechanisms. We circumvent this impasse by proposing a new structured PP framework along with a natural information-theoretic formulation, which lends itself well for analysis and enables devising mechanisms. We termed this information-theoretic formulation as mutual information PP, in terms of conditional mutual information between the mechanism and the secret, given the public information.

2. **T. Nuradha** and M. M. Wilde, "Fidelity-Based Smooth-Min-Relative Entropy: Properties and Applications", in IEEE Transactions on Information Theory, June 2024, doi: 10.1109/TIT.2024.3378590. [98]

   We comprehensively study an information-theoretic quantity referred to as smooth min-relative entropy, for which the smoothing is based on a fidelity constraint. We derive properties, including data processing and its second-order asymptotics. Furthermore, we expand these second-order asymptotic findings to a large class of smoothed Rényi divergences, while showcasing a unification of them into two main variants. We found the use of this fidelity-based divergence in general resource theories when the

target state is impure, with a particular focus on randomness distillation.

3. **T. Nuradha**, H. K. Mishra, F. Leditzky, and M. M. Wilde, "Multivariate Fidelities", in Journal of Physics A: Mathematical and Theoretical, April 2025, doi: 10.1088/1751-8121/adc645. [96]

   Bivariate classical fidelity (also known as Bhattacharya distance) is a widely used measure of similarity between two probability distributions. There exist a few extensions of the notion of similarity to more than two probability distributions. Hitherto, quantum generalizations of multivariate classical fidelities have not been systematically explored. In this work, we introduce several multivariate quantum fidelities, show that they satisfy several desirable properties, and provide operational interpretations for some of them.

4. H. C. Cheng, N. Datta, N. Liu, **T. Nuradha**, R. Salzmann, and M. M. Wilde, "An invitation to the sample complexity of quantum hypothesis testing", arXiv:2403.17868 (Accepted to npj Quantum Information, March 2025). [22]

   In this work, we study the sample complexity of quantum hypothesis testing (QHT), wherein the goal is to determine the minimum number of samples needed to reach a desired error probability. We characterize the sample complexity of binary QHT in the symmetric and asymmetric settings, and we provide bounds on the sample complexity of multiple QHT. We also outline how the sample complexity of QHT is relevant to a broad swathe of research areas and can enhance understanding of many fundamental concepts, including quantum algorithms for simulation and search, and quantum learning and classification.

5. Z. Goldfeld, K. Greenewald, **T. Nuradha**, and G. Reeves, "k-Sliced mu-

tual information: a quantitative study of scalability with dimension", In Conference on Neural Information Processing Systems (NeurIPS), New Orleans, USA, 2022. [52]

Mutual information (MI) is a fundamental measure of dependence between random variables. Sliced MI (SMI) was introduced as a surrogate dependence measure that preserves much of the classic structure of MI while being more scalable for computation and estimation in high dimensions. In this work, we provide a multifaceted account of how SMI itself and estimation rates thereof depend on the ambient dimension, under a broader framework termed k-SMI, which considers projections to k-dimensional subspaces. We derive sharp bounds on the error of Monte Carlo (MC)-based estimates of k-SMI and combine the MC integrator with the neural estimation framework to provide an end-to-end k-SMI estimator, for which optimal convergence rates are established.

6. V. Singh, **T. Nuradha**, M. M. Wilde. "Extendible quantum measurements and limitations on classical communication". arXiv:2412.18556 (Accepted to International Symposium of Information Theory- ISIT 2025). [120]

   In this work, we generalize the framework of unextendibility to quantum measurements and define k-extendible measurements for every integer $k \geq 2$. Our definition provides a hierarchy of semidefinite constraints that specify a set of measurements containing every measurement that can be realized by local operations and one-way classical communication. Furthermore, the set of k-extendible measurements converges to the set of measurements that can be realized by local operations and one-way classical communication as $k \to \infty$.

7. I. George, C. Hirche, **T. Nuradha**, M. M. Wilde. "Quantum Doeblin Coef-

ficients: Interpretations and Applications". arXiv:2503.22823. [50]

The Doeblin coefficient of a classical channel provides an efficiently computable upper bound on the total-variation contraction coefficient of the channel. Here, we investigate quantum Doeblin coefficients by defining various new quantum Doeblin coefficients and analysing their properties. We also develop various interpretations of two of these quantum Doeblin coefficients, including representations as minimal singlet fractions, exclusion values, among others. We also show how our findings provide insights on limitations on quantum learning algorithms that use parameterized quantum circuits (noise-induced barren plateaus), on error mitigation protocols, and on the sample complexity of noisy quantum hypothesis testing.

8. H. Westerheim, J. Chen, Z. Holmes, I. Luo, **T. Nuradha**, D. Patel, S. Rethinasamy, K. Wang, and M. M. Wilde, "Dual-VQE: A quantum algorithm to lower bound the ground-state energy", 2023, arXiv:2312.03083. [136]

   The variational quantum eigensolver (VQE) is a hybrid quantum-classical variational algorithm that produces an upper-bound estimate of the ground-state energy of a Hamiltonian. In this work, we propose a dual variational quantum eigensolver (dual-VQE) that produces a lower-bound estimate of the ground-state energy. As such, VQE and dual-VQE can serve as quality checks on their solutions; in the ideal case, the VQE upper bound and the dual-VQE lower bound form an interval containing the true optimal value of the ground-state energy.

9. J. Chen, H. Westerheim, Z. Holmes, I. Luo, **T. Nuradha**, D. Patel, S. Rethinasamy, K. Wang, and M. M. Wilde, "QSlack: A slack-variable approach for variational quantum semi-definite programming", 2023,

arXiv:2312.03830. [20]

Solving optimization problems is a key task for which quantum computers could possibly provide a speedup over the best-known classical algorithms. We propose the QSlack method for estimating the optimal values of semi-definite programs (SDP). QSlack works by introducing slack variables to transform inequality constraints to equality constraints, transforming a constrained optimization to an unconstrained one via the penalty method, and replacing the optimizations over all possible non-negative variables by optimizations over parameterized quantum states.

1.3 Organization

This thesis is organized as follows. First, in Chapter 2, we establish the notations that we use and the background about privacy frameworks. Next, in Chapter 3, we propose a flexible privacy framework for quantum systems while analyzing desired properties, mechanisms, and providing operational interpretations to quantum information-theoretic quantities. Then, in Chapter 4, we define measured hockey-stick divergences and show how they find use in the analysis of flexible quantum privacy frameworks. Furthermore, we study how the contraction of quantum divergences behaves with the privacy constraints imposed by quantum local differential privacy in Chapter 5. Then, in Chapter 6, we utilize those tools developed on contraction of divergences to study the impact of privacy in statistical learning tasks, including hypothesis testing of quantum states and learning expectation of observables with access to only privatized quantum states. Finally, in Chapter 7, we summarize the main contributions of this thesis and state some potential future directions in the scope of the study of this thesis.

CHAPTER 2

PRELIMINARIES AND BACKGROUND

In this chapter, we present the notation used throughout the thesis while providing background on concepts related to existing privacy frameworks for classical and quantum systems.

2.1 Notations and Quantum Divergences

2.1.1 Notation

Sets are denoted by calligraphic letters, e.g., $\mathcal{X}$. For $k, n \in \mathbb{N}$, we use $\mathcal{X}^{n\times k}$ to denote the database space of $n \times k$ matrices; columns correspond to different attributes, while rows correspond to different individuals. The (i, j)th entry of $x \in \mathcal{X}^{n\times k}$ is denoted as $x(i, j)$. The ith row and jth column of x are denoted by $x(i, \cdot)$ and $x(\cdot, j)$, respectively. We denote by $(\Omega, \mathcal{F}, \mathbb{P})$ the underlying probability space on which all random variables (RVs) are defined, with $\mathbb{E}$ designating expectation. RVs are denoted by uppercase letters, e.g., X, with P_X representing the corresponding probability law. For $X \sim P_X$, we interchangeably use $\mathrm{supp}(X)$ and $\mathrm{supp}(P_X)$ for the support. The joint law of (X, Y) is denoted by P_{XY}, while $P_{Y|X}$ designates the (regular) conditional probability of Y given X. Conventions for $n \times k$-dimensional random variables are the same as for deterministic elements. The space of all Borel probability measures on $\mathcal{S} \subseteq \mathbb{R}^d$ is denoted by $\mathcal{P}(\mathcal{S})$. The Kullback–Leibler (KL) divergence between $P, Q \in \mathcal{P}(\mathcal{X})$ with $P \ll Q$ (absolutely continuous) is given by $D(P\|Q) := \mathbb{E}_P\left[\ln\left(\frac{dP}{dQ}\right)\right]$, where $\frac{dP}{dQ}$ is the Radon–Nikodym derivative of P with respect to Q. For $(X, Y) \sim P_{XY}$, the mutual information be-

tween X and Y is denoted by $I(X;Y) \coloneqq D(P_{XY}||P_X \otimes P_Y)$.

We now review basic concepts from quantum information theory and refer to [75,137] for more details. A (classical or quantum) system R is identified with a finite-dimensional Hilbert space $\mathcal{H}_R$. We denote the set of linear operators acting on $\mathcal{H}_R$ by $\mathcal{L}(\mathcal{H}_R)$. The support of a linear operator $X \in \mathcal{L}(\mathcal{H}_R)$ is defined to be the orthogonal complement of its kernel, and it is denoted by $\text{supp}(X)$. Let $\text{T}(C)$ denote the transpose of C. The partial transpose of $C \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ on the subsystem A is represented as $\text{T}_A(C)$. Let $\text{Tr}[C]$ denote the trace of C, and let $\text{Tr}_A[C]$ denote the partial trace of C over the subsystem A. The trace norm of a matrix B is defined as $\|B\|_1 \coloneqq \text{Tr}\left[\sqrt{B^\dagger B}\right]$. For operators A and B, the notation $A \geq B$ indicates that $A - B$ is a positive semi-definite (PSD) operator, while $A > B$ indicates that $A - B$ is a positive definite operator.

A quantum state $\rho_R \in \mathcal{L}(\mathcal{H}_R)$ on R is a PSD, unit-trace operator acting on $\mathcal{H}_R$. We denote the set of all density operators in $\mathcal{L}(\mathcal{H}_R)$ by $\mathcal{D}(\mathcal{H}_R)$. A state ρ_R of rank one is called pure, and we may choose a normalized vector $|\psi\rangle \in \mathcal{H}_R$ satisfying $\rho_R = |\psi\rangle\langle\psi|$ in this case. Otherwise, ρ_R is called a mixed state. By the spectral decomposition theorem, every mixed state can be written as a convex combination of pure, orthogonal states. A quantum channel $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ is a linear, completely positive and trace-preserving (CPTP) map from $\mathcal{L}(\mathcal{H}_A)$ to $\mathcal{L}(\mathcal{H}_B)$. We denote the adjoint of $\mathcal{N}$ by $\mathcal{N}^\dagger$. A measurement of a quantum system R is described by a positive operator-valued measure (POVM) $\{M_y\}_{y \in \mathcal{Y}}$, which is defined to be a collection of PSD operators satisfying $\sum_{y \in \mathcal{Y}} M_y = I_{\mathcal{H}_R}$, where $\mathcal{Y}$ is a finite alphabet. The Born rule dictates that, after applying the above POVM to $\rho \in \mathcal{D}(\mathcal{H}_R)$, the probability of observing the outcome y is given by $\text{Tr}\left[M_y \rho\right]$.

2.1.2 Quantum Divergences

We define several quantum divergences that will be used throughout this thesis.

Definition 1 (Generalized Quantum Divergence)**.** *We call a distinguishability measure $\boldsymbol{D}(\cdot\|\cdot)$ a generalized divergence [119] if it satisfies the data-processing inequality; i.e., for every channel $\mathcal{N}$, state ρ, and PSD operator σ,*

$$\boldsymbol{D}(\rho\|\sigma) \geq \boldsymbol{D}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))\,. \tag{2.1}$$

Definition 2 (Glossary of Divergences and Information Measures)**.** *We define several quantum divergences as follows.*

1. *The normalized trace distance between the states ρ and σ is defined as*

$$T(\rho,\sigma) \coloneqq \frac{1}{2}\left\|\rho-\sigma\right\|_1. \tag{2.2}$$

   It generalizes the total-variation distance between two probability distributions.

2. *For $\gamma \geq 0$, the quantum hockey-stick divergence is defined as [99, 119]*

$$E_\gamma(\rho\|\sigma) \coloneqq \mathrm{Tr}[(\rho-\gamma\sigma)_+] - (1-\gamma)_+, \tag{2.3}$$

   where $(A)_+ \coloneqq \sum_{i:a_i\geq 0} a_i|i\rangle\langle i|$ for a Hermitian operator $A = \sum_i a_i|i\rangle\langle i|$, and $(x)_+ \coloneqq \max\{0,x\}$ for real number x. For $\gamma = 1$, observe that $E_1(\rho\|\sigma) = T(\rho,\sigma)$.

3. *Uhlamnn Fidelity between two states ρ and σ is defined as [127]*

$$F(\rho,\sigma) \coloneqq \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2 \tag{2.4}$$

   and the square of the Bures distance $d_B(\rho,\sigma)$ as

$$[d_B(\rho,\sigma)]^2 \coloneqq 2\left(1-\sqrt{F}(\rho,\sigma)\right). \tag{2.5}$$

4. *The diamond distance between the two channels* $\mathcal{N}, \mathcal{M} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ *is defined as [79]*

$$\|\mathcal{N} - \mathcal{M}\|_\diamond := \sup_{\rho_{RA}} \|\mathcal{N}_{A\to B}(\rho_{RA}) - \mathcal{M}_{A\to B}(\rho_{RA})\|_1 , \tag{2.6}$$

where the optimization in the definition is over every reference system R *and bipartite density operator* ρ_{RA} *(with* R *allowed to be arbitrarily large). It is well known, however, that it suffices to perform the optimization over pure bipartite states such that the dimension of the reference system* R *is equal to the dimension of the channel input system* A*.*

5. *The Petz–Rényi quantum relative entropy of order* $\alpha \in (0,1) \cup (1,\infty)$ *of a state* ρ *with respect to a PSD operator* σ *is given by [108, 109]: if* $\alpha \in (0,1) \vee (\alpha > 1 \wedge \operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma))$ *then*

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \ln \operatorname{Tr}[\rho^\alpha \sigma^{1-\alpha}] \tag{2.7}$$

and ∞ *otherwise. It is a generalized divergence for* $\alpha \in [0,1) \cup (1,2]$ *[109]. The special case of* $\alpha \to 1$ *is called the quantum relative entropy and amounts to*

$$D(\rho\|\sigma) \equiv D_1(\rho\|\sigma) := \lim_{\alpha\to 1} D_\alpha(\rho\|\sigma) = \operatorname{Tr}[\rho(\ln\rho - \ln\sigma)] \tag{2.8}$$

when $\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma)$ *and it is equal to* $+\infty$ *otherwise. The quantum entropy of a state* ρ *is defined as*

$$S(\rho) := -\operatorname{Tr}[\rho \ln \rho] . \tag{2.9}$$

Equivalently, $S(\rho) = -D_1(\rho\|I)$*, where* I *is the identity operator.*

6. *Fix* $\alpha \in (0,1) \cup (1,\infty)$*. The sandwiched Rényi relative entropy of a state* ρ *and a PSD operator* σ *is defined as [89, 139]: if* $\alpha \in (0,1) \vee (\alpha \in (1,\infty) \wedge \operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma))$ *then*

$$\widetilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \ln \operatorname{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right] \tag{2.10}$$

and ∞ otherwise. It is a generalized divergence for $\alpha \in [1/2, 1) \cup (1, \infty)$ [46] (see also [138]).

7. *Fix $\delta \in [0, 1]$, a state ρ, and a PSD operator σ. The Datta–Leditzky information spectrum divergences are defined as follows [29]:*

$$\underline{D}^{\delta}(\rho\|\sigma) := \sup\{\gamma \in \mathbb{R} : \mathrm{Tr}[(\rho - e^{\gamma}\sigma)_+] \geq 1 - \delta\}, \tag{2.11a}$$

$$\overline{D}^{\delta}(\rho\|\sigma) := \inf\{\gamma \in \mathbb{R} : \mathrm{Tr}[(\rho - e^{\gamma}\sigma)_+] \leq \delta\}, \tag{2.11b}$$

where $(A)_+ := \sum_{i:a_i \geq 0} a_i |i\rangle\langle i|$ for a Hermitian operator $A = \sum_i a_i |i\rangle\langle i|$. Hereafter, we abbreviate these divergences as DL divergences. Proposition 4.3 of [29] shows that

$$\underline{D}^{\delta}(\rho\|\sigma) = \overline{D}^{1-\delta}(\rho\|\sigma), \tag{2.12}$$

and so we can speak of a single DL divergence, which we set hereafter to be $\overline{D}^{\delta}$ from (2.11b). Slightly rewriting (2.11), we have the equivalent representations:

$$\underline{D}^{\delta}(\rho\|\sigma) = \ln \sup\{\lambda \geq 0 : \mathrm{Tr}[(\rho - \lambda\sigma)_+] \geq 1 - \delta\} \tag{2.13a}$$

$$\overline{D}^{\delta}(\rho\|\sigma) = \ln \inf\{\lambda \geq 0 : \mathrm{Tr}[(\rho - \lambda\sigma)_+] \leq \delta\}. \tag{2.13b}$$

8. *The max-relative entropy of a state ρ and a PSD operator σ is defined as [28]*

$$D_{\max}(\rho\|\sigma) := \ln \inf\{\lambda : \rho \leq \lambda\sigma\} \tag{2.14}$$

$$= \ln \sup_{0 \leq M \leq I} \frac{\mathrm{Tr}[M\rho]}{\mathrm{Tr}[M\sigma]}, \tag{2.15}$$

and the smooth max-relative entropy is defined for $\delta \in [0, 1]$ as

$$D^{\delta}_{\max}(\rho\|\sigma) := \inf_{\widetilde{\rho} : \frac{1}{2}\|\widetilde{\rho} - \rho\|_1 \leq \delta} D_{\max}(\widetilde{\rho}\|\sigma), \tag{2.16}$$

with the optimization taken over every state $\widetilde{\rho}$. These quantities have been given an operational meaning in [130]. The Thompson metric [125] is defined in terms of the max-relative entropy as

$$D_T(\rho\|\sigma) := \max\{D_{\max}(\rho\|\sigma), D_{\max}(\sigma\|\rho)\}, \tag{2.17}$$

and it has been given an operational meaning in [114, 117].

2.2 Classical and Quantum Privacy Frameworks

In this section, we provide background on the existing definitions of privacy for both classical and quantum systems, starting from classical DP and proceeding to quantum DP.

2.2.1 Classical Differential and Pufferfish Privacy

DP allows for answering queries about aggregate quantities while protecting the individual entries in a database [40]. To this end, the output of a differential privacy mechanism should be indistinguishable for neighboring databases, defined as those that differ only in a single record (row). Formally, we say that $x, x' \in \mathcal{X}^{n\times k}$ are neighbors, denoted $x \sim x'$, if $x(i,\cdot) \neq x'(i,\cdot)$ for some $i \in \{1,\dots,n\}$, and they agree on all other rows. We also note that a randomized privacy mechanism A, as mentioned below, is described by a conditional probability distribution $P_{A|X}$ for its output given the data.

Definition 3 (Classical Differential Privacy)**.** *Fix $\varepsilon \geq 0$ and $\delta \in [0,1]$. A randomized mechanism $A : \mathcal{X}^{n\times k} \to \mathcal{Y}$ is (ε,δ)-differentially private if*

$$\mathbb{P}(A(x) \in \mathcal{B}) \leq e^{\varepsilon}\, \mathbb{P}(A(x') \in \mathcal{B}) + \delta, \tag{2.18}$$

for all $x \sim x'$ with $x, x' \in \mathcal{X}^{n\times k}$ and $\mathcal{B} \subseteq \mathcal{Y}$ measurable.

As is evident from the above definition, DP aims to conceal whether any particular individual (row) is in fact part of the database or not. While being a

powerful and widely applicable privacy framework, it is often appropriate to consider even broader frameworks. Pufferfish privacy [77] is a versatile generalization of DP that not only allows flexibility in the definition of secrets but also enables the integration of domain knowledge of the database space $\mathcal{X}^{n\times k}$. The PP framework consists of three components:

1. A set of secrets $\mathcal{S} \subseteq \mathcal{X}^{n\times k}$ of measurable subsets;
2. A set of secret pairs $\mathcal{Q} \subseteq \mathcal{S}\times\mathcal{S}$ that need to be indistinguishable in the (ε, δ) sense (c.f., (2.19) below),
3. A class of data distributions $\Theta \subseteq \mathcal{P}(\mathcal{X}^{n\times k})$ that captures prior beliefs or domain knowledge.

As formulated next, PP aims to guarantee that all secret pairs in $\mathcal{Q}$ are indistinguishable with respect to the prior beliefs $P_X \in \Theta$.

Definition 4 (Classical Pufferfish Privacy)**.** *Fix $\varepsilon \geq 0$ and $\delta \in [0, 1]$. A randomized mechanism $A : \mathcal{X}^{n\times k} \to \mathcal{Y}$ is (ε, δ)-private in the pufferfish framework $(\mathcal{S}, \mathcal{Q}, \Theta)$ if*

$$\mathbb{P}(A(X) \in \mathcal{B} \big| \mathcal{R}) \leq e^{\varepsilon}\, \mathbb{P}(A(X) \in \mathcal{B} \big| \mathcal{T}) + \delta, \tag{2.19}$$

for all $P_X \in \Theta$, $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$ with $P_X(\mathcal{R}), P_X(\mathcal{T}) > 0$, and $\mathcal{B} \subseteq \mathcal{Y}$ measurable.

DP from Definition 3 is a special case of PP when $\mathcal{S} = \mathcal{X}^{n\times k}$, the set $\mathcal{Q}$ contains all neighboring pairs of databases, and $\Theta = \mathcal{P}(\mathcal{X}^{n\times k})$ (i.e., there are no distributional assumptions, and privacy is guaranteed in the worst case). Other important examples that are subsumed by PP include (i) generic DP [104], which allows for arbitrary neighboring relationships, and (ii) attribute privacy [143], which privatizes global properties of a database (e.g., a column that corresponds to some sensitive information, such as salary).

2.2.2 Quantum Differential Privacy

Quantum differential privacy (QDP) lifts the notion of DP to the space of quantum states, with the neighboring relation typically defined either in terms of closeness in trace distance [144], reachability by a single local operation [1],[1] or by quantum Wasserstein distance of order 1 [32]. We denote two states being neighbors by $\rho \sim \sigma$.

Definition 5 (Quantum Differential Privacy [62, 144])**.** *Fix $\varepsilon \geq 0$ and $\delta \in [0, 1]$. Let $\mathcal{D}$ be a set of quantum states, and let $\mathcal{A}$ be a quantum algorithm (viz., a quantum channel). The algorithm $\mathcal{A}$ is (ε, δ)-differentially private if*

$$\mathrm{Tr}[M\mathcal{A}(\rho)] \leq e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)] + \delta. \tag{2.20}$$

for every measurement operator M (i.e., satisfying $0 \leq M \leq I$) and all $\rho, \sigma \in \mathcal{D}$ such that $\rho \sim \sigma$.

This definition reduces to classical DP for discrete-output mechanisms with an appropriate choice of the measurement set. See Section 3.2.3 below for further details.

[1]Given two quantum states ρ and σ of n registers each, call them neighbors if it is possible to reach either σ from ρ or ρ from σ by performing a general quantum channel on a single register only.

CHAPTER 3

QUANTUM PUFFERFISH PRIVACY: A FLEXIBLE PRIVACY FRAMEWORK FOR QUANTUM SYSTEMS

3.1 Introduction

With a surging interest in quantum and hybrid classical–quantum systems, ensuring privacy of both classical and quantum data has become pivotal. Privacy-preserving data analysis has been widely studied for classical systems by means of statistical privacy frameworks. Differential privacy (DP) is an important statistical privacy framework that enables answering aggregate queries about a database while keeping individual records private [40, 41].

Quantum DP (QDP) is a generalization of the classical DP notion and has been proposed in [144]. See also [1] for DP of quantum measurements and [62] for an information-theoretic interpretation of QDP. Connections to quantum stability through private learning have been studied in [112]. Moreover, [37] has explored how quantum classifiers can be made private by using the intrinsic noise of existing quantum systems. See also [6, 69, 118, 131] for applications of DP in quantum machine learning. Additionally, privacy amplification of quantum and quantum-inspired algorithms has been analyzed using QDP and classical DP notion in [4]. However, similar to the classical case, the versatility of QDP is limited.

In this chapter, we propose a flexible privacy framework for quantum systems, termed quantum PP (QPP), that addresses these limitations. We provide a comprehensive study of QPP, encompassing properties, mechanisms, privacy-

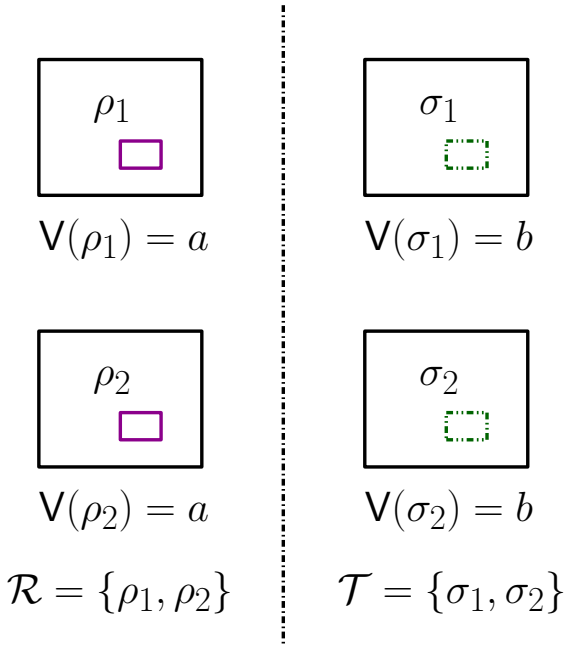


Figure 3.1: Depiction of a setup where the goal is to hide whether the amount of entanglement V present in the bipartite states ρ_1, ρ_2, σ_1, and σ_2 equals a or b. In this diagram, large squares represent the entire quantum state, while small rectangles correspond to a specific attribute of that state (i.e., the amount of entanglement as quantified by the function V). The specific attribute can take on one of two values, a or b, represented by solid or dotted lines, respectively. As the goal is to conceal only the entanglement level, and not necessarily the specific quantum state, we want the sets $\mathcal{R} = \{\rho_1, \rho_2\}$ and $\mathcal{T} = \{\sigma_1, \sigma_2\}$ to be indistinguishable.

utility tradeoffs, as well as the first operational meaning of the Datta–Leditzky information spectrum divergence [29] (hereafter abbreviated as the DL divergence), which arises from our framework.

3.1.1 Motivation

We seek to address key limitations of QDP by exploring more *flexible privacy frameworks* for quantum information processing. As delineated next, flexible secrets, embedding domain knowledge, and relaxing the need for worst-case

measurements are considerations central to our approach.

Flexible secrets. QDP guarantees that any pair of states that are classified as neighbors are approximately indistinguishable, i.e., cannot be identified under any possible measurement. However, scenarios may arise in which one wants to hide specific properties of the states, as opposed to the state itself (e.g., whether the states possess a certain symmetry or property, have any entanglement with a special subsystem, or have secret correlations with other systems). In such situations, QDP may be an overly pessimistic notion of privacy, which, in turn, hinders utility. As an example, consider hiding the amount of entanglement V present in the bipartite states in the set $\{\rho_1, \rho_2, \sigma_1, \sigma_2\}$, for which $\mathsf{V}(\rho_1) = \mathsf{V}(\rho_2) = a$ and $\mathsf{V}(\sigma_1) = \mathsf{V}(\sigma_2) = b$. As illustrated in Fig. 3.1, hiding whether V equals a or b amounts to making the classes $\{\rho_1, \rho_2\}$ and $\{\sigma_1, \sigma_2\}$ indistinguishable. This can be achieved by applying a QDP mechanism to the state space $\{\rho_1, \rho_2, \sigma_1, \sigma_2\}$ by choosing (ρ_i, σ_j) for all $i, j \in \{1, 2\}$ as neighbors, with the criterion that ρ and σ are neighbors if and only if $|\mathsf{V}(\rho) - \mathsf{V}(\sigma)| = |a - b|$. However, doing so provides a stricter guarantee than required. The source of the issue is the inability of QDP to account for secrets concerning collections of states (as opposed to singletons), which is the first issue we aim to address.

Domain knowledge. In QDP, a worst-case privacy guarantee is provided for all neighboring states. However, one may possess knowledge about the likelihood of observing different states, e.g., via expert feedback. Referring back to the setting from Fig. 3.1, if we have domain knowledge such that observing the states $\rho_1, \rho_2, \sigma_1, \sigma_2$ is prescribed by the probability vector $(p/2, (1-p)/2, 1/2, 0)$, for $p \in (0, 1)$, then the requirement simplifies to the indistinguishability of $\{\rho_1, \rho_2\}$ versus $\{\sigma_1\}$. Classically, it has been demonstrated that domain knowledge can

be leveraged to design privacy mechanisms with increased accuracy and utility [12, 77]. This calls for a quantum privacy framework that can also encode domain knowledge.

Relaxing worst-case measurements. Another worst-case aspect of QDP is its account of all possible measurements. However, such a requirement might be too stringent in practice, especially in quantum systems. As an example, while a joint measurement can accurately distinguish between entangled but physically separated states, oftentimes only local operations and classical communications (LOCC) are available (e.g., as considered in quantum data-hiding protocols [34, 42, 57, 58, 80, 85, 124]). In such cases, one may achieve improved accuracy and utility by relaxing the privacy requirement to account for LOCC measurements only.

In sum, the rapid advancements in quantum technologies require designing flexible privacy frameworks that can be adjusted to timely needs. Furnishing such a framework is the main objective of this chapter.

3.1.2 Contributions

This chapter proposes a quantum analog of the PP framework that accounts for the three aforementioned aspects. Our formalism enables reasoning about the privacy of quantum systems using information-theoretic tools. We provide a comprehensive study of QPP, encompassing properties, mechanisms, and privacy-utility tradeoffs. Our paradigm also gives rise to the first operational interpretation of the DL divergence [29]. The proposed QPP framework comprises four key ingredients:

1. the set of potential secrets,
2. the set of discriminative pairs that are required to be indistinguishable at the output of the mechanism,
3. the set of data distributions, which encodes domain knowledge on the occurrence of quantum or classical data, and
4. the set of measurements to be accounted for, which is specified based on physical, ethical, or any other constraints.

See Definition 6 for a formal definition. QPP guarantees the indistinguishability, under any allowable measurement, of sets of states formed based on the above ingredients.

After defining the operational privacy framework, we observe that when the measurement class contains all possible measurements, QPP can be equivalently posed as a DL divergence constraint. To the best of our knowledge, this provides the first operational interpretation of the DL divergence. We then derive an efficiently computable formulation of the DL divergence as a semi-definite program (SDP), which may be of independent interest. This SDP is utilized to prove properties of the DL divergence, which are then used in the analysis of QPP mechanisms. These properties include joint quasi-convexity and the data-processing inequality under positive and trace non-increasing maps (see Section 3.3). Our results also generalize the connection between the hockey-stick divergence and QDP, originally established in [62]. Moreover, we show that existing privacy frameworks such as classical DP [40, 41], classical PP [77], utility-optimized local DP (not subsumed by classical PP) [90], and QDP [62, 144] are special cases of our QPP framework.

We then move on to derive properties of QPP mechanisms, encompassing convexity, post-processing, and composability (both parallel and adaptive). As a specific example, we characterize the flip parameter that guarantees QPP of the depolarization mechanism. We also describe how QPP mechanisms implementable on quantum devices can be instantiated to achieve classical PP. We consider the associated privacy-utility tradeoff for QPP mechanisms. Our utility metric captures how invertible the privacy mechanism is, which is formulated as the infimized diamond distance between a post-processing of the mechanism's output and the identity channel. We show that this utility metric can be computed as an SDP, and we analyze the privacy-utility tradeoff of the depolarization mechanism. Lastly, we study optimal privacy-utility tradeoffs of QPP mechanisms and characterize the achievable region in several settings.

Another application we consider is privacy auditing, which refers to certifying whether a black-box mechanism satisfies a target privacy guarantee. While several auditing methods are available for classical frameworks, there is currently no approach that can handle quantum data. We fill this gap by proposing the first auditing pipeline for quantum privacy mechanisms. In contrast to existing approaches for classical DP and PP, which require first relaxing the privacy notion and only then auditing, our approach audits for QDP directly.

Finally, we explore connections between QPP, existing quantum privacy frameworks, and figures of merit. First, we provide bounds on quantum Rényi divergences and the trace distance, which stem from QPP. This inspires relaxations of QPP that are defined via these divergences. Lastly, we present a variant of QPP that can incorporate entanglement into the framework with the use of reference systems.

3.2 Quantum Pufferfish Privacy (QPP)

Inspired by the versatility of the classical PP framework, we propose a quantum variant thereof. Termed QPP, our framework allows for customizing the notion of private states, tailoring the feasible set of measurements to the application of interest, and incorporating domain knowledge of the state distribution into the model. As such, the QPP framework can generate a rich class of privacy definitions for both classical and quantum systems, and for hybrid classical–quantum systems as well.

3.2.1 Framework

The QPP framework requires a domain expert to specify four components: a set $\mathcal{S}$ of potential secrets, a set $\mathcal{Q} \subseteq \mathcal{S} \times \mathcal{S}$ of discriminative pairs, a set Θ of data distributions, and a set $\mathcal{M}$ of measurements. We expand on and explicitly define each component next.

Set $\mathcal{S}$ of potential secrets: Secrets are modeled as subsets of density operators that share a certain property (these subsets are merely singletons in the QDP case). The set $\mathcal{S}$ is a collection of such secret subsets. For example, if one aims to privatize the resource value V of a state, then the corresponding set of secrets is $\mathcal{S} = \bigcup_{i=1}^{n} \mathcal{T}_i$, where

$$\mathcal{T}_i = \{\rho \in \mathcal{D}(\mathcal{H}) : \mathsf{V}(\rho) = a_i\} \tag{3.1}$$

and $\{a_i\}_{i=1}^{n}$ are the possible values that V can take (recall that, in Fig. 3.1, we considered a setup relevant to hiding the resource value V being a or b).

Set $\mathcal{Q}$ of discriminative pairs: This is a subset of $\mathcal{S} \times \mathcal{S}$ that specifies which pairs of elements from $\mathcal{S}$ should be indistinguishable. Namely, if $(\mathcal{T}_1, \mathcal{T}_2) \in \mathcal{Q}$, then the goal of the privacy mechanism is to conceal whether the input belongs to $\mathcal{T}_1$ or $\mathcal{T}_2$. Note that $\rho \in \mathcal{T}_1 \Rightarrow \rho \notin \mathcal{T}_2$. We require that $\mathcal{Q}$ is symmetric, i.e., that $(\mathcal{T}_i, \mathcal{T}_j) \in \mathcal{Q}$ if and only if $(\mathcal{T}_j, \mathcal{T}_i) \in \mathcal{Q}$. Proceeding with the same example, we can set

$$\mathcal{Q} = \bigcup_{i \neq j} \{(\mathcal{T}_i, \mathcal{T}_j)\}. \tag{3.2}$$

Set Θ of data distributions: A collection of probability distributions $P_X \in \mathcal{P}(\mathcal{X})$ over a finite space $\mathcal{X}$ that indexes an ensemble of density operators $\{\rho^x\}_{x \in \mathcal{X}}$. Taking $X \sim P_X \in \Theta$, the matrix-valued random variable ρ^X models a density operator that is randomly chosen according to P_X. Proceeding with the same example, $\{\rho^x\}_{x \in \mathcal{X}} = \{\sigma : \sigma \in \mathcal{T}_i,\ \mathcal{T}_i \in \mathcal{S}\} \subset \mathcal{D}(\mathcal{H})$. The set Θ can be understood as capturing beliefs that the adversary has regarding the state of the system.

In the above example, we have considered a subset of density operators (i.e., $\{\rho^x\}_{x \in \mathcal{X}} \subset \mathcal{D}(\mathcal{H})$). There could be applications where we have to consider all density operators. To incorporate this, we choose the following: Fix $k \in \mathbb{N}$ and let $\mathfrak{F}_k \subset 2^{\mathcal{D}(\mathcal{H})}$ be the collection of all finite subsets of $\mathcal{D}(\mathcal{H})$ with k elements. For each $\mathcal{F} \in \mathfrak{F}_k$, we write $\mathcal{P}(\mathcal{F})$ for the class of all distributions supported on $\mathcal{F}$, and define

$$\mathcal{P}_k(\mathcal{D}(\mathcal{H})) \coloneqq \bigcup_{\mathcal{F} \in \mathfrak{F}_k} \mathcal{P}(\mathcal{F}). \tag{3.3}$$

Every distribution $P \in \mathcal{P}_k(\mathcal{D}(\mathcal{H}))$ is supported on exactly k density operators. Note that all density operators outside of the underlying finite set comprise of the null set. We associate a random variable $X \sim P = P_X$ with each such distribution and write $\mathcal{X} = \mathrm{supp}(P_X)$ for its support. Note the slight abuse in notation, as the support of P_X changes with the distribution, which is not reflected in the

generic indexing set $\mathcal{X}$. The set of data distributions in the QPP framework is now taken as $\Theta \subseteq \mathcal{P}_k(\mathcal{D}(\mathcal{H}))$ for some $k \in \mathbb{N}$.

Set $\mathcal{M}$ of measurements: This set is a subset of all possible measurements, i.e., $\mathcal{M} \subseteq \{M : 0 \leq M \leq I\}$. The choice of $\mathcal{M}$ gives the flexibility to consider only measurements that are possible under physical, legal, or ethical constraints.

Now, we are ready to present a formal definition of the quantum analog of PP, which we call QPP.

Definition 6 (Quantum Pufferfish Privacy)**.** *Fix $\varepsilon \geq 0$ and $\delta \in [0, 1]$. A quantum algorithm $\mathcal{A}$ is (ε, δ)-private in the quantum pufferfish framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ if for all $P_X \in \Theta$, $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$ with $P_X(\mathcal{R}), P_X(\mathcal{T}) > 0$, and all $M \in \mathcal{M}$, the following inequality holds:*

$$\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{R}})\Big] \leq e^{\varepsilon}\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big] + \delta, \tag{3.4}$$

where

$$\rho^{\mathcal{R}} \coloneqq \sum_{\{x:\rho^x \in \mathcal{R}\}} q_{\mathcal{R}}(x)\rho^x, \tag{3.5}$$

$$q_{\mathcal{R}}(x) \coloneqq \frac{P_X(x)}{P_X(\mathcal{R})}, \tag{3.6}$$

$$P_X(\mathcal{R}) \coloneqq \sum_{\{x:\rho^x \in \mathcal{R}\}} P_X(x), \tag{3.7}$$

and $\rho^{\mathcal{T}}$ is defined similarly but with $\mathcal{T}$ instead of $\mathcal{R}$. We say that an algorithm $\mathcal{A}$ satisfies ε-QPP if it satisfies $(\varepsilon, 0)$-QPP.

Evidently, discriminative secret pairs in $\mathcal{Q}$ are indistinguishable at the output of a QPP mechanism $\mathcal{A}$ in the (ε, δ)-sense, under every measurement from the class $\mathcal{M}$.

Remark 1 (Semantics of the QPP Framework)**.** *The QPP framework provides the following privacy guarantee for fixed $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$ and $P_X \in \Theta$: For a state ρ^X chosen*

according to $X \sim P_X$ and input to the quantum channel $\mathcal{A}$, an adversary applying a measurement $M \in \mathcal{M}$ on the channel output $\mathcal{A}(\rho^X)$ draws the same conclusions regardless of whether ρ^X belongs to $\mathcal{R}$ or $\mathcal{T}$.

Remark 2 (Incorporating Entanglement)**.** *We can incorporate entanglement in the QPP framework by introducing a reference system. Specifically, we can modify the QPP framework from $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ to $(\mathcal{S}, \mathcal{G}, \Theta, \mathcal{M}')$, where*

$$\mathcal{G} := \left\{ \begin{array}{c} \omega_{RA}^{\mathcal{R}}, \omega_{RA}^{\mathcal{T}} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_A), \\ (\omega_{RA}^{\mathcal{R}}, \omega_{RA}^{\mathcal{T}}) : \mathrm{Tr}_R\left[\omega_{RA}^{\mathcal{R}}\right] = \rho^{\mathcal{R}}, \mathrm{Tr}_R\left[\omega_{RA}^{\mathcal{T}}\right] = \rho^{\mathcal{T}}, \\ (\mathcal{R}, \mathcal{T}) \in \mathcal{Q} \end{array} \right\} \tag{3.8}$$

is a set of pairs of bipartite states with $\rho^{\mathcal{R}}$ and $\rho^{\mathcal{T}}$ defined similar to Definition 6. We then say that $\mathcal{A}$ is (ε, δ)-QPP in that framework if for all $P_X \in \Theta$, $M' \in \mathcal{M}'$, and $(\omega_{RA}^{\mathcal{R}}, \omega_{RA}^{\mathcal{T}}) \in \mathcal{G}$, we have

$$\mathrm{Tr}\left[M'(\mathcal{I} \otimes \mathcal{A})(\omega_{RA}^{\mathcal{R}})\right] \leq e^{\varepsilon}\mathrm{Tr}\left[M'(\mathcal{I} \otimes \mathcal{A})(\omega_{RA}^{\mathcal{T}}))\right] + \delta. \tag{3.9}$$

However, it is unclear whether such a stronger privacy notion would be useful in practical applications. For example, consider $\sigma_1 := |0\rangle\langle 0| \otimes \rho^{\mathcal{R}}$ and $\sigma_2 := |1\rangle\langle 1| \otimes \rho^{\mathcal{T}}$ with $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$. If a measurement on the reference system can be applied, then a computational-basis measurement distinguishes σ_1 and σ_2 perfectly. Thus, it is important to choose $\mathcal{A}$ appropriately with a practically applicable $\mathcal{M}'$, such that the required indistinguishability is achieved.

We shall revisit a variant of this framework with quantum divergences in Section 3.8.2. The strength of the privacy framework is determined by the underlying quantum divergence. However, note that the problems discussed previously are not completely solved by the variant proposed therein.

3.2.2 Equivalent Formulation of QPP with DL Divergence

We present an equivalent formulation for (ε, δ)-QPP by means of the DL divergence from (2.11b). To the best of our knowledge, this provides the first operational interpretation of the DL divergence.

Proposition 1 (Equivalent Formulation of (ε, δ)-QPP)**.** *Fix the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$, with $\bar{\mathcal{M}}$ corresponding to the set of all possible measurements. Then algorithm $\mathcal{A}$ satisfies (ε, δ)-QPP with respect to the framework $(\mathcal{S}, \mathcal{Q}, \mathcal{M}, \Theta)$ if and only if for all $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$, we have*

$$\overline{D}^{\delta}\left(\mathcal{A}(\rho^{\mathcal{R}}) \middle\| \mathcal{A}(\rho^{\mathcal{T}})\right) \leq \varepsilon. \tag{3.10}$$

Proof. We first show that (ε, δ)-QPP implies (3.10). For fixed $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$, observe that (ε, δ)-QPP corresponds to

$$\sup_{M \in \bar{\mathcal{M}}} \operatorname{Tr}\left[M\left(\mathcal{A}(\rho^{\mathcal{R}}) - e^{\varepsilon}\mathcal{A}(\rho^{\mathcal{T}})\right)\right] \leq \delta. \tag{3.11}$$

Since

$$\sup_{M \in \bar{\mathcal{M}}} \operatorname{Tr}\left[M\left(\mathcal{A}(\rho^{\mathcal{R}}) - e^{\varepsilon}\mathcal{A}(\rho^{\mathcal{T}})\right)\right] = \operatorname{Tr}\left[\left(\mathcal{A}(\rho^{\mathcal{R}}) - e^{\varepsilon}\mathcal{A}(\rho^{\mathcal{T}})\right)_{+}\right], \tag{3.12}$$

as a consequence of, e.g., [62, Lemma II.1], the inequality in (3.11) is equivalent to

$$\operatorname{Tr}\left[\left(\mathcal{A}(\rho^{\mathcal{R}}) - e^{\varepsilon}\mathcal{A}(\rho^{\mathcal{T}})\right)_{+}\right] \leq \delta. \tag{3.13}$$

By the definition in (2.11b), this leads to ε being a possible candidate for the optimization in $\overline{D}^{\delta}\left(\mathcal{A}(\rho^{\mathcal{R}}) \middle\| \mathcal{A}(\rho^{\mathcal{T}})\right)$, and thus implies

$$\overline{D}^{\delta}\left(\mathcal{A}(\rho^{\mathcal{R}}) \| \mathcal{A}(\rho^{\mathcal{T}})\right) \leq \varepsilon. \tag{3.14}$$

As this holds for every $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$, we obtain the desired implication (ε, δ)-QPP $\Rightarrow$ (3.10).

Next, we show that (3.10) implies (ε, δ)-QPP. Suppose that for all $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$, we have $\overline{D}^{\delta}\left(\mathcal{A}(\rho^{\mathcal{R}}) \| \mathcal{A}(\rho^{\mathcal{T}})\right) \leq \varepsilon$. Then, for fixed $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$, let

$$\overline{D}^{\delta}\left(\mathcal{A}(\rho^{\mathcal{R}}) \middle\| \mathcal{A}(\rho^{\mathcal{T}})\right) = \nu, \tag{3.15}$$

which implies that $\mathrm{Tr}\left[\left(\mathcal{A}(\rho^{\mathcal{R}}) - e^{\nu}\mathcal{A}(\rho^{\mathcal{T}})\right)_+\right] \leq \delta$. Recalling that $\nu \leq \varepsilon$ and noting that $\lambda \mapsto \mathrm{Tr}\left[\left(\mathcal{A}(\rho^{\mathcal{R}}) - e^{\lambda}\mathcal{A}(\rho^{\mathcal{T}})\right)_+\right]$ is a monotonically decreasing function (c.f., [29, Lemma 4.2]), we have

$$\sup_{M \in \bar{\mathcal{M}}} \mathrm{Tr}\left[M\left(\mathcal{A}(\rho^{\mathcal{R}}) - e^{\varepsilon}\mathcal{A}(\rho^{\mathcal{T}})\right)\right] = \mathrm{Tr}\left[\left(\mathcal{A}(\rho^{\mathcal{R}}) - e^{\varepsilon}\mathcal{A}(\rho^{\mathcal{T}})\right)_+\right] \tag{3.16}$$

$$\leq \mathrm{Tr}\left[\left(\mathcal{A}(\rho^{\mathcal{R}}) - e^{\nu}\mathcal{A}(\rho^{\mathcal{T}})\right)_+\right] \tag{3.17}$$

$$\leq \delta. \tag{3.18}$$

As $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$ are arbitrary, (ε, δ)-QPP follows. □

In the following remark, we highlight how the DL divergence also provides a novel characterization for classical PP.

Remark 3 (Classical PP through DL Divergence)**.** *For discrete probability distributions $p, q \in \mathcal{P}(\mathcal{Y})$, the DL divergence in Eq. (2.13b) reduces to*

$$\overline{D}_c^{\delta}(p\|q) \coloneqq \ln \inf \left\{ \lambda \geq 0 : \sum_{y \in \mathcal{Y}} \max\{p(y) - \lambda\, q(y), 0\} \leq \delta \right\}. \tag{3.19}$$

A randomized mechanism $A : \mathcal{X}^{n \times k} \to \mathcal{Y}$ is (ε, δ)-classical PP in the framework $(\mathcal{S}_c, \mathcal{Q}_c, \Theta_c)$ if for all $P_X \in \Theta_c$, $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}_c$ with $P_X(\mathcal{R}), P_X(\mathcal{T}) > 0$,

$$\overline{D}_c^{\delta}(P_{A(X)|\mathcal{R}} \| P_{A(X)|\mathcal{T}}) \leq \varepsilon, \tag{3.20}$$

where $P_{A(X)|\mathcal{R}}, P_{A(X)|\mathcal{T}}$ are the output distributions conditioned on the secret events $\mathcal{R}$ and $\mathcal{T}$, respectively. See also Remark 7 for further connections to information-theoretic quantities characterizing classical privacy frameworks.

We further note that Lemma 1 below provides a semi-definite programming characterization of the DL divergence, which in the classical case reduces to a linear program.

Remark 4 (Operational Interpretation of DL Divergence)**.** *For fixed $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$, the DL divergence $\overline{D}^{\delta}\left(\mathcal{A}(\rho^{\mathcal{R}}) \| \mathcal{A}(\rho^{\mathcal{T}})\right)$ is equal to the minimal ε that can be achieved for fixed δ via the indistinguishability condition of the QPP framework $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$ stated in (3.4).*

Remark 5 (Equivalent Formulation with Hockey-Stick Divergence)**.** *Another equivalent formulation of QPP arises as a generalization of the information-theoretic equivalence for QDP [62]. Specifically, $\mathcal{A}$ is (ε, δ)-QPP with respect to the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$, where $\bar{\mathcal{M}} = \{M : 0 \leq M \leq I\}$, if*

$$E_{e^{\varepsilon}}\left(\mathcal{A}(\rho^{\mathcal{R}}) \middle\| \mathcal{A}(\rho^{\mathcal{T}})\right) \leq \delta, \tag{3.21}$$

for all $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$, where $E_{\nu}(\rho \| \sigma) := \mathrm{Tr}[(\rho - \nu\sigma)_+]$ is the hockey-stick divergence for $\nu \geq 1$ [119]. Fixing P_X and $(\mathcal{R}, \mathcal{T})$, the quantity $E_{e^{\varepsilon}}\left(\mathcal{A}(\rho^{\mathcal{R}}) \| \mathcal{A}(\rho^{\mathcal{T}})\right)$ is the minimal δ that can be achieved for fixed ε under the indistinguishability condition from (3.4).

3.2.3 Reduction to Existing Privacy Frameworks

The proposed QPP framework subsumes other important privacy frameworks as special cases. These reductions are presented next.

Quantum DP

In QDP (Definition 5), secrets are singletons, discriminative pairs comprise states satisfying a neighboring relation, while the measurement class $\mathcal{M}$ in-

cludes all possible measurements. QPP recovers the QDP setting by making the following choices while recalling (3.3) [1]:

$$\begin{aligned} \mathcal{S} &= \mathcal{D}, \\ \mathcal{Q} &= \{(\rho, \sigma) : \rho, \sigma \in \mathcal{D}, \rho \sim \sigma\}, \\ \Theta &= \mathcal{P}_2(\mathcal{D}(\mathcal{H})), \\ \mathcal{M} &= \{M : 0 \leq M \leq I\}. \end{aligned} \tag{3.22}$$

More generally, one may add flexibility to the QDP formulation by considering other subsets Θ (i.e., $\Theta \subset \mathcal{P}_2(\mathcal{D}(\mathcal{H}))$) and $\mathcal{M}$ (i.e., $\mathcal{M} \subset \{M : 0 \leq M \leq I\}$). This can be used, for instance, to treat situations in which only certain neighboring pairs are of interest, namely, by choosing the distributions that assign positive probabilities only to those selected density operators, and situations where only certain classes of measurements are physically possible to perform. This can be interpreted as adding domain knowledge to the original QDP framework.

Quantum Local DP

In quantum local DP (QLDP) [62] [2], we choose secret pairs to be pairs of arbitrary distinct states, while the measurement class includes all possible measurements. Thus, QLDP realizes the same $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ framework as QDP, except that $\mathcal{Q} = \{(\rho, \sigma) : \rho, \sigma \in \mathcal{D}\}$ for QLDP.

[1] For each pair of states (ρ, σ), there exists at least one probability distribution that assigns positive probability for these two states, which recovers the definitions of QDP.

[2] QLDP is also known as Local differential privacy (under the 'extreme setting', as compared to standard QDP) in Section V.A of [62].

Classical PP

Consider a classical PP framework $(\mathcal{S}_c, \mathcal{Q}_c, \Theta_c)$, as specified in Definition 4. Assume that $p_X \in \Theta_c$ are discrete probability distributions over the probability space $\mathcal{P}(\mathcal{X}^{n\times k})$. Let the encoding of the database $x \in \mathcal{X}^{n\times k}$ be $\rho^x := |x\rangle\langle x|$, and denote a projective measurement operator corresponding to outcome y as $|y\rangle\langle y|$. Here note that $\{|x\rangle\}_{x\in\mathcal{X}}$ and $\{|y\rangle\}_{y\in\mathcal{Y}}$ are respective orthonormal bases formed related to the input and output alphabets of the classical PP mechanism $\mathcal{A}_c$. Then, classical PP is obtained from QPP by setting

$$\begin{aligned}
\mathcal{S} &= \{\{\rho^x : x \in \mathcal{R}_c\} : \mathcal{R}_c \in \mathcal{S}_c\},\\
\mathcal{Q} &= \{(\{\rho^x : x \in \mathcal{R}_c\}, \{\rho^x : x \in \mathcal{T}_c\}) : (\mathcal{R}_c, \mathcal{T}_c) \in \mathcal{Q}_c\}\},\\
\Theta &= \Theta_c,\\
\mathcal{M} &= \left\{\sum_{y\in\mathcal{B}} |y\rangle\langle y| : \mathcal{B} \subseteq \mathcal{Y}\right\}.
\end{aligned} \tag{3.23}$$

In this scenario, assuming the output of the algorithm is discrete, we have that

$$\mathcal{A}(\rho^x) = \sum_{y\in\mathcal{Y}, x'\in\mathcal{X}} p(y|x)|y\rangle\langle x'|\rho^x|x'\rangle\langle y| \tag{3.24}$$

where $p(y|x) = \mathbb{P}(\mathcal{A}_c(x) = y)$.

Remark 6 (Utility-Optimized Privacy Models)**.** *As is evident from above, the measurement set corresponding to classical PP entails every subset $\mathcal{B} \subseteq \mathcal{Y}$. However, when some of the outcomes are not sensitive, we may want to relax this requirement to gain utility (c.f., e.g., [90]). While classical PP does not allow for that, QPP gives extra flexibility in choosing $\mathcal{M}$ and adapting it to the application of interest. Indeed, if we only need to privatize outcomes within the set $\mathcal{Y}' \subsetneq \mathcal{Y}$, the smaller measurement set $\mathcal{M} = \left\{\sum_{y\in\mathcal{Y}'} |y\rangle\langle y| : \mathcal{B} \subseteq \mathcal{Y}'\right\}$ is sufficient.*

3.3 Datta–Leditzky Information Spectrum Divergence

We now focus on the DL divergence [29], whose operational interpretation in terms of QPP was provided in the previous section (see Remark 4), and we study structural properties thereof, which will be useful when analyzing the QPP framework. We first formulate a primal and dual SDP to compute the DL divergence and then use that to prove joint-quasi convexity, the data-processing inequality under positive, trace-preserving maps, and connections to the smooth max-relative entropy.

3.3.1 SDP Formulations

We now present several SDPs for computing the DL divergence in (2.11b), which may be of independent interest. (Recall that the other DL divergence in (2.11a) is easily obtained by applying the equality in (2.12).)

Lemma 1 (SDP Formulation of the DL Divergence)**.** *For $\delta \in (0,1)$, a state ρ, and a PSD operator σ, the following equalities hold*

$$\overline{D}^{\delta}(\rho\|\sigma) = \ln \inf_{\lambda, Z\geq 0} \{\lambda : \mathrm{Tr}[Z] \leq \delta,\ Z \geq \rho - \lambda\sigma\} \tag{3.25a}$$

$$= \ln \sup_{\mu, W\geq 0} \{\mathrm{Tr}[W\rho] - \mu\delta : \mathrm{Tr}[W\sigma] \leq 1,\ W \leq \mu I\}\,. \tag{3.25b}$$

Proof. Considering (2.13b), fix $\lambda > 0$ and first observe that

$$\mathrm{Tr}[(\rho - \lambda\sigma)_+] = \sup_{\Lambda:\, 0\leq\Lambda\leq I} \mathrm{Tr}[\Lambda\,(\rho - \lambda\sigma)]. \tag{3.26}$$

Indeed, this follows because, for every $0 \leq \Lambda \leq I$, we have that

$$\begin{aligned} \operatorname{Tr}[\Lambda\,(\rho - \lambda\sigma)] &= \operatorname{Tr}[\Lambda\,((\rho - \lambda\sigma)_+ - (\rho - \lambda\sigma)_-)] \\ &\leq \operatorname{Tr}[\Lambda\,(\rho - \lambda\sigma)_+] \\ &\leq \operatorname{Tr}[(\rho - \lambda\sigma)_+]\,, \end{aligned} \tag{3.27}$$

and the inequalities above are all attained by setting Λ to be the projection onto the support of $(\rho - \lambda\sigma)_+$. The SDP dual of this quantity is given by

$$\operatorname{Tr}[(\rho - \lambda\sigma)_+] = \inf_{Z \geq 0} \{\operatorname{Tr}[Z] : Z \geq \rho - \lambda\sigma\}\,. \tag{3.28}$$

We then find from (2.11b), (2.13b), and (3.28) that

$$\begin{aligned} \overline{D}_s^{\delta}(\rho\|\sigma) &= \ln \inf \{\lambda \geq 0 : \operatorname{Tr}[(\rho - \lambda\sigma)_+] \leq \delta\} \\ &= \ln \inf_{\lambda, Z \geq 0} \{\lambda : \operatorname{Tr}[Z] \leq \delta,\ Z \geq \rho - \lambda\sigma\}\,, \end{aligned} \tag{3.29}$$

which completes the proof of (3.25a).

The dual forms of these optimization problems are derived from the canonical primal and dual formulations of SDPs, which are respectively given by (c.f., [75, Definition 2.20])

$$\begin{aligned} &\inf_{Y \geq 0} \left\{\operatorname{Tr}[BY] : \Phi^\dagger(Y) \geq A\right\}, \\ &\sup_{X \geq 0} \{\operatorname{Tr}[AX] : \Phi(X) \leq B\}\,, \end{aligned} \tag{3.30}$$

where A and B are Hermitian matrices and Φ is a Hermiticity-preserving superoperator. Comparing the former to (3.29), we make the following choices so that

the general optimization problem recovers (3.29) (inside the logarithm):

$$Y = \begin{bmatrix} \lambda & 0 \\ 0 & Z \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \tag{3.31}$$

$$\Phi^\dagger(Y) = \begin{bmatrix} -\operatorname{Tr}[Z] & 0 \\ 0 & Z + \lambda\sigma \end{bmatrix}, \tag{3.32}$$

$$A = \begin{bmatrix} -\delta & 0 \\ 0 & \rho \end{bmatrix}. \tag{3.33}$$

Then, setting

$$X = \begin{bmatrix} \mu & 0 \\ 0 & W \end{bmatrix}, \tag{3.34}$$

we solve for the map $\Phi(X)$ to find that

$$\operatorname{Tr}[X\Phi^\dagger(Y)] = \operatorname{Tr}\left[\begin{bmatrix} \mu & 0 \\ 0 & W \end{bmatrix}\begin{bmatrix} -\operatorname{Tr}[Z] & 0 \\ 0 & Z + \lambda\sigma \end{bmatrix}\right] \tag{3.35}$$

$$= -\mu \operatorname{Tr}[Z] + \operatorname{Tr}[W(Z + \lambda\sigma)] \tag{3.36}$$

$$= \operatorname{Tr}[(W - \mu I) Z] + \lambda \operatorname{Tr}[W\sigma] \tag{3.37}$$

$$= \operatorname{Tr}\left[\begin{bmatrix} \lambda & 0 \\ 0 & Z \end{bmatrix}\begin{bmatrix} \operatorname{Tr}[W\sigma] & 0 \\ 0 & W - \mu I \end{bmatrix}\right] \tag{3.38}$$

$$= \operatorname{Tr}[Y\Phi(X)], \tag{3.39}$$

so that

$$\Phi(X) = \begin{bmatrix} \operatorname{Tr}[W\sigma] & 0 \\ 0 & W - \mu I \end{bmatrix}. \tag{3.40}$$

Plugging into the dual form, we obtain

$$\sup_{X \geq 0} \{\operatorname{Tr}[AX] : \Phi(X) \leq B\} = \sup_{\mu, W \geq 0} \{\operatorname{Tr}[W\rho] - \mu\delta : \operatorname{Tr}[W\sigma] \leq 1, W \leq \mu I\}. \tag{3.41}$$

Choose $\mu = \mu_1 \in (0, 1)$ and $W = \mu_2 I$ such that $\mu_1\delta < \mu_2 < \mu_1$, as a strictly feasible solution to the above. For the other SDP formulation from (3.29), set λ

such that $\mathrm{Tr}[(\rho - \lambda\sigma)_+] \leq \delta$, and $Z = (\rho - \lambda\sigma)_+ \geq 0$ as a feasible solution. By Slater's condition, we conclude that strong duality holds, and the primal and dual optimal values coincide. □

Corollary 1 (Another formulation of the DL divergence)**.** *DL divergence has the following equivalent formulation:*

$$\overline{D}^{\delta}(\rho\|\sigma) = \ln \sup_{0\leq W\leq I, \mathrm{Tr}[W\rho]\geq\delta} \frac{\mathrm{Tr}[W\rho] - \delta}{\mathrm{Tr}[W\sigma]}. \tag{3.42}$$

Proof. Consider the SDP formulation in (3.25b) and set $W' = \frac{W}{\mu}$ therein to arrive at

$$\overline{D}^{\delta}(\rho\|\sigma) = \ln \sup_{\mu, W'\geq 0} \left\{ \begin{array}{c} \mu\,\mathrm{Tr}[W'\rho] - \mu\delta : \\ \mu\,\mathrm{Tr}[W'\sigma] \leq 1,\ W' \leq I \end{array} \right\} \tag{3.43}$$

$$= \ln \sup_{0\leq W'\leq I, \mathrm{Tr}[W'\rho]\geq\delta} \frac{\mathrm{Tr}[W'\rho] - \delta}{\mathrm{Tr}[W'\sigma]}, \tag{3.44}$$

where the last equality follows from identifying that $\mu = 1/\,\mathrm{Tr}[W'\sigma]$ is the μ that maximizes the former, given that $\mathrm{Tr}[W'\rho] \geq \delta$. When $\mathrm{Tr}[W'\rho] < \delta$, optimum $\mu = 0$ and the objective within the supremum becomes zero. Replacing W' by W concludes the proof. □

Remark 7 (Approximate-Max Divergence)**.** *In [41], δ-approximate-max divergence is defined as*

$$D_{\infty}^{\delta}(p_Y\|p_Z) \coloneqq \ln \max_{S\in\mathrm{Supp}(Y), \Pr[Y\in S]\geq\delta} \frac{\Pr[Y \in S] - \delta}{\Pr[Z \in S]}, \tag{3.45}$$

where Y and Z are random variables distributed according to $Y \sim p_Y$ and $Z \sim p_Z$ By substituting classical distributions into Corollary 1, we observe that DL divergence reduces to approximate-max divergence. Note that approximate-max divergence has been used to characterize (ε, δ)-(classical) DP in [41, Remark 3.1]. Thus, this showcases that the equivalence we established for QPP with DL divergence herein reduces to the existing equivalence on (classical) DP.

3.3.2 Properties

We derive several properties of the DL divergence from (2.13b), which are subsequently used in the analysis of the QPP framework. Basic properties of the DL divergence, including the data-processing inequality, have been proven in [29, Proposition 4.3]. Here, we generalize the data-processing inequality to hold for arbitrary positive, trace non-increasing maps (beyond the set of quantum channels) and also establish joint-quasi convexity of the DL divergence, along with its connection to the smooth max-relative entropy (recall the definition in (2.16)). The proofs of these properties rely on the SDP formulation from Lemma 1.

Proposition 2 (Properties of the DL Divergence)**.** *Fix $\delta \in (0,1)$, and let $\rho, \rho_1, \ldots, \rho_k$ and $\sigma, \sigma_1, \ldots, \sigma_k$ be two collections of states and PSD operators, respectively. The DL divergence in* (2.13b) *satisfies the following properties:*

1. *Data-processing inequality: For every positive, trace non-increasing map $\mathcal{N}$, we have*

$$\overline{D}^{\delta}(\rho\|\sigma) \geq \overline{D}^{\delta}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)). \tag{3.46}$$

2. *Joint-quasi convexity: Let $p_i \in [0,1]$, for $i \in \{1,\ldots,k\}$, with $\sum_{i=1}^{k} p_i = 1$. Then*

$$\overline{D}^{\delta}\left(\sum_{i=1}^{k} p_i\rho_i \middle\| \sum_{i=1}^{k} p_i\sigma_i\right) \leq \max_i \overline{D}^{\delta}(\rho_i\|\sigma_i)\,, \tag{3.47}$$

and, more generally,

$$\overline{D}^{\delta'}\left(\sum_{i=1}^{k} p_i\rho_i \middle\| \sum_{i=1}^{k} p_i\sigma_i\right) \leq \max_i \overline{D}^{\delta_i}(\rho_i\|\sigma_i)\,, \tag{3.48}$$

where $\delta' := \sum_{i=1}^{k} p_i\delta_i$ with $\delta_1, \ldots, \delta_k \in (0,1)$.

3. *Relation to smooth max-relative entropy:*

$$\overline{D}^{\delta}(\rho\|\sigma) \leq D_{\max}^{\delta}(\rho\|\sigma) \leq \overline{D}^{\delta'}(\rho\|\sigma) - \ln(1-\delta')\,, \tag{3.49}$$

where $\delta' := 1 - \sqrt{1-\delta^2} \in (0,1)$*, and the second inequality above can be equivalently written as*

$$D_{\max}^{\sqrt{\delta(2-\delta)}}(\rho\|\sigma) \leq \overline{D}^{\delta}(\rho\|\sigma) - \ln(1-\delta)\,. \tag{3.50}$$

4. *Quasi subadditivity: Let* $\delta_1, \delta_2 \in (0,1)$ *satisfy* $\delta_1' + \delta_2' < 1$*, with* $\delta_i' := \sqrt{\delta_i(2-\delta_i)} \in (0,1)$ *for* $i \in \{1,2\}$*. Then*

$$\overline{D}^{\delta_1'+\delta_2'}(\rho_1\otimes\rho_2\|\sigma_1\otimes\sigma_2) \;\leq\; \overline{D}^{\delta_1}(\rho_1\|\sigma_1) + \overline{D}^{\delta_2}(\rho_2\|\sigma_2) - \ln\big((1-\delta_1)(1-\delta_2)\big). \tag{3.51}$$

Furthermore,

(a) if $\delta_1 = \delta_2 = 0$*, then*

$$\overline{D}^{0}(\rho_1\otimes\rho_2\|\sigma_1\otimes\sigma_2) \leq \overline{D}^{0}(\rho_1\|\sigma_1) + \overline{D}^{0}(\rho_2\|\sigma_2). \tag{3.52}$$

(b) if σ_1, σ_2 *are states, then*

$$\overline{D}^{\delta}(\rho_1\otimes\rho_2\|\sigma_1\otimes\sigma_2) \leq \overline{D}^{\delta_1}(\rho_1\|\sigma_1) + \overline{D}^{\delta_2}(\rho_2\|\sigma_2), \tag{3.53}$$

where

$$\delta := \min\Big\{\delta_1 + e^{\overline{D}^{\delta_1}(\rho_1\|\sigma_1)}\delta_2,\; \delta_2 + e^{\overline{D}^{\delta_2}(\rho_2\|\sigma_2)}\delta_1\Big\}. \tag{3.54}$$

Proof. Property 1: The statement was proven in [29, Proposition 4.3] for a quantum channel $\mathcal{N}$ by using the inequality

$$\mathrm{Tr}[(\rho - e^{\gamma}\sigma)_+] \geq \mathrm{Tr}[(\mathcal{N}(\rho) - e^{\gamma}\mathcal{N}(\sigma))_+]\,, \tag{3.55}$$

which holds for all $\gamma \in \mathbb{R}$. Here, we prove the data-processing inequality, but we generalize it to hold for a positive, trace non-increasing map $\mathcal{N}$. Our derivation relies on the SDP formulation of the DL divergence from (3.25a).

Let $\lambda^\star$ and $Z^\star$ be optimal choices[3] in the optimization for $\overline{D}^\delta(\rho\|\sigma)$, so that $\overline{D}^\delta(\rho\|\sigma) = \ln\lambda^\star$, $Z^\star \geq \rho - \lambda^\star\sigma$ with $\mathrm{Tr}[Z^\star] \leq \delta$, and $Z^\star \geq 0$ (indeed, note that the infimum is achieved with $\mathrm{Tr}[(\rho - \lambda^\star\sigma)_+] = \delta$). Since $Z^\star - (\rho - \lambda^\star\sigma) \geq 0$, it follows that $\mathcal{N}(Z^\star - (\rho - \lambda^\star\sigma)) \geq 0$ from the assumption that $\mathcal{N}$ is a positive map. Consequently, we obtain

$$Z' \coloneqq \mathcal{N}(Z^\star) \geq \mathcal{N}(\rho) - \lambda^\star\mathcal{N}(\sigma). \tag{3.56}$$

Furthermore, $Z' \geq 0$ since $Z^\star \geq 0$ and $\mathcal{N}$ is a positive map. Additionally, since $\mathcal{N}$ is trace non-increasing, it follows that

$$\mathrm{Tr}[Z'] \leq \mathrm{Tr}[Z^\star] \leq \delta. \tag{3.57}$$

Thus, $\lambda^\star$ is a feasible point for $\overline{D}^\delta(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))$. We conclude the proof by noting that the quantity $\overline{D}^\delta(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))$ involves a minimization over all such feasible points, implying the desired inequality:

$$\overline{D}^\delta(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \leq \ln(\lambda^\star) = \overline{D}^\delta(\rho\|\sigma)\,. \tag{3.58}$$

Note that this property can also be derived using the proof of [119, Lemma 4].

Property 2: We again consider the SDP from (3.25a). Let $\lambda_i^\star$ and $Z_i^\star$ be optimal for $\overline{D}^\delta(\rho_i\|\sigma_i)$, so that $\overline{D}^\delta(\rho_i\|\sigma_i) = \ln(\lambda_i^\star)$, $Z_i^\star \geq \rho_i - \lambda_i^\star\sigma_i$ with $\mathrm{Tr}\left[Z_i^\star\right] \leq \delta$, and $Z_i^\star \geq 0$. Define

$$Z \coloneqq \sum_{i=1}^k p_i Z_i^\star \geq \sum_{i=1}^k p_i\rho_i - \sum_{i=1}^k \lambda_i^\star p_i\sigma_i. \tag{3.59}$$

and observe that $\mathrm{Tr}[Z] \leq \delta$ and $Z \geq 0$. This implies that $Z \geq \sum_{i=1}^k p_i\rho_i - \max_i \lambda_i^\star \sum_{j=1}^k p_j\sigma_j$, which suggests that $\max_i \lambda_i^\star$ and Z are (candidate) infimizers in the SDP formulation of $\overline{D}^\delta\left(\sum_{i=1}^k p_i\rho_i\middle\|\sum_{i=1}^k p_i\sigma_i\right)$. Consequently, we obtain

$$\overline{D}^\delta\left(\sum_{i=1}^k p_i\rho_i\middle\|\sum_{i=1}^k p_i\sigma_i\right) \leq \ln\left(\max_i \lambda_i^\star\right) = \max_i \overline{D}^\delta(\rho_i\|\sigma_i)\,. \tag{3.60}$$

[3]When the DL divergence is finite, the infimum is achieved by a standard continuity plus compactness argument. The stated relations trivially hold when the DL divergence is infinite.

The proof of the general case follows along the same lines by observing that $\mathrm{Tr}[Z] \leq \sum_{i=1}^{k} p_i \delta_i$.

Property 3: From [130, Appendix B], we have that

$$D_{\max}^{\delta}(\rho\|\sigma) = \ln \inf_{\lambda,\widetilde{\rho},Y\geq 0} \left\{ \begin{array}{l} \lambda : \widetilde{\rho} \leq \lambda\sigma, \;\; \mathrm{Tr}[Y] \leq \delta, \\ \mathrm{Tr}[\widetilde{\rho}] = 1, \;\; Y \geq \rho - \widetilde{\rho} \end{array} \right\}. \tag{3.61}$$

Let λ, Y, and $\widetilde{\rho}$ be arbitrary operators satisfying the constraints for $D_{\max}^{\delta}(\rho\|\sigma)$. Then by combining the inequalities $\widetilde{\rho} \leq \lambda\sigma$ and $Y \geq \rho - \widetilde{\rho}$, we get

$$Y \geq \rho - \lambda\sigma. \tag{3.62}$$

We see that λ and Y satisfy the constraints needed for λ and Z, respectively, in the SDP for $\overline{D}^{\delta}(\rho\|\sigma)$, whereby

$$\overline{D}^{\delta}(\rho\|\sigma) \leq \lambda. \tag{3.63}$$

Since the argument holds for all λ, Y, and $\widetilde{\rho}$ satisfying the constraints in the definition of $D_{\max}^{\delta}(\rho\|\sigma)$, we further obtain

$$\overline{D}^{\delta}(\rho\|\sigma) \leq D_{\max}^{\delta}(\rho\|\sigma). \tag{3.64}$$

The proof is concluded by invoking the following lemma (proven in Appendix A.1).

Lemma 2. *Fix $\lambda > 0$, let ρ be a state and σ a positive semi-definite operator, and define $\delta := \mathrm{Tr}[(\rho - \lambda\sigma)_+]$. Then*

$$D_{\max}^{\sqrt{\delta(2-\delta)}}(\rho\|\sigma) \leq \ln\lambda - \ln(1-\delta). \tag{3.65}$$

For fixed $\delta \in (0,1)$, by definition, we have $\overline{D}^{\delta}(\rho\|\sigma) = \ln(\lambda^\star)$ with $\delta = \mathrm{Tr}[(\rho - \lambda^\star\sigma)_+]$. With that, Lemma 2 with the reparametrization $\delta \to 1 - \sqrt{1-\delta^2}$,

yields

$$D^{\delta}_{\max}(\rho\|\sigma) \leq \overline{D}^{1-\sqrt{1-\delta^2}}(\rho\|\sigma) + \ln\left(\frac{1}{\sqrt{1-\delta^2}}\right). \tag{3.66}$$

This completes the proof.

Property 4: This follows by invoking Property 3 and using the fact that the smooth max-relative entropy satisfies subadditivity (Appendix A.2) with

$$D^{\delta_1+\delta_2}_{\max}(\rho_1\otimes\rho_2\|\sigma_1\otimes\sigma_2) \leq D^{\delta_1}_{\max}(\rho_1\|\sigma_1) + D^{\delta_2}_{\max}(\rho_2\|\sigma_2). \tag{3.67}$$

Part (a) now follows by taking the limits $\delta_1 \to 0$ and $\delta_2 \to 0$ in (3.67), and applying Property 3.

To prove Part (b), we use the SDP formulation in (3.25a). Let $\overline{D}^{\delta_i}(\rho_i\|\sigma_i) = \ln(\lambda_i^\star)$ for $i \in \{1,2\}$. It follows that $Z_i \geq \rho_i - \lambda_i^\star\sigma_i$ with $\mathrm{Tr}[Z_i] \leq \delta$ and $Z_i \geq 0$. Consider that

$$\begin{aligned}(\rho_1\otimes\rho_2) - \lambda_1^\star\lambda_2^\star(\sigma_1\otimes\sigma_2) &= (\rho_1\otimes\rho_2) - \lambda_1^\star\sigma_1\otimes\rho_2 + \lambda_1^\star\sigma_1\otimes\rho_2 - \lambda_1^\star\lambda_2^\star(\sigma_1\otimes\sigma_2)\\ &= (\rho_1-\lambda_1^\star\sigma_1)\otimes\rho_2 + \lambda_1^\star\sigma_1\otimes(\rho_2-\lambda_2^\star\sigma_2)\\ &\leq Z_1\otimes\rho_2 + \lambda_1^\star\sigma_1\otimes Z_2 =: Z. \end{aligned} \tag{3.68}$$

Observe that $Z \geq 0$ and $\mathrm{Tr}[Z] = \mathrm{Tr}[Z_1] + \lambda_1^\star\mathrm{Tr}[Z_2]$, since $\mathrm{Tr}[\rho_1] = \mathrm{Tr}[\sigma_1] = 1$. Consequently, we have $\mathrm{Tr}[Z] \leq \delta_1 + \lambda_1^\star\delta_2$, and $\lambda_1^\star\lambda_2^\star$ is a candidate infimizer. For $\delta' = \delta_1 + \lambda_1^\star\delta_2$, we now arrive at

$$\overline{D}^{\delta'}(\rho_1\otimes\rho_2\|\sigma_1\otimes\sigma_2) \leq \ln(\lambda_1^\star\lambda_2^\star) \tag{3.69}$$

$$= \ln(\lambda_1^\star) + \ln(\lambda_2^\star) \tag{3.70}$$

$$= \overline{D}^{\delta_1}(\rho_1\|\sigma_1) + \overline{D}^{\delta_2}(\rho_2\|\sigma_2)\,. \tag{3.71}$$

The above holds for $\delta' = \delta_2 + \lambda_2^\star\delta_1$ as well, by adding and subtracting $\lambda_2^\star\rho_1\otimes\sigma_1$ instead of $\lambda_1^\star\sigma_1\otimes\rho_2$, and then following the same argument. □

3.4 Properties and Mechanisms for QPP

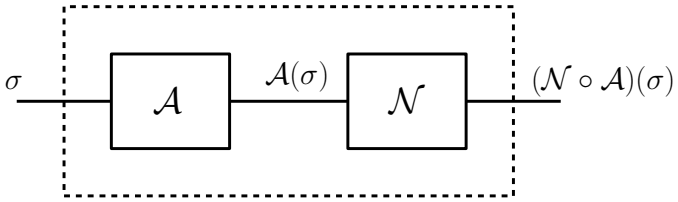
3.4.1 Properties of QPP Mechanisms

Modern guidelines for privacy frameworks [76] render properties such as convexity and post-processing (also known as transformation invariance) as basic requirements for privacy frameworks. Composability is another important property, which implies that a combination of privacy mechanisms is itself private. These properties are known to hold for the classical mutual information PP framework, and all of them, except for composability, hold for the classical PP framework; c.f., [92, Theorem 2] and [77, Theorem 5.1], respectively.

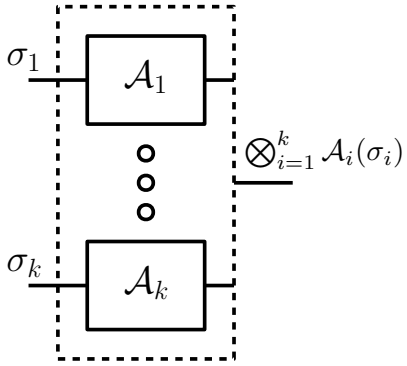
Before proving these properties for the QPP framework, we discuss their operational interpretation. Convexity means that applying a QPP mechanism that is randomly chosen from a given set of such mechanisms still satisfies QPP. Post-processing ensures that passing the output of a QPP mechanism $\mathcal{A}$ through a channel $\mathcal{N}$ preserves QPP; see Fig. 3.2a. Parallel composability is illustrated in Fig. 3.2b and guarantees that QPP holds after applying

$$\mathcal{A}^{(k)} := \bigotimes_{i=1}^{k} \mathcal{A}_i = \mathcal{A}_1 \otimes \mathcal{A}_2 \otimes \cdots \otimes \mathcal{A}_k \tag{3.72}$$

to the input $\rho^{X_1} \otimes \rho^{X_2} \otimes \cdots \otimes \rho^{X_k}$, with $X_i \sim P_X \in \Theta$, where each X_i is independently chosen. Informally, the semantic meaning of this property is that after applying $\mathcal{A}^{(k)}$, the same conclusions can be drawn about the input $\rho^{X_1} \otimes \rho^{X_2} \otimes \cdots \otimes \rho^{X_k}$ regardless of whether each ρ^{X_i} belongs to $\mathcal{R}_i$ or $\mathcal{T}_i$, where $(\mathcal{R}_i, \mathcal{T}_i) \in \mathcal{Q}$ for all



(a)



(b)

Figure 3.2: Properties of QPP mechanisms: (a) refers to post-processing of QPP algorithm $\mathcal{A}$; If $\mathcal{A}$ satisfies QPP, then $\mathcal{N} \circ \mathcal{A}$ also satisfies QPP. (b) refers to parallel composition of k QPP mechanisms; composition of k mechanisms independently in a parallel fashion satisfies QPP if each $\mathcal{A}_i$ satisfies QPP.

$i \in \{1, \ldots, k\}$. In this setting, the set of discriminative pairs is taken as

$$Q^{(k)} := \left\{ \begin{array}{c} \mathcal{R}^{(k)} := (\mathcal{R}_1, \ldots, \mathcal{R}_k), \\ (\mathcal{R}^{(k)}, \mathcal{T}^{(k)}) : \mathcal{T}^{(k)} := (\mathcal{T}_1, \ldots, \mathcal{T}_k) \\ \forall i \in \{1, \ldots, k\} \ (\mathcal{R}_i, \mathcal{T}_i) \in Q \end{array} \right\}. \tag{3.73}$$

Furthermore, the class of product measurements is $\bigotimes_{i=1}^{k} \mathcal{M}_i$ (i.e., the output of algorithm $\mathcal{A}_i$ is followed by a measurement from $\mathcal{M}_i$, for all $i \in \{1, \ldots, k\}$), while the set of all possible measurements on the k systems, including joint measurements, is denoted by $\bar{\mathcal{M}}^k$. We note here that one could consider other classes of limited measurements, such as local operations and classical communication (LOCC) measurements and positive-partial-transpose (PPT) measurements [87].

The formal statement of these properties is as follows.

Theorem 1 (Properties of QPP Mechanisms)**.** *The following properties hold:*

1. *Convexity: Let $\mathcal{A}_1, \ldots, \mathcal{A}_k$ be (ε, δ)-QPP mechanisms in the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$. Take I to be a k-ary categorical random variable with probability distribution $(p_1, \ldots, p_k)$. Then the mechanism $\mathcal{A} := \mathcal{A}_I$ (i.e., $\mathcal{A} = \mathcal{A}_i$ with probability p_i, for $i \in \{1, \ldots, k\}$) also satisfies (ε, δ)-QPP in the same framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$.*

2. *Post-processing: If a mechanism $\mathcal{A}$ satisfies (ε, δ)-QPP in the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$, then, for a quantum channel $\mathcal{N}$, the processed mechanism $\mathcal{N} \circ \mathcal{A}$ also satisfies (ε, δ)-QPP in the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M}')$, where $\mathcal{M}' \subseteq \left\{M' : \mathcal{N}^\dagger(M') \in \mathcal{M}\right\}$.*

3. *Parallel composability (non-adaptive): Let $\mathcal{A}_1, \ldots, \mathcal{A}_k$ be mechanisms such that $\mathcal{A}_i$ is $(\varepsilon_i, \delta_i)$-QPP in the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M}_i)$, for each $i \in \{1, \ldots, k\}$. Then the composed mechanism*

$$\mathcal{A}^{(k)} : \bigotimes_{i=1}^{k} \sigma_i \mapsto \mathcal{A}_1(\sigma_1) \otimes \cdots \otimes \mathcal{A}_k(\sigma_k) \tag{3.74}$$

satisfies $\left(\sum_{i=1}^k \varepsilon_i, \sum_{i=1}^k \delta_i\right)$-QPP in the framework $\left(\mathcal{S}, \mathcal{Q}^{(k)}, \Theta, \bigotimes_{i=1}^k \mathcal{M}_i\right)$.

Proof. See Appendix A.3. □

More broadly, parallel composition (i.e., Property 3 of Theorem 1) holds under separable measurements that are defined as follows:

$$\left\{\sum_j M_1^{(j)} \otimes \cdots \otimes M_k^{(j)} : \forall i \in \{1, \ldots, k\} \ \sum_j M_i^{(j)} \in \mathcal{M}_i\right\}. \tag{3.75}$$

where product measurements considered in Theorem 1 is a special case.

The latter two properties of Theorem 1 change if one considers a measurement class that comprises all possible measurements $\bar{\mathcal{M}}^k$, as opposed to only

product measurements. This is one of the main distinctions between the semi-classical and quantum cases, where, for the latter, joint measurements may infer more information and thus privacy degrades. The following theorem accounts for this latter scenario.

Theorem 2 (Properties of QPP with $\mathcal{M} = \bar{\mathcal{M}}$)**.** *The following properties hold for the case in which the measurement class is* $\bar{\mathcal{M}}$*:*

1. Convexity: *Let* $\mathcal{A}_1, \ldots, \mathcal{A}_k$ *be* (ε, δ)*-QPP mechanisms in the framework* $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$*. Take* I *to be a k-ary categorical random variable with parameters* $(p_1, \ldots, p_k)$*. Then the mechanism* $\mathcal{A} \coloneqq \mathcal{A}_I$ *(i.e.,* $\mathcal{A} = \mathcal{A}_i$ *with probability* p_i*, for* $i \in \{1, \ldots, k\}$*) also satisfies* (ε, δ)*-QPP in the framework* $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$*.*

2. Post-processing: *If a mechanism* $\mathcal{A}$ *satisfies* (ε, δ)*-QPP in the framework* $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$*, then, for a quantum channel* $\mathcal{N}$*, the mechanism* $\mathcal{N} \circ \mathcal{A}$ *also satisfies* (ε, δ)*-QPP in the framework* $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$*.*

3. Parallel composability: *If* $\mathcal{A}_i$ *satisfies* $(\varepsilon_i, \delta_i)$*-QPP in* $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$ *for* $i \in \{1, 2\}$*, then the composed mechanism* $\mathcal{A}_1 \otimes \mathcal{A}_2$ *satisfies* (ε', δ')*-QPP in the framework* $\left(\mathcal{S}, \mathcal{Q}^{(2)}, \Theta, \bar{\mathcal{M}}^2\right)$ *where*

$$\varepsilon' \coloneqq \varepsilon_1 + \varepsilon_2 + \ln\left(\frac{1}{(1-\delta_1)(1-\delta_2)}\right), \tag{3.76}$$

$$\delta' \coloneqq \sqrt{\delta_1(2-\delta_1)} + \sqrt{\delta_2(2-\delta_2)}. \tag{3.77}$$

and $\mathcal{A}_1 \otimes \mathcal{A}_2$ *also satisfies* $(\varepsilon_1 + \varepsilon_2, \delta)$*-QPP where*

$$\delta \coloneqq \min\{\delta_1 + e^{\varepsilon_1}\delta_2, \delta_2 + e^{\varepsilon_2}\delta_1\}. \tag{3.78}$$

Observe that, if $\delta_i = 0$ *for* $i \in \{1, 2\}$*, then* $\mathcal{A}_1 \otimes \mathcal{A}_2$ *satisfies* $(\varepsilon_1 + \varepsilon_2)$*-QPP for the parallel composed framework.*

Proof. The proof of Theorem 2 relies on properties of the DL divergence established in Proposition 2. Items (1), (2), and (3) follow from joint quasi-convexity (Property 2), data processing (Property 1), and quasi subadditivity (Property 4), respectively. □

Remark 8 (Comparison to Existing Results)**.** *In [62, Corollary III.3], the parallel composition of two mechanisms that satisfy $(\varepsilon_i, \delta_i)$-QDP for $i \in \{1, 2\}$ is shown to be $(\varepsilon_1 + \varepsilon_2, \delta)$-QDP, where δ is given in (3.78). The proof technique is, however, different from ours. Property 3 of Theorem 1 also reveals that if one considers a restricted class of measurements (e.g., product measurements), then it is possible to achieve tighter privacy guarantees (namely, $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-QDP) than those obtained when allowing all joint measurements on the two systems. Also note that δ' in (3.77) is independent of ε_1 and ε_2, whereas δ in (3.78) depends on them. Depending on the particular values that the parameters ε_i and δ_i take, for $i \in \{1, 2\}$, one of these aforementioned results provides sharper privacy guarantees.*

Adaptive Composability

Adaptive composition refers to the case when each subsequently composed mechanism is chosen based on the outputs of the preceding ones. The goal is to quantify the overall privacy leakage at the output of the adaptively composed mechanism. This idea has been studied in detail for classical privacy settings [41], and here we explore it for the quantum setting with QPP.

We first focus on the setting depicted in Fig. 3.3. Fix $X_i \sim P_X \in \Theta$ for $i \in \{1, 2\}$, which are independently chosen, and let the input state be

$$\sigma_I := \rho^{X_1} \otimes \rho^{X_2}. \tag{3.79}$$

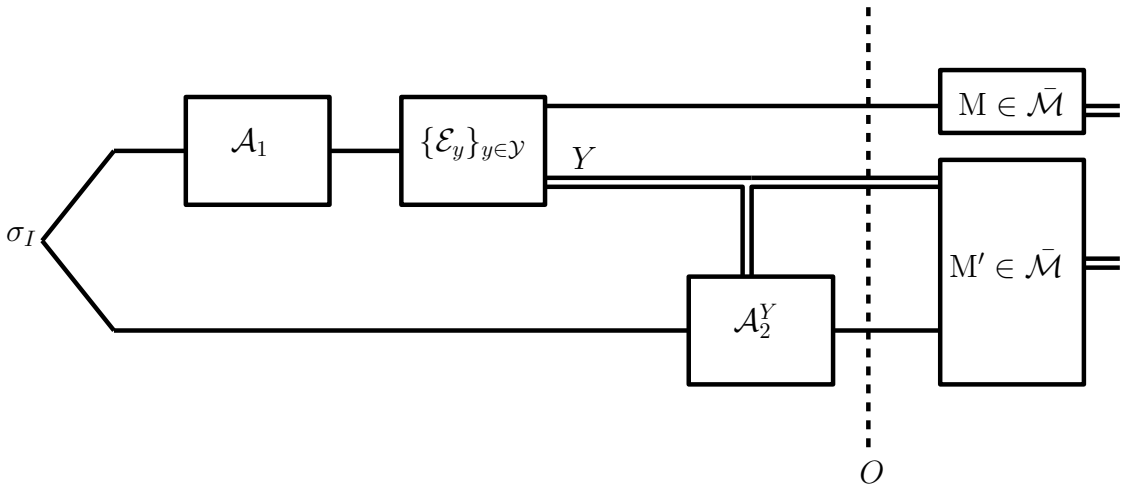


Figure 3.3: Setup for adaptive composition: On the top system, the channel $\mathcal{A}_1$ is followed by the quantum instrument $\{\mathcal{E}_y\}_{y\in\mathcal{Y}}$, and then the random classical outcome Y is used to choose the channel $\mathcal{A}_2^Y$. In this setting, we analyze how well an adversary can learn properties of the input state σ_I by applying measurements on the output state.

On the top subsystem in Fig. 3.3, the channel $\mathcal{A}_1$ is followed by the quantum instrument $\{\mathcal{E}_y\}_{y\in\mathcal{Y}}$, which is a collection of completely positive maps such that the sum map

$$\overline{\mathcal{E}} := \sum_{y\in\mathcal{Y}} \mathcal{E}_y \tag{3.80}$$

is trace preserving [30, 31, 103]. Depending on the measurement outcome y, the channel $\mathcal{A}_2^y$ is chosen and applied to the bottom subsystem. The combined output state at stage O, as marked in the figure, is

$$\sigma_O := \sum_{y\in\mathcal{Y}} \mathcal{E}^y\Big(\mathcal{A}_1(\rho^{X_1})\Big) \otimes |y\rangle\langle y| \otimes \mathcal{A}_2^y(\rho^{X_2}). \tag{3.81}$$

We focus on the adaptive composition of two quantum mechanisms in the above-described setting. Suppose that $\mathcal{A}_1$ is an $(\varepsilon_1, \delta_1)$-QPP mechanism in the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M}_1)$. Suppose furthermore that, for each outcome $y \in \mathcal{Y}$, the mechanism $\mathcal{A}_2^y$ satisfies $(\varepsilon_2, \delta_2)$-QPP in the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M}_2)$ in the

following sense: for all $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$ and $M \in \mathcal{M}_2$,

$$\mathrm{Tr}\Big[M\mathcal{A}_2^y(\rho^{\mathcal{R}})\Big] \leq e^{\varepsilon_2}\mathrm{Tr}\Big[M\mathcal{A}_2^y(\rho^{\mathcal{T}})\Big] + \delta_2. \tag{3.82}$$

Under adaptive composition, we want to guarantee the indistinguishability of pairs of states

$$\sigma_I^{\mathcal{R}} := \rho^{\mathcal{R}_1} \otimes \rho^{\mathcal{R}_2} \text{ and } \sigma_I^{\mathcal{T}} := \rho^{\mathcal{T}_1} \otimes \rho^{\mathcal{T}_2}, \tag{3.83}$$

where $(\mathcal{R}_i, \mathcal{T}_i) \in \mathcal{Q}$ for $i \in \{1, 2\}$. This means, informally, that the adversary would draw the same conclusions regardless of whether ρ^{X_i} belongs to $\mathcal{R}_i$ or $\mathcal{T}_i$, for $i \in \{1, 2\}$, when the initial input σ_I to the system in Fig. 3.3 is given by (3.79). The following proposition provides parameters under which QPP of the adaptively composed mechanism is guaranteed.

Proposition 3 (Adaptive composition of QPP)**.** *Fix the framework* $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$*. Suppose that* $\mathcal{A}_1$ *satisfies* $(\varepsilon_1, \delta_1)$*-QPP and* $\mathcal{A}_2^y$ *satisfies* $(\varepsilon_2, \delta_2)$*-QPP for every measurement outcome* y*, as in* (3.82)*. Then the mechanism in Fig. 3.3 satisfies* $(\varepsilon_1+\varepsilon_2, \delta_2+\delta_1|\mathcal{Y}|)$*-QPP in the framework* $(\mathcal{S}, \mathcal{Q} \times \mathcal{Q}, \Theta, \bar{\mathcal{M}} \otimes \bar{\mathcal{M}})$ *where* $|\mathcal{Y}|$ *denotes the cardinality of the set* $\mathcal{Y}$*.*

Proof. See Appendix A.4. □

Note that, when $\delta_i = 0$ for $i \in \{1, 2\}$, the privacy parameters are additive. However, when $\delta_i \neq 0$, the privacy parameter $\delta_2 + \delta_1|\mathcal{Y}|$ degrades linearly with an increasing number of measurement outcomes.

Remark 9 (Composability with Correlated States)**.** *In Property 3 of Theorem 1 and Proposition 3, we considered the case in which two mechanisms composed in parallel, receive independent inputs (i.e., the input being* $\rho^{X_1} \otimes \rho^{X_2}$ *where* $X_i \sim P_X \in \Theta$ *for* $i = \{1, 2\}$*, which are chosen independently). In Appendix A.5 we study the setting in*

which the inputs are correlated. There, we observe that QPP shares similar properties related to the composability of classical PP frameworks, where the class of Θ *plays a key role in composability to hold in general.*

In Proposition 3, we assume a local structure of measurements conducted in the process, as shown in Fig. 3.3. This assumption is mainly motivated by technical considerations, as we can treat the resulting setting using our existing set of tools. Exploration of advanced adaptive composition techniques, which hold for more general classes of measurements is an interesting avenue for future work. In Section 3.8.2, we present a variant of QPP where adaptive composition holds for general measurements and strategies (refer to Fig. 3.8 and Remark 23).

3.4.2 Mechanisms for QPP

We propose mechanisms to achieve ε-QPP and (ε, δ)-QPP using the depolarization channel. In addition, we provide a general procedure to generate (ε, δ)-(classical) PP mechanisms using a quantum mechanism satisfying (ε, δ)-QPP.

Depolarization Mechanism

Let

$$\mathcal{A}^p_{\mathrm{Dep}}(\rho) \coloneqq (1-p)\rho + \frac{p}{d} I, \tag{3.84}$$

where $p \in [0,1]$ and d is the dimension of the Hilbert space on which ρ acts.

Theorem 3 (ε-QPP depolarization mechanism)**.** *Fix* $p \in [0,1]$ *and a privacy framework* $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$*. Let* $\mathcal{E}$ *be a quantum channel. Then* $\mathcal{A}^p_{\mathrm{Dep}}(\mathcal{E}(\cdot))$ *(in Fig. 3.4) is* ε*-QPP*

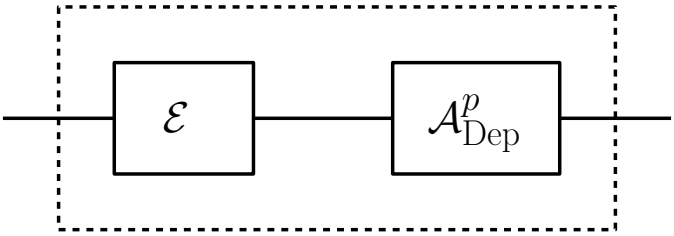


Figure 3.4: Depolarization mechanism to achieve QPP: This corresponds to a channel $\mathcal{E}$ followed by a depolarizing channel. Note that we can choose $\mathcal{E} = \mathcal{I}$ to be the identity channel as well.

if

$$p \geq \frac{dK}{dK + e^{\varepsilon} - 1}, \tag{3.85}$$

where

$$K := \sup_{M \in \mathcal{M}} \frac{\|M\|_\infty}{\mathrm{Tr}[M]} \times \sup_{\Theta, (\mathcal{R},\mathcal{T}) \in Q} \frac{\left\| \mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}}) \right\|_1}{2}. \tag{3.86}$$

This further implies that the depolarization channel with parameter p achieves ε-QPP whenever

$$\varepsilon \geq \ln\left(1 + \frac{(1-p)dK}{p}\right). \tag{3.87}$$

Proof. Fix $P_X \in \Theta$, $(\mathcal{R}, \mathcal{T}) \in Q$, and $M \in \mathcal{M}$, and consider that

$$\begin{aligned}
&\frac{\mathrm{Tr}\left[M \mathcal{A}^p_{\mathrm{Dep}}\left(\mathcal{E}(\rho^{\mathcal{R}})\right)\right]}{\mathrm{Tr}\left[M \mathcal{A}^p_{\mathrm{Dep}}(\mathcal{E}(\rho^{\mathcal{T}}))\right]} - 1 \\
&= \frac{(1-p)\mathrm{Tr}\left[M\mathcal{E}(\rho^{\mathcal{R}})\right] + \frac{p}{d}\mathrm{Tr}[M]}{(1-p)\mathrm{Tr}[M\mathcal{E}(\rho^{\mathcal{T}})] + \frac{p}{d}\mathrm{Tr}[M]} - 1 && (3.88) \\
&= \frac{(1-p)\mathrm{Tr}\left[M\left(\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})\right)\right]}{(1-p)\mathrm{Tr}[M\mathcal{E}(\rho^{\mathcal{T}})] + \frac{p}{d}\mathrm{Tr}[M]} && (3.89) \\
&\leq \frac{(1-p)\left|\mathrm{Tr}\left[M\left(\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})\right)\right]\right|}{\frac{p}{d}\mathrm{Tr}[M]} && (3.90)
\end{aligned}$$

Given the above, if

$$\varepsilon \geq \ln\left(1 + \frac{d(1-p)\left|\mathrm{Tr}\left[M\left(\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})\right)\right]\right|}{p\mathrm{Tr}[M]}\right), \tag{3.91}$$

then

$$\frac{\mathrm{Tr}\Big[M\mathcal{A}^{p}_{\mathrm{Dep}}\Big(\mathcal{E}(\rho^{\mathcal{R}})\Big)\Big]}{\mathrm{Tr}\Big[M\mathcal{A}^{p}_{\mathrm{Dep}}(\mathcal{E}(\rho^{\mathcal{T}}))\Big]} \leq e^{\varepsilon}. \tag{3.92}$$

Recalling that $(\mathcal{R},\mathcal{T}) \in \mathcal{Q}$ if and only if $(\mathcal{T},\mathcal{R}) \in \mathcal{Q}$, the roles of $\rho^{\mathcal{R}}$ and $\rho^{\mathcal{T}}$ can be interchanged, and we conclude that

$$e^{-\varepsilon} \leq \frac{\mathrm{Tr}\Big[M\mathcal{A}^{p}_{\mathrm{Dep}}\Big(\mathcal{E}(\rho^{\mathcal{R}})\Big)\Big]}{\mathrm{Tr}\Big[M\mathcal{A}^{p}_{\mathrm{Dep}}(\mathcal{E}(\rho^{\mathcal{T}}))\Big]}. \tag{3.93}$$

Consider that

$$\mathrm{Tr}\Big[M\Big(\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})\Big)\Big] \leq \|M\|_{\infty} \frac{\Big\|\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})\Big\|_1}{2}. \tag{3.94}$$

Indeed, consider the following Jordan–Hahn decomposition $\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}}) = P - Q$, where P and Q are the positive and negative parts of $\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})$, respectively, satisfying $P, Q \geq 0$ and $PQ = 0$. Then

$$\begin{aligned}
&\mathrm{Tr}\Big[M\Big(\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})\Big)\Big] \\
&= \mathrm{Tr}[M(P-Q)] && (3.95)\\
&\leq \mathrm{Tr}[MP] && (3.96)\\
&\leq \|M\|_{\infty} \frac{\Big\|\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})\Big\|_1}{2}, && (3.97)
\end{aligned}$$

where the last inequality follows from Hölder's inequality and because $\Big\|\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})\Big\|_1 = \mathrm{Tr}[P] + \mathrm{Tr}[Q] = 2\mathrm{Tr}[P]$ since $\mathrm{Tr}\Big[\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})\Big] = 0 = \mathrm{Tr}[P-Q]$.

Collecting all terms and supremizing over $\mathcal{M}$ and Θ and all secret pairs of $\mathcal{Q}$ yields the desired result. □

Note that the parameter K derived from Theorem 3 represents the domain knowledge accessible and incorporated into the privacy model of the $(\mathcal{S},\mathcal{Q},\Theta,\mathcal{M})$ QPP framework.

Corollary 2 (ε-QDP with Domain Knowledge)**.** *Fix $p \in [0,1]$, and a privacy framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ for QDP that encodes domain knowledge. Let $\mathcal{E}$ be a quantum channel. Then $\mathcal{A}^p_{\text{Dep}}(\mathcal{E}(\cdot))$ is ε-QPP with*

$$\varepsilon \geq \ln\left(1 + \frac{(1-p)d}{2p} k' \sup_{M \in \mathcal{M}} \frac{\|M\|_\infty}{\text{Tr}[M]}\right), \tag{3.98}$$

where

$$k' := \sup_{(\rho^{x_1}, \rho^{x_2}) \in \mathcal{W}_\Theta} \|\mathcal{E}(\rho^{x_1}) - \mathcal{E}(\rho^{x_2})\|_1 , \tag{3.99}$$

$$\mathcal{W}_\Theta := \{(\rho^{x_1}, \rho^{x_2}) \in \mathcal{Q} \mid \exists P_X \in \Theta \ P_X(x_1), P_X(x_2) > 0\}. \tag{3.100}$$

Note that the domain knowledge encoded into the QDP framework may guide towards an improved accuracy/utility, as opposed to considering all neighboring states as secret pairs and all possible measurements. For the QDP framework without domain knowledge, [144, Theorem 3] shows that

$$\varepsilon \geq \ln\left(1 + \frac{(1-p)d}{2p} \sup_{\rho \sim \sigma} \|\rho - \sigma\|_1\right) \tag{3.101}$$

is a sufficient condition to ensure $(\varepsilon, 0)$-QDP. Compared with that due to the condition

$$\sup_{M \in \mathcal{M}} \frac{\|M\|_\infty}{\text{Tr}[M]} \times \sup_{(\rho^{x_1}, \rho^{x_2}) \in \mathcal{W}_\Theta} \|\mathcal{E}(\rho^{x_1}) - \mathcal{E}(\rho^{x_2})\|_1 \leq \sup_{\rho \sim \sigma} \|\rho - \sigma\|_1 , \tag{3.102}$$

a QDP framework that has the capability to incorporate domain knowledge may cause less perturbation to the useful channel output of $\mathcal{E}$ in some cases. The rightmost inequality holds because $\mathcal{W}_\Theta$ includes only the neighboring pairs of states such that their occurrence has a positive probability, while $\rho \sim \sigma$ denotes all possible neighboring pairs. Furthermore, we always have that $\frac{\|M\|_\infty}{\text{Tr}[M]} \leq 1$ for every measurement operator M.

Remark 10 (Local DP)**.** *For the setup in Section 3.2.3, Theorem 3 reduces to*

$$p \geq \frac{d}{d + e^\varepsilon - 1}, \tag{3.103}$$

with the choice of the identity channel instead of $\mathcal{E}$ in Fig. 3.4. This occurs because $\|\rho - \sigma\|_1 \leq 2$, with equality for pairs of orthogonal states, and $\|M\|_\infty \leq \mathrm{Tr}[M]$, with equality holding whenever M is a rank-one measurement operator. This is analogous to a version of the randomized response technique used to achieve classical local DP [43, 44, 73]. For a finite alphabet $\mathcal{X}$ with cardinality $|\mathcal{X}|$, the randomized response mechanism outputs the true value with probability $1 - q$, and it outputs a randomly chosen realization with probability $q/|\mathcal{X}|$. Then, if

$$q \geq \frac{|\mathcal{X}|}{|\mathcal{X}| + e^{\varepsilon} - 1}, \tag{3.104}$$

ε-local differential privacy is achieved. This analogy further suggests that the depolarization mechanism can be considered as a quantum version of the randomized response mechanism that achieves classical privacy guarantees.

Considering the scenario in which we want to provide a privacy guarantee for all possible measurements (i.e., $\mathcal{M} = \bar{\mathcal{M}}$), next we derive the parameter p to achieve (ε, δ)-QPP.

Proposition 4 ((ε, δ)-QPP Depolarization Mechanism)**.** *Fix $p \in [0, 1]$ and the privacy framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ with $\mathcal{M} = \{M : 0 \leq M \leq I\}$. Let $\mathcal{E}$ be a quantum channel. Then $\mathcal{A}^p_{\mathrm{Dep}}(\mathcal{E}(\cdot))$ is ε-QPP if*

$$p \geq \min\left\{0, \frac{d(K' - \delta)}{dK' + e^{\varepsilon} - 1}\right\}, \tag{3.105}$$

where

$$K' := \sup_{\Theta, (\mathcal{R}, \mathcal{T}) \in Q} \frac{\left\|\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})\right\|_1}{2}. \tag{3.106}$$

Proof. The proof follows from the use of the equivalent formulation through the hockey-stick divergence and the properties of this divergence. By [62,

Lemma IV.I], we have

$$E_{e^\varepsilon}\Big(\mathcal{A}_{\text{Dep}}\big(\mathcal{E}(\rho^{\mathcal{R}})\big)\|\mathcal{A}_{\text{Dep}}\big(\mathcal{E}(\rho^{\mathcal{T}})\big)\Big) \leq (1-e^\varepsilon)\frac{p}{d} + (1-p)E_{e^\varepsilon}\Big(\mathcal{E}(\rho^{\mathcal{R}})\|\mathcal{E}(\rho^{\mathcal{T}})\Big). \quad (3.107)$$

We also have the property [62, Lemma II.4]

$$E_{e^\varepsilon}\Big(\mathcal{E}(\rho^{\mathcal{R}})\|\mathcal{E}(\rho^{\mathcal{T}})\Big) \leq \frac{\big\|\mathcal{E}(\rho^{\mathcal{R}}) - \mathcal{E}(\rho^{\mathcal{T}})\big\|_1}{2}. \quad (3.108)$$

Combining these relations, and supremizing over Θ, and secret pairs $(\mathcal{R},\mathcal{T}) \in \mathcal{Q}$, we can choose

$$\delta \geq (1-e^\varepsilon)\frac{p}{d} + (1-p)K'. \quad (3.109)$$

Then, rearranging the terms we arrive at

$$p \geq \frac{d(K'-\delta)}{dK' + e^\varepsilon - 1}. \quad (3.110)$$

Since $p \geq 0$, when $K' - \delta \leq 0$, we set $p = 0$. □

Classical PP Mechanisms from QPP Mechanisms

The QPP formalism provides a direct methodology to design classical PP mechanisms with the assistance of QPP mechanisms. [4] In this case, we use quantum encoding to convert classical data to quantum data. We denote the quantum encoding of classical data $x \in \mathcal{X}^{n\times k}$ as $\rho^x := |x\rangle\langle x|$ (recall that $p_X \in \Theta_c$ are discrete probability distributions over the probability space $\mathcal{P}(\mathcal{X}^{n\times k})$ and this lead to a finite collection of $\{\rho^x\}_x$ quantum encodings). Then, we ensure the privacy of the quantum data (quantum encoding) such that the privacy is ensured for the underlying classical data.

[4] In the context of classical PP mechanisms, the main attempt has been the introduction of Wasserstein mechanisms, based on the infinity order Wasserstein distance and its modifications. In particular, [121] and [91] introduced these mechanisms to achieve $(\varepsilon,0)$-PP and (ε,δ)-PP, respectively. However, it is important to note that these approaches may encounter computational intractability challenges.

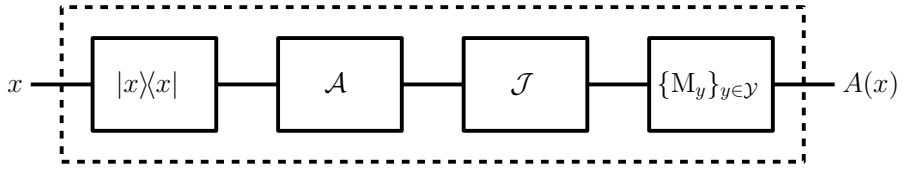


Figure 3.5: Generation of classical PP mechanisms from QPP mechanism $\mathcal{A}$: First, the classical data is encoded using quantum encoding techniques, then the QPP mechanism $\mathcal{A}$, and if needed any other channel $\mathcal{J}$, and finally the measurement channel.

Proposition 5 ((ε,δ)- (Classical) PP Mechanism)**.** *Given an (ε,δ)-QPP mechanism $\mathcal{A}$ within the framework $(\mathcal{S},\mathcal{Q},\Theta,\mathcal{M})$ when $\rho^x := |x\rangle\langle x|$ with*

$$\begin{aligned}
\mathcal{S} &= \{\{\rho^x : x \in \mathcal{R}_c\} : \mathcal{R}_c \in \mathcal{S}_c\}, \\
\mathcal{Q} &= \{(\{\rho^x : x \in \mathcal{R}_c\}, \{\rho^x : x \in \mathcal{T}_c\}) : (\mathcal{R}_c, \mathcal{T}_c) \in \mathcal{Q}_c\}\}, \\
\Theta &= \{\{p_X(x), \rho^x\}_x : p_X \in \Theta_c\}, \\
\mathcal{M} &= \{M : 0 \leq M \leq I\},
\end{aligned} \tag{3.111}$$

any post-processing of $\mathcal{A}$ by a quantum channel $\mathcal{J}$ followed by applying a POVM $\{M_y\}_{y\in\mathcal{Y}}$ denoted as $A : \mathcal{X}^{n\times k} \to \mathcal{Y}$ as shown in Fig. 3.5 is (ε,δ)-PP in the framework $(\mathcal{S}_c,\mathcal{Q}_c,\Theta_c)$.

Furthermore, for a selected post-processing $\mathcal{J}$ and POVM $\{M_y\}_{y\in\mathcal{Y}}$, it is sufficient for $\mathcal{M}_c^{\mathcal{J}} \subseteq \mathcal{M}$ for A to be (ε,δ)-PP, where

$$\mathcal{M}_c^{\mathcal{J}} := \left\{ \mathcal{J}^\dagger\left(\sum_{y\in\mathcal{B}} M_y\right) : \mathcal{B} \in \mathcal{Y} \right\}. \tag{3.112}$$

Proof. Fix $\mathcal{B} \subseteq \mathcal{Y}$. Consider that

$$\begin{aligned}
&\mathbb{P}(A(X) \in \mathcal{B} | \mathcal{R}_c) \\
&= \frac{\mathbb{P}(\{A(X) \in \mathcal{B}\} \cap \mathcal{R}_c)}{\mathbb{P}(\mathcal{R}_c)} && (3.113)\\
&\overset{(a)}{=} \frac{\sum_{x \in \mathcal{R}_c} p(x) \mathbb{P}(A(x) \in \mathcal{B})}{P_X(\mathcal{R})} && (3.114)\\
&\overset{(b)}{=} \frac{\sum_{x \in \mathcal{R}_c} p(x) \sum_{y \in \mathcal{B}} \mathbb{P}(A(x) = y)}{P_X(\mathcal{R})} && (3.115)\\
&\overset{(c)}{=} \frac{\sum_{\rho^x \in \mathcal{R}} p(x) \sum_{y \in \mathcal{B}} \mathrm{Tr}\Big[M_y \mathcal{J} \circ \mathcal{A}(\rho^x)\Big]}{P_X(\mathcal{R})} && (3.116)\\
&\overset{(d)}{=} \mathrm{Tr}\left[\sum_{y \in \mathcal{B}} M_y \mathcal{J} \circ \mathcal{A}\left(\sum_{\rho^x \in \mathcal{R}} \frac{p(x)}{P_X(\mathcal{R})} \rho^x\right)\right] && (3.117)\\
&\overset{(e)}{=} \mathrm{Tr}\left[\mathcal{J}^\dagger\left(\sum_{y \in \mathcal{B}} M_y\right) \mathcal{A}\left(\rho^{\mathcal{R}}\right)\right] && (3.118)\\
&\overset{(f)}{=} \mathrm{Tr}\Big[M \mathcal{A}(\rho^{\mathcal{R}})\Big], && (3.119)
\end{aligned}$$

where: (a) from $\mathcal{R} \coloneqq \{\rho^x : x \in \mathcal{R}_c\}$; (b) from $\mathcal{B}$ being a collection of $y \in \mathcal{Y}$; (c) from M_y being the measurement applied to obtain the outcome y; (d) from the linearity of trace operator and quantum channels $\mathcal{A}, \mathcal{J}$; (e) from the definition of $\rho^{\mathcal{R}}$, and $\mathcal{J}^\dagger$ being the adjoint of $\mathcal{J}$; and finally (f) from $M \coloneqq \mathcal{J}^\dagger\left(\sum_{y \in \mathcal{B}} M_y\right)$.

Similarly, $\mathbb{P}(A(X) \in \mathcal{B} | \mathcal{T}_c) = \mathrm{Tr}\Big[M \mathcal{A}(\rho^{\mathcal{T}})\Big]$. Then with the assumption that $\mathcal{A}$ is (ε, δ)-QPP for $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ mentioned in the proposition statement, we have

$$\mathbb{P}(A(X) \in \mathcal{B} | \mathcal{R}) \leq e^{\varepsilon}\, \mathbb{P}(A(X) \in \mathcal{B} | \mathcal{T}) + \delta. \tag{3.120}$$

concluding the proof. □

Depolarization is a common kind of noise considered in quantum information processing. Thus, this offers a method for designing classical PP mechanisms by combining the results presented in Proposition 5 and Theorem 3.

However, it is essential to recognize that quantum encoding of classical data would require additional computational resources, shifting the complexity of the mechanism design phase to the encoding phase.

3.5 Quantifying Privacy-Utility Tradeoff

In this section, we aim to assess the utility achievable through the implementation of a privatization mechanism while adhering to privacy constraints and characterize the inherent tradeoffs involved in this process. To achieve this, we define a utility metric grounded in an operational approach and demonstrate its representation via an SDP. Subsequently, we leverage this metric to conduct an in-depth analysis of privacy-utility tradeoffs, with a specific emphasis on the depolarization mechanism.

3.5.1 Utility Metric

Let $\mathcal{A}$ denote a privacy mechanism. We focus on assessing the potential of reversing the effects of $\mathcal{A}$ by applying a post-processing mechanism $\mathcal{B}$ to recover the initial input state to $\mathcal{A}$ up to an error $1-\gamma$, we define γ-utility in terms of how distinguishable $\mathcal{B}\circ\mathcal{A}$ is from the identity channel employing the normalized diamond distance as the distinguishability measure.

Definition 7 (γ-Utility)**.** *Let $\mathcal{A}:\mathcal{L}(\mathcal{H}_A)\to\mathcal{L}(\mathcal{H}_C)$ be a privacy mechanism, and fix $\gamma\in[0,1]$. We say that $\mathcal{A}$ satisfies γ-utility if*

$$\mathsf{U}(\mathcal{A}) \coloneqq 1-\inf_{\mathcal{B}}\frac{1}{2}\|\mathcal{I}-\mathcal{B}\circ\mathcal{A}\|_{\diamond}\geq\gamma, \tag{3.121}$$

where the infimum is taken over every quantum channel $\mathcal{B} : \mathcal{L}(\mathcal{H}_C) \to \mathcal{L}(\mathcal{H}_D)$.

The defined utility metric can be reformulated as an SDP by using the dual SDP form of the diamond distance [133, Section 4] and rewriting the quantum channel $\mathcal{B}$ in terms of its Choi matrix $\Gamma^{\mathcal{B}}_{CD}$, as well as translating the conditions for $\mathcal{B}$ to be a channel to conditions on its Choi matrix, namely $\Gamma^{\mathcal{B}}_{CD} \geq 0$ and $\mathrm{Tr}_D\left[\Gamma^{\mathcal{B}}_{CD}\right] = I_c$.

Proposition 6 (SDP Formulation of γ-Utility)**.** *The γ-utility of a privacy mechanism $\mathcal{A}$ can be formulated as the following SDP:*

$$\mathsf{U}(\mathcal{A}) = 1 - \inf_{\substack{\mu \geq 0 \\ Z_{AD} \geq 0 \\ \Gamma^{\mathcal{B}}_{CD} \geq 0}} \left\{ \begin{array}{c} \mu : Z_{AD} \geq \Gamma_{AD} - \Gamma^{\mathcal{B}\circ\mathcal{A}}_{AD}, \\ \mu I_A \geq \mathrm{Tr}_D[Z_{AD}], \\ \mathrm{Tr}_D\left[\Gamma^{\mathcal{B}}_{CD}\right] = I_C \end{array} \right\} \geq \gamma, \tag{3.122}$$

where

$$\Gamma^{\mathcal{B}\circ\mathcal{A}}_{AD} := \mathrm{Tr}_C\left[\left(I_A \otimes \Gamma^{\mathcal{B}}_{CD}\right)\left(\mathrm{T}_C(\Gamma^{\mathcal{A}}_{AC}) \otimes I_D\right)\right], \tag{3.123}$$

and Γ *represents the Choi matrix with the subscripts showing the input and the output system of the channel, while the superscript indicates the channel considered, with no superscript for the identity channel.*[5]

Note that a similar SDP formulation for approximate degradability, where the identity channel in Definition 7 is replaced by the complementary channel, is presented in [122, Proposition 9].

Remark 11 (Characterizing Optimal Privacy-Utility Tradeoffs)**.** *The optimal utility attained by an (ε, δ)-QPP mechanism can be characterized as an SDP. To achieve this, we combine the equivalent formulation of QPP via the DL divergence from Proposition 1 with the SDP formulation of DL divergence in Lemma 1. Combining this with*

[5]The Choi matrix of the composed channel $\mathcal{B} \circ \mathcal{A}$ is denoted by $\Gamma^{\mathcal{B}\circ\mathcal{A}}_{AD}$ and (3.123) follows from [75, Eq. (3.2.22)].

the SDP formulated from Proposition 6 enables computing the privacy requirements and quantifying utility together. Additionally, we determine the optimal privacy parameters for fixed utility requirements. For a comprehensive discussion of this point, please refer to Appendix A.6. The utilization of the SDP derived for the DL divergence in this operational task highlights an advantage of the equivalent formulation for QPP using the DL divergence.

3.5.2 Analysis of Depolarization Mechanism

We now instantiate $\mathcal{A}$ as the depolarizing channel with parameter p, denoted as $\mathcal{A}^p_{\text{Dep}}$ (as defined in (3.84)), and proceed to analyze $\mathcal{U}(\mathcal{A}^p_{\text{Dep}})$.

Proposition 7 (Utility from Depolarization Mechanism)**.** *Fix $p \in [0, 1]$. The depolarization mechanism satisfies γ-utility if and only if*

$$\mathsf{U}(\mathcal{A}^p_{\text{Dep}}) = 1 - \frac{p(d^2 - 1)}{d^2} \geq \gamma. \tag{3.124}$$

Proof. The proof below relies on observing that the optimization term (in the utility metric) is minimized by setting $\mathcal{B} = \mathcal{I}$, and then evaluating $\left\| \mathcal{I} - \mathcal{A}^p_{\text{Dep}} \right\|_\diamond$ using the Choi states of the channels $\mathcal{A}^p_{\text{Dep}}$ and $\mathcal{I}$, due to the joint covariance of the two channels under unitaries (i.e., $\mathcal{A}^p_{\text{Dep}} \circ \mathcal{U} = \mathcal{U} \circ \mathcal{A}^p_{\text{Dep}}$ and $\mathcal{I} \circ \mathcal{U} = \mathcal{U} \circ \mathcal{I}$ for every unitary channel $\mathcal{U}$).

Consider that

$$\begin{aligned}
&\left\| \mathcal{I} - \mathcal{B} \circ \mathcal{A}^p_{\text{Dep}} \right\|_\diamond \\
&\overset{(a)}{=} \left\| \mathcal{U} \circ (\mathcal{I} - \mathcal{B} \circ \mathcal{A}^p_{\text{Dep}}) \circ \mathcal{U}^\dagger \right\|_\diamond && (3.125) \\
&\overset{(b)}{=} \left\| \mathcal{I} - \mathcal{U} \circ \mathcal{B} \circ \mathcal{U}^\dagger \circ \mathcal{A}^p_{\text{Dep}}) \right\|_\diamond && (3.126) \\
&\overset{(c)}{=} \int \mathsf{d}\mathcal{U} \left\| \mathcal{I} - \mathcal{U} \circ \mathcal{B} \circ \mathcal{U}^\dagger \circ \mathcal{A}^p_{\text{Dep}}) \right\|_\diamond && (3.127) \\
&\overset{(d)}{\geq} \left\| \mathcal{I} - \left(\int \mathsf{d}\mathcal{U}\, \mathcal{U} \circ \mathcal{B} \circ \mathcal{U}^\dagger \right) \circ \mathcal{A}^p_{\text{Dep}}) \right\|_\diamond , && (3.128)
\end{aligned}$$

where: (a) follows from the unitary invariance of the diamond norm with $\mathcal{U}$ representing a unitary channel; (b) from the commutative property of $\mathcal{A}^p_{\text{Dep}}$ with every unitary channel; (c) with $\mathsf{d}\mathcal{U}$ denoting the Haar measure over the unitary group and from the left-hand side being independent of $\mathcal{U}$; and (d) from the convexity of the diamond norm.

Next, observe that $\mathcal{B}^\star := \int \mathsf{d}\mathcal{U}\, \mathcal{U} \circ \mathcal{B} \circ \mathcal{U}^\dagger$ is a quantum channel, and it is in fact equal to a depolarization channel [67]. Then, $\mathcal{B}^\star = \mathcal{A}^q_{\text{Dep}}$ for some $q \in [0, 1]$.

The composition of two depolarization channels

$$\mathcal{B}^\star = \mathcal{A}^q_{\text{Dep}} \circ \mathcal{A}^p_{\text{Dep}} \tag{3.129}$$

is also a depolarization channel with the parameter $p^\star := 1 - (1 - p)(1 - q)$. The minimum value is attained by the choice $q = 0$, where $\mathcal{A}^q_{\text{Dep}}$ in that case corresponds to the identity channel.

With that, we arrive at

$$\inf_{\mathcal{B}} \left\| \mathcal{I} - \mathcal{B} \circ \mathcal{A}^p_{\text{Dep}} \right\|_\diamond = \left\| \mathcal{I} - \mathcal{A}^p_{\text{Dep}} \right\|_\diamond . \tag{3.130}$$

With the property of joint covariance of $\mathcal{I}$ and $\mathcal{A}^p_{\text{Dep}}$ under unitaries [75, Proposition 7.82], we simplify this to

$$\left\| \mathcal{I} - \mathcal{A}^p_{\text{Dep}} \right\|_\diamond = \frac{1}{d} \left\| \Gamma_{AD} - \Gamma_{AD}^{\mathcal{A}^p_{\text{Dep}}} \right\|_1 . \tag{3.131}$$

Then consider that

$$\begin{aligned}
&\frac{1}{d}\left\|\Gamma_{AD}-\Gamma_{AD}^{\mathcal{A}^p_{\text{Dep}}}\right\|_1 \\
&\overset{(a)}{=}\frac{1}{d}\left\|\Gamma_{AD}-\left((1-p)\Gamma_{AD}+\frac{p}{d}I_{d^2}\right)\right\|_1 && (3.132)\\
&=\frac{p}{d}\left\|\Gamma_{AD}-\frac{1}{d}I_{d^2}\right\|_1 && (3.133)\\
&=p\left\|\frac{\Gamma_{AD}}{d}\left(1-\frac{1}{d^2}\right)-\frac{1}{d^2}\left(I_{d^2}-\frac{\Gamma_{AD}}{d}\right)\right\|_1 && (3.134)\\
&\overset{(b)}{=}p\left(1-\frac{1}{d^2}\right)\left(\left\|\frac{\Gamma_{AD}}{d}\right\|_1+\left\|\frac{I_{d^2}-\frac{\Gamma_{AD}}{d}}{d^2-1}\right\|_1\right) && (3.135)\\
&\overset{(c)}{=}2p\left(1-\frac{1}{d^2}\right), && (3.136)
\end{aligned}$$

where: (a) from $\Gamma_{AD}^{\mathcal{A}^p_{\text{Dep}}}=(1-p)\Gamma_{AD}+\frac{p}{d}I_{d^2}$; (b) from $\frac{\Gamma_{AD}}{d}$, and $I_{d^2}-\frac{\Gamma_{AD}}{d}$ being orthogonal; and (c) from trace norm of quantum states being equal to one.

Combining the above chain of arguments together completes the proof. □

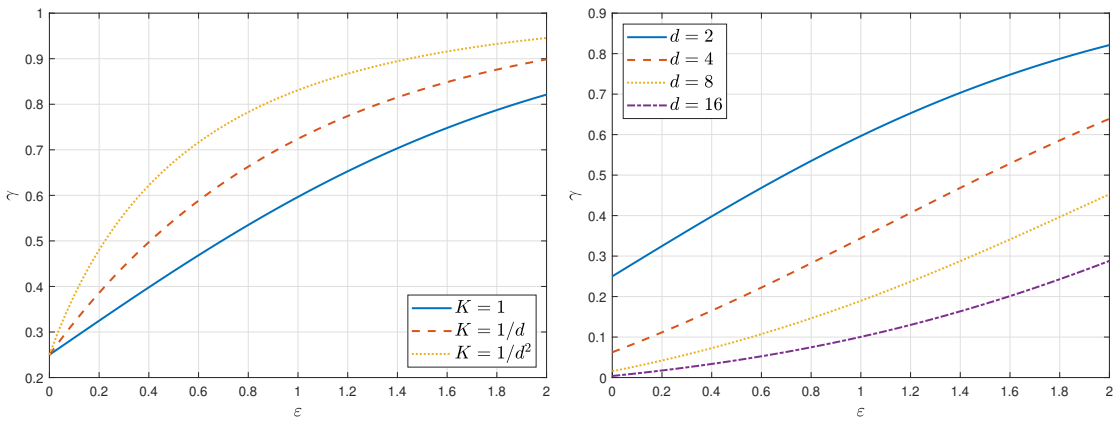
Next, we focus on understanding the privacy-utility tradeoff with respect to the parameter p governing a depolarization mechanism. From Proposition 7, to achieve γ-utility, we require that

$$p\leq\frac{(1-\gamma)d^2}{(d^2-1)}. \tag{3.137}$$

Conversely, to achieve ε-QPP in the chosen privacy framework $(\mathcal{S},\mathcal{Q},\Theta,\mathcal{M})$, from Theorem 3, we require that

$$p\geq\frac{dK}{dK+e^{\varepsilon}-1}. \tag{3.138}$$

These two inequalities provide insight into the privacy-utility tradeoff associated with the depolarization mechanism. Consequently, it is essential to carefully adjust the parameter p based on the desired utility, characterized by γ, as well as the privacy parameter ε.



(a) (b)

Figure 3.6: (a) For fixed $d = 2$, the figure depicts the optimum utility γ for ε achievable with the depolarization mechanism in Theorem 3. The value of K encodes the domain knowledge available, where $K = 1$ corresponds to no such additional information being available. (b) For fixed $K = 1$, the figure depicts the optimum utility γ for ε achievable with the depolarization mechanism in Theorem 3 for $d \in \{2, 4, 8, 16\}$.

Effect of Domain Knowledge: Fig. 3.6a illustrates the optimal utility achievable using the ε-QPP depolarization mechanism presented in Theorem 3. Notably, as the value of K reduces, the attainable utility region expands. The parameter K derived from Theorem 3 represents the domain knowledge accessible and incorporated into the privacy model of the $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ QPP framework. This observation underscores the significance of incorporating such domain knowledge to enhance utility gains while simultaneously ensuring the necessary privacy assurances.

Effect of Dimension: In Fig. 3.6b, we observe a prominent privacy-utility trade-off as the dimension increases for the depolarization mechanism presented in Theorem 3. Regarding the utility of the depolarization mechanism (given by $1 - \frac{p(d^2-1)}{d^2}$), we can always establish the following lower bound for every d:

$$1 - \frac{p(d^2 - 1)}{d^2} \geq 1 - p, \tag{3.139}$$

where this lower bound is attained as $d \to \infty$. However, the achievable privacy level ε in (3.87) degrades at most by an order of $\ln(d)$. Hence, it is crucial to identify the optimal privacy parameters achieved by private mechanisms, particularly in high-dimensional scenarios.

Remark 12 (Application Specific Privacy-Utility Tradeoffs)**.** *In the previous analysis concerning the depolarization mechanism in Fig. 3.4, we chose $\mathcal{E} = \mathcal{I}$, the identity channel. However, it would be an interesting future work to explore the utility of user-specific $\mathcal{E}$ channels. Specifically, we can choose $1 - \frac{1}{2}\inf_{\mathcal{B}\in\mathrm{CPTP}} \|\mathcal{E} - \mathcal{B}\circ\mathcal{A}^{p}_{\mathrm{Dep}}\circ\mathcal{E}\|_{\diamond}$ as the utility metric. If $\mathcal{E}$ possesses certain symmetries, one can potentially utilize arguments akin to those presented in the proof of Proposition 7. This investigation could shed light on tailoring privacy mechanisms to specific application needs, leading to more effective privacy-utility tradeoffs.*

3.6 Auditing Privacy Frameworks

Auditing for privacy aims to detect violations in privacy guarantees and reject incorrect algorithms (see [33,35,71,94] for classical approaches). In this section, our focus is on utilizing quantum information theory tools and quantum algorithms to audit the privacy of quantum systems. Specifically, we concentrate on auditing algorithms for QDP guarantees, and it should be noted that these ideas can be extended to audit algorithms for privacy guarantees demanded by QPP (see Remark 15).

The main idea behind auditing classical privacy frameworks (DP and PP) involves translating the privacy requirement into a weaker privacy notion that can be efficiently computed. For example, in [94], sliced mutual-information-based

DP is used to audit for DP by utilizing neural estimation techniques proposed in [52]. By doing so, algorithms failing to meet the privacy conditions imposed by the relaxed privacy notion are concluded to violate the original privacy requirement. However, a pitfall of this approach is the inability to quantify the gap between the constraints stemming from the original DP or PP notion and the relaxed privacy notions. In other words, even if we verify that the relaxed privacy notion is satisfied, we cannot determine whether the original privacy requirement is also satisfied. In contrast, in this chapter, we focus on auditing QDP without translating it into a relaxed privacy notion.

3.6.1 Techniques for Auditing QDP

Using Semi-Definite Programs:

By leveraging the equivalent formulation from Proposition 1 and adopting the specific choices of $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ as provided in Section 3.2.3, for (ε, δ)-QDP, we have that

$$\sup_{\rho \sim \sigma} \overline{D}^{\delta}(\mathcal{A}(\rho) \| \mathcal{A}(\sigma)) \leq \varepsilon. \tag{3.140}$$

Then, we can compute the left-hand side above by using the SDP formulation of $\overline{D}^{\delta}(\cdot\|\cdot)$ presented in Lemma 1. This approach is particularly beneficial for low-dimensional setups, as the time complexity of SDP computation is polynomial in the dimension of the quantum states. However, it is essential to note that the complexity of this approach grows exponentially with the number of qubits, making it less feasible for higher-dimensional systems.

Additionally, using the equivalent formulation in Remark 5, consider that

(ε, δ)-QDP is equivalent to

$$\sup_{\rho \sim \sigma} E_{e^{\varepsilon}}(\mathcal{A}(\rho) \| \mathcal{A}(\sigma)) \leq \delta, \tag{3.141}$$

where

$$E_{\gamma}(\rho \| \sigma) \coloneqq \mathrm{Tr}[(\rho - \gamma\sigma)_{+}] \tag{3.142}$$

$$= \frac{1}{2} \|\rho - \gamma\sigma\|_1 + \frac{(1-\gamma)}{2}, \tag{3.143}$$

with $\gamma \geq 1$ for quantum states ρ and σ [62, Eq. (II.2)]. As shown in (3.26) and (3.28), the quantity on the right-hand side of (3.142) can be evaluated by means of an SDP. Then, auditing QDP reduces to computing $E_{\gamma}(\mathcal{A}(\rho) \| \mathcal{A}(\sigma))$ for $\rho \sim \sigma$. However, similar to the previous approach, the time complexity of this SDP also grows exponentially with the number of qubits. Thus, computing these SDPs remains challenging for higher-dimensional quantum systems.

Using Quantum Circuits:

Another approach is to borrow the results of [132, 134] and use the connection of $E_{\gamma}(\rho \| \sigma)$ to the trace distance given in (3.143). Despite this connection, evaluating $E_{\gamma}(\rho \| \sigma)$ remains computationally challenging, even for quantum computers [132, 134]. Nevertheless, there are proposals for evaluating the trace distance using variational quantum algorithms [21, 115] (which, however, do not give particular runtimes), and for cases in which the quantum states have low rank [129].

In the subsequent analysis, we explore an approach from [21, 115] (via variational algorithms with parameterized quantum circuits) to estimate the quantity $\|\mathcal{A}(\rho) - e^{\varepsilon}\mathcal{A}(\sigma)\|_1$. Using such an estimate, for a fixed value of the privacy parameter ε, we can validate on which values of δ the needed guarantees are satisfied.

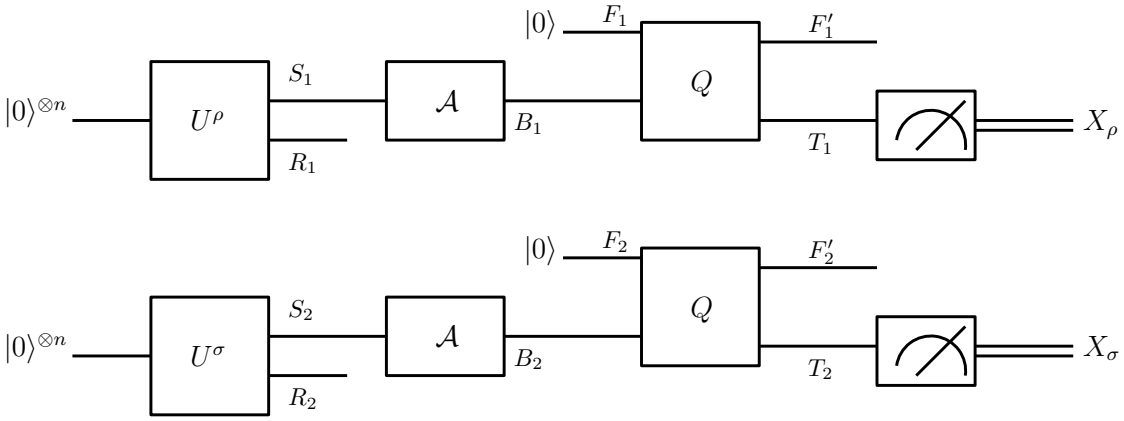


Figure 3.7: Quantum circuit assisted in estimating QDP: U^ρ, U^σ are the unitaries used to prepare ρ and σ by tracing out R_1, R_2 systems, respectively. Then $\mathcal{A}$ is applied on the systems S_i for $i \in \{1, 2\}$. The unitary Q takes inputs F_i, B_i and outputs F_i', T_i, where F_i and T_i are qubit systems. Finally, each of the T_i systems is measured and the (classical) output random variable is denoted as X_ρ for $i = 1$ and X_σ for $i = 2$. Here $X_\rho, X_\sigma \in \{0, 1\}$. This procedure is repeated a sufficient number of times, and the outcomes of the trials are used to estimate $\mathbb{P}(X_\rho = 0)$ and $\mathbb{P}(X_\sigma = 1)$.

Firstly, let us focus on how to estimate $E_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))$ for a fixed $\rho \sim \sigma$ and ε. With the ideas developed in [115, Algorithm 14], we discuss how a so-called quantum interactive proof can be used for estimating the privacy level. For that, refer to the quantum circuit in Fig. 3.7: the unitaries U^ρ and U^σ are used to prepare the states ρ and σ, by tracing out systems R_1 and R_2, respectively. Then the algorithm $\mathcal{A}$ is applied on the systems S_i for $i \in \{1, 2\}$. The unitary Q takes inputs F_i, B_i and outputs F_i', T_i, where F_i and T_i are qubit systems. Finally, both of the T_i systems are measured in the standard basis $\{|0\rangle, |1\rangle\}$, and the (classical) output random variable is denoted as X_ρ for $i = 1$ and X_σ for $i = 2$. Here X_ρ and X_σ take values in $\{0, 1\}$. This procedure is repeated a sufficient number of times, and we use the results to estimate $\mathbb{P}(X_\rho = 0)$ and $\mathbb{P}(X_\sigma = 1)$.

Next, consider a scenario in which one could maximize the following utility

function over all possible choices of Q:

$$g(Q,\rho,\sigma,\mathcal{A},\varepsilon) := \frac{1}{e^{\varepsilon}+1}\mathbb{P}(X_\rho = 0) + \frac{e^{\varepsilon}}{e^{\varepsilon}+1}\mathbb{P}(X_\sigma = 1). \tag{3.144}$$

In quantum complexity terminology, this action could be conducted by a quantum prover who has unbounded computational power (we discuss how to relax this assumption in Remark 13). From [115, Eq. (128)] and the discussion therein, we conclude that

$$f(\rho,\sigma,\mathcal{A},\varepsilon) := \sup_{Q} g(Q,\rho,\sigma,\mathcal{A},\varepsilon) \tag{3.145}$$

$$= \frac{1}{2}\left(1 + \frac{1}{e^{\varepsilon}+1}\left\|\mathcal{A}(\rho) - e^{\varepsilon}\mathcal{A}(\sigma)\right\|_1\right), \tag{3.146}$$

where the optimization is over every unitary Q.

If

$$f(\rho,\sigma,\mathcal{A},\varepsilon) \leq \frac{1}{2}\left(1 + \frac{2\delta + e^{\varepsilon} - 1}{e^{\varepsilon}+1}\right), \tag{3.147}$$

then $E_{e^{\varepsilon}}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \delta$, due to (3.142). Then, by changing the role of ρ and σ in (3.144), we obtain $f(\sigma,\rho,\mathcal{A},\varepsilon)$. Next, we select the maximum of these quantities

$$\widehat{f}(\rho,\sigma,\mathcal{A},\varepsilon) := \max\left\{f(\rho,\sigma,\mathcal{A},\varepsilon), f(\sigma,\rho,\mathcal{A},\varepsilon)\right\}. \tag{3.148}$$

To this end, if

$$\sup_{\rho\sim\sigma} \widehat{f}(\rho,\sigma,\mathcal{A},\varepsilon) \leq \frac{1}{2}\left(1 + \frac{2\delta + e^{\varepsilon} - 1}{e^{\varepsilon}+1}\right), \tag{3.149}$$

it can be verified that $\mathcal{A}$ is (ε,δ)-QDP.

Remark 13 (Relaxing the Computationally Unbounded Assumption)**.** *As the number of qubits increases, classical methods (e.g., using SDPs) often become intractable due to the exponential growth in computational complexity. So in light of this, the above approach is desired with the increasing dimension of the quantum system.*

However, finding the Q that achieves the optimum utility in (3.144) is practically infeasible. To relax this assumption, we replace the action of the prover (who finds this Q) with a parameterized circuit Q_θ. Then we can use (3.144) as a utility function for training a variational quantum algorithm [13, 19] to estimate a lower bound for (3.145). With that we obtain a lower bound on $E_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))$, which we denote as $E^{\mathrm{LB}}_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))$. A lower bound gives a sufficient condition for ruling out algorithms that do not satisfy (ε,δ)-QDP. This claim follows because the estimated lower bound $E^{\mathrm{LB}}_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) > \delta$ implies that $E_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) > \delta$.

Remark 14 (Neighboring Pairs)**.** *To verify (ε,δ)-QDP, it is required to compute whether $\widehat{f}(\rho,\sigma,\mathcal{A},\varepsilon) \leq \frac{1}{2}\left(1+\frac{2\delta+e^\varepsilon-1}{e^\varepsilon+1}\right)$ for all neighboring pairs $(\rho,\sigma)\in\mathcal{Q}$. However, checking this requirement has increasing computational complexity as the cardinality of the set $\mathcal{Q}$ increases. If the privacy requirements can be relaxed so as to encode domain knowledge as in the QPP framework, the effective set $\mathcal{Q}$ may be a small set in some applications of interest. For example, consider an application where hypothesis testing is carried out between the states ρ and σ under privacy constraints where $\mathcal{Q}=\{(\rho,\sigma),(\sigma,\rho)\}$.*

Remark 15 (Auditing QPP[6])**.** *To audit for (ε,δ)-QPP in the framework $(\mathcal{S},\mathcal{Q},\Theta,\bar{\mathcal{M}})$, one can use the ideas described above by choosing $\rho^{\mathcal{R}}$ and $\rho^{\mathcal{T}}$ for all $P_X\in\Theta,(\mathcal{R},\mathcal{T})\in\mathcal{Q}$ instead of ρ and σ. In that case, the complexity of the approach relies on the set of $\mathcal{Q}$ as well as distributions contained in Θ.*

3.6.2 Formal Guarantees for Auditing QDP

The question of quantifying the success of a privacy auditing approach, specifically in correctly accepting and rejecting an algorithm with given privacy re-

[6]Note that the following ideas can also be extended to auditing for variants of QPP in Section 3.8 as well (See also Remark 19).

quirements, is a crucial consideration in privacy auditing research. The authors of [35, 94] have worked towards answering this question for auditing classical DP, but using a relaxed privacy definition.

To tackle this for the quantum setting, we propose a hypothesis testing-based auditing pipeline tailored specifically for QDP (also for QPP). In this pipeline, we use the trace-norm estimation quantum algorithm proposed in [129]. This quantum algorithm provides an estimation with at most α additive error from the exact value, with high probability, which allows us to achieve the desired significance in the hypothesis test.

Let us define

$$T^{\varepsilon}(\rho, \sigma, \mathcal{A}) := \frac{1}{e^{\varepsilon} + 1} \left\| \mathcal{A}(\rho) - e^{\varepsilon} \mathcal{A}(\sigma) \right\|_1 . \tag{3.150}$$

Fix $(\rho, \sigma) \in \mathcal{Q}$. If algorithm $\mathcal{A}$ satisfies (ε, δ)-QDP, then $E_{e^{\varepsilon}}(\mathcal{A}(\rho) \| \mathcal{A}(\sigma)) \leq \delta$. By applying (3.142), we arrive at

$$T^{\varepsilon}(\rho, \sigma, \mathcal{A}) \leq \frac{2\delta + e^{\varepsilon} - 1}{e^{\varepsilon} + 1} =: \mathsf{g}(\varepsilon, \delta). \tag{3.151}$$

Next, by estimating the quantity on the left-hand side, and using $\mathsf{g}(\varepsilon, \delta)$ as a threshold, we design an auditing pipeline for QDP by means of the following null and alternative hypotheses:

$$H_0 : \max\{T^{\varepsilon}(\rho, \sigma, \mathcal{A}), T^{\varepsilon}(\sigma, \rho, \mathcal{A})\} \leq \mathsf{g}(\varepsilon, \delta), \tag{3.152}$$

$$H_1 : \max\{T^{\varepsilon}(\rho, \sigma, \mathcal{A}), T^{\varepsilon}(\sigma, \rho, \mathcal{A})\} > \mathsf{g}(\varepsilon, \delta). \tag{3.153}$$

Let the estimates of $T^{\varepsilon}(\rho, \sigma, \mathcal{A})$ and $T^{\varepsilon}(\sigma, \rho, \mathcal{A})$ from a randomized algorithm (in our analysis we use the algorithm corresponding to [129, Corollary 3.4]) be $\hat{T}^{\varepsilon}(\rho, \sigma, \mathcal{A})$ and $\hat{T}^{\varepsilon}(\sigma, \rho, \mathcal{A})$, respectively. We choose

$$\hat{T}^{\varepsilon}_{\max}(\rho, \sigma, \mathcal{A}) := \max\{\hat{T}^{\varepsilon}(\rho, \sigma, \mathcal{A}), \hat{T}^{\varepsilon}(\sigma, \rho, \mathcal{A})\} \tag{3.154}$$

as our test statistic.

Estimation of the Test Statistic:

Lemma 3 (Estimating trace distance using samples of ρ, σ — restatement of [129, Corollary 3.4]). *Given access to identical copies of d-dimensional quantum states ρ and σ, there is a quantum algorithm that estimates the normalized trace distance $T(\rho, \sigma)$ (recall (2.2)) within additive error α and with probability not less than $1 - \beta$, by using*

$$O\left(\log\left(\frac{1}{\beta}\right)\frac{r^2}{\alpha^5}\log^2\left(\frac{r}{\alpha}\right)\log^2\left(\frac{1}{\alpha}\right)\right) \tag{3.155}$$

samples of ρ and σ, where r is an upper bound on the rank of ρ and σ.

Lemma 3 is obtained by using the existing result in Corollary 3.4 of [129], and combining its argument in Theorem 2.6 therein on estimating $\mathrm{Tr}[A\rho]$ within additive error α with probability $1 - \beta$, by using $O\left(\frac{1}{\alpha^2}\log\left(\frac{1}{\beta}\right)\right)$ identical samples of ρ. The algorithm proposed in [129] is designed based on the following idea. Let $V := (\rho - \sigma)/2$. Consider its singular value decomposition as $V = W\Sigma U^\dagger$. Then the trace distance can be expressed by the following identity:

$$T(\rho, \sigma) = \frac{1}{2}\left(\mathrm{Tr}[\rho\,\mathrm{sgn}(V)] - \mathrm{Tr}[\sigma\,\mathrm{sgn}(V)]\right), \tag{3.156}$$

where $\mathrm{sgn}(V) := W\mathrm{sgn}(\Sigma)U^\dagger$, and $\mathrm{sgn}(\cdot)$ is the sign function. Then, $\mathrm{Tr}[\rho\,\mathrm{sgn}(V)]$ and $\mathrm{Tr}[\sigma\,\mathrm{sgn}(V)]$ can be estimated separately, by combining the techniques of quantum singular value transformation [51] and the Hadamard test [3]. To this end, to implement unitary block-encodings of ρ and σ approximately, the technique of density matrix exponentiation [84] is used.

The same techniques can be employed to compute $T^\varepsilon(\rho, \sigma, \mathcal{A})$ since

$$T^\varepsilon(\rho, \sigma, \mathcal{A}) = \frac{1}{e^\varepsilon + 1}\left(\mathrm{Tr}[\mathcal{A}(\rho)\,\mathrm{sgn}(\mathrm{V}^\varepsilon)] - e^\varepsilon\mathrm{Tr}[\mathcal{A}(\sigma)\,\mathrm{sgn}(\mathrm{V}^\varepsilon)]\right), \tag{3.157}$$

where $\mathrm{V}^{\varepsilon} := (\mathcal{A}(\rho) - e^{\varepsilon}\mathcal{A}(\sigma))/(e^{\varepsilon}+1)$.

Type-1 Error Analysis

We arrive at the following bound on the Type-1 error of the proposed hypothesis testing pipeline.

Proposition 8 (Type-I Error)**.** *Fix arbitrary $\alpha, \delta > 0$ and consider the above hypothesis testing pipeline. Then*

$$\sup_{\rho\sim\sigma} \mathbb{P}\Big(\hat{T}^{\varepsilon}_{\max}(\rho,\sigma,\mathcal{A}) > \mathsf{g}(\varepsilon,\delta) + \alpha \,\Big|\, H_0\Big) \leq \beta, \tag{3.158}$$

if the algorithm from Lemma 3 has access to

$$O\left(\log\left(\frac{1}{\beta}\right)\frac{r^2}{\alpha^5}\log^2\left(\frac{r}{\alpha}\right)\log^2\left(\frac{1}{\alpha}\right)\right) \tag{3.159}$$

identical copies of the states ρ and σ, such that $\rho \sim \sigma$ and where

$$r := \sup_{\rho\sim\sigma} \max\{\mathrm{rank}(\mathcal{A}(\rho)), \mathrm{rank}(\mathcal{A}(\sigma))\}. \tag{3.160}$$

Proof. Fix ρ and σ such that $\rho \sim \sigma$. Under the null hypothesis and the assumption that $\hat{T}^{\varepsilon}_{\max}(\rho,\sigma,\mathcal{A}) = \hat{T}^{\varepsilon}(\rho,\sigma,\mathcal{A})$, we have that

$$\begin{aligned}
&\mathbb{P}\Big(\hat{T}^{\varepsilon}_{\max}(\rho,\sigma,\mathcal{A}) > \mathsf{g}(\varepsilon,\delta) + \alpha\Big) \\
&= \mathbb{P}\Big(\hat{T}^{\varepsilon}(\rho,\sigma,\mathcal{A}) > \mathsf{g}(\varepsilon,\delta) + \alpha\Big) && (3.161)\\
&\overset{(a)}{\leq} \mathbb{P}\Big(\hat{T}^{\varepsilon}(\rho,\sigma,\mathcal{A}) - T^{\varepsilon}(\rho,\sigma,\mathcal{A}) > \alpha\Big) && (3.162)\\
&\leq \mathbb{P}\Big(|\hat{T}^{\varepsilon}(\rho,\sigma,\mathcal{A}) - T^{\varepsilon}(\rho,\sigma,\mathcal{A})| > \alpha\Big) && (3.163)\\
&\overset{(b)}{\leq} \beta, && (3.164)
\end{aligned}$$

where: (a) follows since $T^{\varepsilon}(\rho,\sigma,\mathcal{A}) \leq \mathsf{g}(\varepsilon,\delta)$ under the null hypothesis; (b) from the high probability statement in Lemma 3. Similarly, the above inequality holds when $\hat{T}^{\varepsilon}_{\max}(\rho,\sigma,\mathcal{A}) = \hat{T}^{\varepsilon}(\sigma,\rho,\mathcal{A})$ concluding the proof. □

Proposition 8 provides a bound on the number of samples of the states required to achieve type-I error (significance) of β. In that case, we would use a threshold of $g(\varepsilon, \delta) + \alpha$ for accepting the null hypothesis, such that the null hypothesis is accepted when the test statistic is less than or equal to $g(\varepsilon, \delta) + \alpha$.

Remark 16 (Computational Complexity with Rank r)**.** *From Proposition 8, it is evident that the copy complexity of the algorithm grows as $O(r^2 \log^2(r))$. Let $\mathcal{A} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ be a quantum channel. Then, $r \leq d_B$, where d_B is the dimension of the Hilbert space $\mathcal{H}_B$. To handle computational complexity, one possibility is to compose $\mathcal{A}$ with another quantum channel $\mathcal{N}$ that translates the space to a low-dimensional setting. However, due to the data-processing inequality for the trace distance, it will only provide a lower bound. With that, it is possible to reject algorithms if $T^{\varepsilon}(\rho, \sigma, \mathcal{N} \circ \mathcal{A}) > g(\varepsilon, \delta)$, which implies that $T^{\varepsilon}(\rho, \sigma, \mathcal{A}) > g(\varepsilon, \delta)$. Consequently, it may lead to limitations similar to classical auditing approaches that use relaxed privacy notions, since the contraction gap between $T^{\varepsilon}(\rho, \sigma, \mathcal{N} \circ \mathcal{A})$, and $T^{\varepsilon}(\rho, \sigma, \mathcal{A})$ is hard to quantify. In the quantification of the gap, finding the contraction coefficient $\eta_{\mathcal{N}}$ of the channel $\mathcal{N}$ would be useful if $\eta_{\mathcal{N}} < 1$ (recall that the contraction coefficient of a channel $\mathcal{N}$ under a generalized divergence $\boldsymbol{D}$, as given in Eq. (2.1), is defined as $\eta_{\mathcal{N}} := \sup_{\rho,\sigma \in \mathcal{D}, \boldsymbol{D}(\rho\|\sigma) \neq 0} \frac{\boldsymbol{D}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))}{\boldsymbol{D}(\rho\|\sigma)}$).*

In summary, we proposed a hypothesis testing pipeline for auditing the privacy of quantum systems, offering formal guarantees on auditing QDP using quantum algorithms designed for estimating trace distance. However, an essential task for further investigation is analyzing the Type-II error of this approach. This analysis would allow us to quantify the power of the test and assess its ability to correctly accept algorithms with the desired privacy requirements, which is left for future work.

3.7 Information-Theoretic Bounds from QPP

In this section, we begin by investigating several information-theoretic bounds that stem from an algorithm satisfying QPP constraints. Later on, we utilize the derived bounds to assess the relative strength of the QPP variants introduced in Section 3.8.

In [62], it was highlighted that finding bounds on quantum relative entropy and mutual information resulting from QDP is an interesting open problem. We address this question in a general setting, encompassing QDP as a special case. We offer bounds for relative entropy and Holevo information, along with bounds for Rényi relative entropies and trace distance. For the rest of the discussion, we adopt the fixed privacy framework to be $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$.

Proposition 9 (Bounds on Quantum Rényi Relative Entropy and Quantum Relative Entropy due to QPP)**.** *Fix $\alpha > 1$. If $\mathcal{A}$ is ε-QPP (i.e., $(\varepsilon, 0)$-QPP) in the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$ for all $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$, then*

$$D_\alpha\Big(\mathcal{A}(\rho^{\mathcal{R}}) \| \mathcal{A}(\rho^{\mathcal{T}})\Big) \leq \min\left\{\frac{\varepsilon^2 \alpha}{2}, \varepsilon\right\}, \tag{3.165}$$

where $D_\alpha(\cdot\|\cdot)$ is an arbitrary quantum Rényi relative entropy satisfying data processing.[7] *Furthermore,*

$$D\Big(\mathcal{A}(\rho^{\mathcal{R}}) \| \mathcal{A}(\rho^{\mathcal{T}})\Big) \leq \min\left\{\frac{\varepsilon^2}{2}, \varepsilon\right\}, \tag{3.166}$$

where D is the quantum relative entropy in (2.8).

Proof. ε-QPP of the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$ implies that for all $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$

$$D_T\Big(\mathcal{A}(\rho^{\mathcal{R}}) \| \mathcal{A}(\rho^{\mathcal{T}})\Big) \leq \varepsilon, \tag{3.167}$$

[7] Note that Petz–Rényi in (2.7) satisfies data processing for $\alpha \in (0, 1) \cup (1, 2]$ and sandwiched Rényi in (2.10) satisfies data processing for $\alpha \in [1/2, 1) \cup (1, \infty)$.

where D_T is the Thompson metric from (2.17). This follows by the definition of the max-relative entropy defined in (2.15) and Thompson metric in (2.17), and recalling the definition of QPP framework for $\delta = 0$ and $\mathcal{M} = \bar{\mathcal{M}}$ (see Definition 6). By this implication and the fact that every quantum Rényi-divergence D_α of order α satisfying data processing is bounded from above by $D_{\max}$ [126, Eq. (4.36)], ε-QPP implies that for all $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$,

$$D_\alpha\Big(\mathcal{A}(\rho^{\mathcal{R}})\|\mathcal{A}(\rho^{\mathcal{T}})\Big) \leq \varepsilon. \tag{3.168}$$

By a different argument via the maximal extension presented in Lemma 4, we can further arrive at the following, for all $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$,

$$D_\alpha\Big(\mathcal{A}(\rho^{\mathcal{R}})\|\mathcal{A}(\rho^{\mathcal{T}})\Big) \leq \frac{\varepsilon^2\alpha}{2}. \tag{3.169}$$

This completes the proof of the first inequality.

Next, noting that the Petz–Rényi relative entropy in (2.7) satisfies data processing for $\alpha \in (0,1) \cup (1,2]$, and then taking the limit $\alpha \to 1^+$, we arrive at the bound on quantum relative entropy by using the equality in (2.8). □

Lemma 4. *Fix $\alpha > 1$, and ρ, σ PSD operators. The following inequality holds:*

$$D_\alpha(\rho\|\sigma) \leq \frac{\alpha}{2}(D_T(\rho\|\sigma))^2. \tag{3.170}$$

Proof. See Appendix A.7. □

Remark 17 (Operational Interpretation of Thompson Metric)**.** *An operational interpretation of the Thompson metric has appeared in symmetric postselected hypothesis testing (a setting allowing for an inconclusive outcome along with two general conclusive outcomes and postselecting on the conclusive outcomes) as the asymptotic error exponent of discriminating two quantum states ρ and σ [114], as well as in the resource*

theory of symmetric distinguishability [117]. Here, by referring to (3.167), the QPP framework also provides another operational interpretation of the Thompson metric. In the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$*, for fixed* $P_X \in \Theta$ *and* $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$,[8] *the Thompson metric given by* $D_T\Big(\mathcal{A}(\rho^{\mathcal{R}}) \| \mathcal{A}(\rho^{\mathcal{T}})\Big)$ *is equal to the minimal* ε *needed to achieve* ε*-QPP.*

Proposition 10 (Bounds on Trace Norm)**.** *If* $\mathcal{A}$ *is* ε*-QPP, then*

$$\sup_{\Theta,(\mathcal{R},\mathcal{T})\in\mathcal{Q}} \left\|\mathcal{A}(\rho^{\mathcal{R}}) - \mathcal{A}(\rho^{\mathcal{T}})\right\|_1 \leq \min\left\{\varepsilon,\ \sqrt{2\varepsilon}\right\}, \tag{3.171}$$

and if $\mathcal{A}$ *is* (ε, δ)*-QPP, we have*

$$\sup_{\Theta,(\mathcal{R},\mathcal{T})\in\mathcal{Q}} \left\|\mathcal{A}(\rho^{\mathcal{R}}) - \mathcal{A}(\rho^{\mathcal{T}})\right\|_1 \leq 2 - \frac{4(1-\delta)}{e^{\varepsilon}+1}. \tag{3.172}$$

Proof. The first inequality holds by applying the quantum Pinsker inequality ($\frac{1}{2}\|\rho - \sigma\|_1^2 \leq D(\rho\|\sigma)$) [102, Theorem 1.15] and Proposition 9.

For the second inequality: $(0, \delta')$-QPP is equivalent to

$$\sup_{\Theta,(\mathcal{R},\mathcal{T})\in\mathcal{Q}} \frac{\left\|\mathcal{A}(\rho^{\mathcal{R}}) - \mathcal{A}(\rho^{\mathcal{T}})\right\|_1}{2} \leq \delta'. \tag{3.173}$$

Then, adapting Lemma 5 and fixing ε' therein to zero leads to the desired result.

□

Lemma 5. *Fix* $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ *privacy framework. Then, we have*

$$(\varepsilon, \delta)\text{-QPP} \implies (\varepsilon', \delta')\text{-QPP}, \tag{3.174}$$

where $\varepsilon' < \varepsilon$ *with*

$$\delta' := 1 - \frac{(e^{\varepsilon'}+1)(1-\delta)}{(e^{\varepsilon}+1)}. \tag{3.175}$$

Proof. The proof follows similarly to the proof of [27, Property 3] for classical DP and is presented in Appendix A.8. □

[8] by definition also for $(\mathcal{T}, \mathcal{R}) \in \mathcal{Q}$

3.8 Variants of Quantum Pufferfish Privacy Framework

We now propose variants of QPP via generalized divergences (as defined in (2.1)).[9] We provide an operational interpretation of generalized divergences as privacy metrics and characterize the relative strength between them.

Here, we focus on QPP in the framework $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$ and formulate QPP based on generalized divergences, which we denote by $(\mathbf{D}, \varepsilon)$-QPP, where $\mathbf{D}$ is a placeholder for the generalized divergence being used.

Definition 8 ($(\mathbf{D}, \varepsilon)$-QPP)**.** *Fix a privacy framework $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$ and $\varepsilon > 0$. An algorithm $\mathcal{A}$ is $(\mathbf{D}, \varepsilon)$-QPP if*

$$\sup_{\Theta,(\mathcal{R},\mathcal{T})\in\mathcal{Q}} \boldsymbol{D}\Big(\mathcal{A}(\rho^{\mathcal{R}})\|\mathcal{A}(\rho^{\mathcal{T}})\Big) \leq \varepsilon, \tag{3.176}$$

where $\boldsymbol{D}$ is an arbitrary generalized divergence and $\rho^{\mathcal{R}}$ and $\rho^{\mathcal{T}}$ are defined as in Definition 6. Note that these two density matrices depend on elements of Θ and $\mathcal{Q}$.

Indeed, Definition 8 encompasses variants of classical DP [15, 88] and QDP [62], which rely on Rényi divergences as the generalized divergence along with the appropriate choice of QPP framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ stated in Section 3.2.3 and Section 3.2.3, respectively.

Remark 18 (Properties of $(\boldsymbol{D}, \varepsilon)$-QPP)**.** *Post-processing holds, by the definition of generalized divergences. Convexity follows if $\boldsymbol{D}$ satisfies the direct-sum property, as defined in [75, Eq. (4.3.7)], from which it follows that $\boldsymbol{D}$ is jointly convex [75, Proposition 4.15].*

Parallel composability: If $\mathcal{A}_i$ satisfies $(\varepsilon_i, \delta_i)$-QPP in $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$ for $i \in \{1, \ldots, k\}$, then the composed mechanism, as defined in Theorem 1, satisfies $\left(\boldsymbol{D}, \sum_{i=1}^{k} \varepsilon_i\right)$-QPP in

[9] Due to the definition of generalized divergences, the privacy notions defined based on them inherently satisfy post-processing.

the framework $\left(\mathcal{S}, \mathcal{Q}^{(k)}, \Theta, \bar{\mathcal{M}}^k\right)$*, if* $\boldsymbol{D}$ *satisfies subadditivity (i.e., if* $\boldsymbol{D}(\rho_1 \otimes \rho_2 \| \sigma_1 \otimes \sigma_2) \leq \boldsymbol{D}(\rho_1 \| \sigma_1) + \boldsymbol{D}(\rho_2 \| \sigma_2)$ *for all states* ρ_1*,* ρ_2*,* σ_1*, and* σ_2*). It is worth noting that by employing this privacy notion based on generalized divergences, we achieve improved composability results, even in scenarios involving joint measurements (recall property 3 of Theorem 2).*

Remark 19 (Auditing Variants of QPP)**.** *The methodologies proposed in Section 3.6 can be used to audit the variants of QPP (based on generalized divergences) as well. In this regard, quantum algorithms and procedures for estimating respective generalized divergence (e.g., Rényi relative entropies) would be useful. This motivates the development of novel techniques for estimating them efficiently and accurately, beyond those already established in [128].*

3.8.1 Variants Based on Rényi Divergences

First, let us recall definitions of the following quantities. The measured Rényi divergence of order $\alpha \in (0,1) \cup (1,\infty)$ is defined as [48, Eqs. (3.116)–(3.117)]

$$\check{D}_\alpha(\rho \| \sigma) \coloneqq \sup_{\mathcal{M}} D_\alpha^c(\mathcal{M}(\rho) \| \mathcal{M}(\sigma)), \tag{3.177}$$

where

$$D_\alpha^c(\mathcal{M}(\rho) \| \mathcal{M}(\sigma)) \coloneqq \frac{1}{\alpha - 1} \ln\left(\sum_{x \in \mathcal{X}} (p(x))^\alpha (q(x))^{1-\alpha}\right), \tag{3.178}$$

with $p(x) \coloneqq \mathrm{Tr}[M_x \rho]$ and $q(x) \coloneqq \mathrm{Tr}[M_x \sigma]$ for $\mathcal{M}$ corresponding to a POVM $\{M_x\}_{x \in \mathcal{X}}$. The Rényi preparation divergence of order $\alpha \in (0,1) \cup (1,\infty)$ is defined as [86]

$$\hat{D}_\alpha(\rho \| \sigma) \coloneqq \inf_{P,Q,\mathcal{P}} D_\alpha^c(P \| Q), \tag{3.179}$$

where $\mathcal{P}$ is a classical–quantum channel, $\mathcal{P}(P) = \rho$, $\mathcal{P}(Q) = \sigma$, and the classical Rényi divergence is defined as

$$D_\alpha^c(P\|Q) := \frac{1}{\alpha - 1} \ln\left(\sum_{x\in\mathcal{X}} (P(x))^\alpha (Q(x))^{1-\alpha}\right). \tag{3.180}$$

The quantities in (3.177) and (3.179) satisfy the data-processing inequality for all $\alpha \in (0,1) \cup (1,\infty)$ by construction, following from this property holding for the underlying classical divergence. Moreover, the following bounds hold

$$\check{D}_\alpha(\rho\|\sigma) \leq D_\alpha(\rho\|\sigma) \leq \hat{D}_\alpha(\rho\|\sigma). \tag{3.181}$$

where D_α is an arbitrary quantum Rényi divergence that satisfies data processing [60, Eq. (3.7)].

Remark 20 (Relative Strength of Privacy Metrics)**.** *Choosing the preparation divergence $\hat{D}_\alpha = \boldsymbol{D}$ in $(\boldsymbol{D}, \varepsilon)$-QPP gives a stronger privacy metric that satisfies the post-processing property for the family of quantum Rényi divergences of order α, while $\check{D}_\alpha = \boldsymbol{D}$ gives a weaker privacy metric from that same family of divergences. That is, we have that*

$$\begin{aligned}
&\sup_{\Theta,(\mathcal{R},\mathcal{T})\in Q} \check{D}_\alpha\Big(\mathcal{A}(\rho^{\mathcal{R}})\|\mathcal{A}(\rho^{\mathcal{T}})\Big) \\
&\leq \sup_{\Theta,(\mathcal{R},\mathcal{T})\in Q} D_\alpha\Big(\mathcal{A}(\rho^{\mathcal{R}})\|\mathcal{A}(\rho^{\mathcal{T}})\Big) &(3.182)\\
&\leq \sup_{\Theta,(\mathcal{R},\mathcal{T})\in Q} \hat{D}_\alpha\Big(\mathcal{A}(\rho^{\mathcal{R}})\|\mathcal{A}(\rho^{\mathcal{T}})\Big), &(3.183)
\end{aligned}$$

so that

$$(\hat{D}_\alpha, \varepsilon)\text{-QPP} \implies (D_\alpha, \varepsilon)\text{-QPP} \implies (\check{D}_\alpha, \varepsilon)\text{-QPP}. \tag{3.184}$$

The above relations follow by the direct application of (3.181) *and Definition 8.*

Remark 21 (Operational Interpretation as Privacy Metrics)**.** *Choose*

$$\mathcal{S} = \{\rho, \sigma\}, \tag{3.185}$$

$$\mathcal{Q} = \{(\rho, \sigma)\}, \tag{3.186}$$

$$\Theta = \{\{P_X(x), \rho^x\}_{x\in\mathcal{X}} : P_X \in \mathcal{P}(\mathcal{X}),\ \rho^x \in \{\rho, \sigma\}\}, \tag{3.187}$$

$$\mathcal{M} = \bar{\mathcal{M}}. \tag{3.188}$$

The strongest privacy metric that can be generated from the family of quantum Rényi divergences of order α, which is also a generalized divergence, is devised by setting $\boldsymbol{D} = \hat{D}_\alpha$. In the chosen QPP framework $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$, the value of ε that divides the region where (D, ε)-QPP is achieved and the region where it is violated, for an identity channel, is $\hat{D}_\alpha(\rho\|\sigma)$.

The sandwiched Rényi relative entropy $\widetilde{D}$ in (2.10) satisfies data processing for $\alpha \in [1/2, 1) \cup (1, \infty)$ [46, 138], and the quantum relative entropy D in (2.8) also satisfies data processing [83]. Thus, both of these are candidates for a generalized divergence.

Proposition 11. *Fix $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$, $\alpha \in [1/2, 1) \cup (1, \infty)$, and $\delta \in (0, 1)$. Then, for an algorithm $\mathcal{A}$, we have that*

$$\varepsilon\text{-QPP} \implies (\widetilde{D}_\alpha, \varepsilon')\text{-QPP} \implies (\varepsilon^\star, \delta)\text{-QPP}, \tag{3.189}$$

where

$$\varepsilon' := \min\left\{\varepsilon, \frac{\varepsilon^2\alpha}{2}\right\}, \tag{3.190}$$

$$\varepsilon^\star := \varepsilon' + \frac{1}{\alpha - 1}\ln\left(\frac{1}{\delta^2}\right) + \ln\left(\frac{1}{1-\delta^2}\right). \tag{3.191}$$

We also have

$$\varepsilon\text{-QPP} \implies (D, \varepsilon'')\text{-QPP} \implies (\hat{\varepsilon}, \delta)\text{-QPP}, \tag{3.192}$$

where

$$\varepsilon'' := \min\left\{\varepsilon, \frac{\varepsilon^2}{2}\right\}, \tag{3.193}$$

$$K := \sup_{\Theta,(\mathcal{R},\mathcal{T})\in Q} T\Big(\mathcal{A}(\rho^{\mathcal{R}}), \mathcal{A}(\rho^{\mathcal{T}})\Big) \tag{3.194}$$

$$\hat{\varepsilon} := \frac{1}{\delta^2}(\varepsilon' + K) + \ln\left(\frac{1}{1-\delta^2}\right). \tag{3.195}$$

Proof. The first implication in both (3.189) and (3.192) follows from Proposition 9. Then for the second implication, recall from item 3 of Proposition 2 that $\overline{D}^{\delta}(\rho\|\sigma) \leq D^{\delta}_{\max}(\rho\|\sigma)$. Then applying [130, Propositions 5 and 6], we arrive at:

$$\overline{D}^{\delta}(\rho\|\sigma) \leq \widetilde{D}_{\alpha}(\rho\|\sigma) + \frac{1}{\alpha-1}\ln\left(\frac{1}{\delta^2}\right) + \ln\left(\frac{1}{1-\delta^2}\right), \tag{3.196}$$

$$\overline{D}^{\delta}(\rho\|\sigma) \leq \frac{1}{\delta^2}(D(\rho\|\sigma) + T(\rho,\sigma)) + \ln\left(\frac{1}{1-\delta^2}\right). \tag{3.197}$$

Finally, to conclude the proof, invoke Proposition 1 to establish the required relationship to the QPP framework from there. □

Note that the chain of implications in (3.189) holds for every Rényi divergence D_{α} satisfying data processing, beyond just $\widetilde{D}_{\alpha}$, because data processing is the key property to adapt Proposition 9 and [130, Proposition 6] (see Eqs. (K51) and (K52) therein).

Remark 22 (Comparison to Existing Results for QDP)**.** *In the special case of QDP, the dependence on the* (ε, α) *parameters in* (3.189) *provides a strict improvement over previous results. Specifically, Lemmas V.4 and V.5 of [62] show that*

$$\varepsilon\text{-QDP} \implies (D_{\alpha}, \varepsilon)\text{-QDP} \implies (\bar{\varepsilon}, \delta)\text{-QDP}, \tag{3.198}$$

with

$$\bar{\varepsilon} := \varepsilon + \frac{\ln\left(1/(1-\sqrt{1-\delta^2})\right)}{\alpha-1} \approx \varepsilon + \frac{\ln(2/\delta^2)}{\alpha-1}, \tag{3.199}$$

where the approximation holds for small δ. Our first implication in (3.189) is tighter compared to this since $\varepsilon' \leq \varepsilon$ and the second implication is also tighter (i.e, $\varepsilon^\star \leq \bar{\varepsilon}$) if we choose $\delta^2 \leq 1 - 2^{\left(-\frac{1}{\alpha-1}\right)}$ in the small δ regime.

3.8.2 Variant Incorporating Entanglement

In this subsection, we introduce a novel variant of QPP that incorporates reference systems. This extension potentially allows us to explore the impact of entanglement on the privacy of the system of interest.

Definition 9 (($\boldsymbol{D}^R, \varepsilon$)-QPP with Reference Systems)**.** *With the same setup $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ in Definition 6, a quantum algorithm $\mathcal{A} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ is $(\varepsilon, \boldsymbol{D})$-QPP with reference systems if*

$$\sup_{\Theta, (\omega^{\mathcal{R}}_{RA}, \omega^{\mathcal{T}}_{RA}) \in \mathcal{G}} \boldsymbol{D}\Big((\mathcal{I}_R \otimes \mathcal{A})(\omega^{\mathcal{R}}_{RA}) \| (\mathcal{I}_R \otimes \mathcal{A})(\omega^{\mathcal{T}}_{RA})\Big) \leq \varepsilon, \tag{3.200}$$

where the set $\mathcal{G}$ is defined in (3.8).

Since there is no restriction on the reference systems, the supremum in Definition 9 is also taken over the dimension of the reference system R, which is an unbounded set. However, following from isometric invariance of generalized divergences, together with purification and the Schmidt decomposition, it suffices to take the supremum over pure bipartite states with the reference system R isomorphic to the channel input system A.

Remark 23 (Properties)**.** *Similar to variants of QPP (recall Remark 18), this variant, incorporating reference systems, also satisfies post-processing, convexity, and parallel composability.*

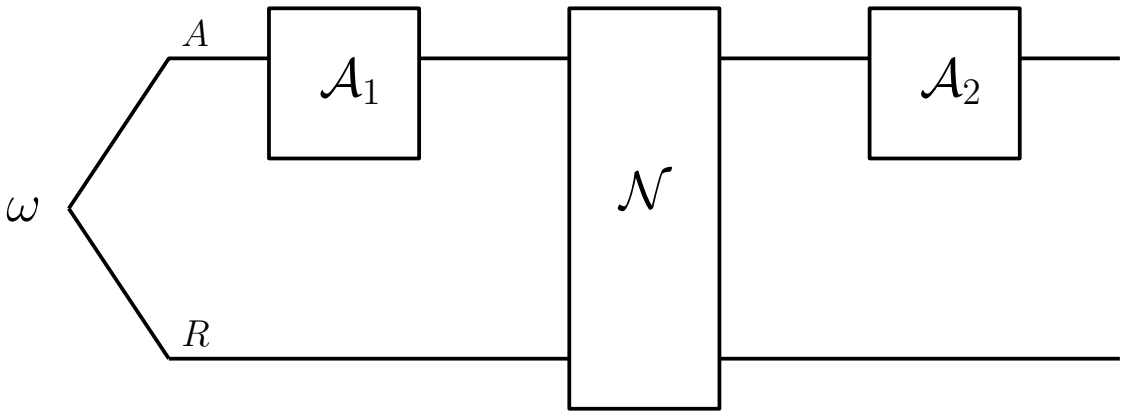


Figure 3.8: Adaptive composition with reference systems: First $\mathcal{A}_1$ is applied on the upper system, then the quantum channel $\mathcal{N}$ on both the systems, and lastly $\mathcal{A}_2$ on the upper system. Then the adaptive composition is QPP if $\mathcal{A}_1$ and $\mathcal{A}_2$ are.

Adaptive composition: Let $\mathcal{A}_i$ be a $(\boldsymbol{D}^R, \varepsilon_i)$-QPP algorithm for $i \in \{1, 2\}$. Consider the adaptive composition of $\mathcal{A}_1$ and $\mathcal{A}_2$, as shown in Fig. 3.8: First $\mathcal{A}_1$ is applied on the upper system, then the quantum channel $\mathcal{N}$ acts on both systems, and lastly $\mathcal{A}_2$ acts on the upper system. Adaptive composition of $\mathcal{A}_1$ and $\mathcal{A}_2$ also satisfies $(\boldsymbol{D}^R, \varepsilon_1)$-QPP. This shows that the adaptive composition illustrated in Fig. 3.8 does not degrade privacy, showcasing the strength of the privacy framework in Definition 9.

Next, we observe that Definition 9 is a stronger privacy guarantee than Definition 8, which does not take into account reference systems.

Corollary 3 (Strength Compared to $(\boldsymbol{D}, \varepsilon)$-QPP)**.** *$(\boldsymbol{D}^R, \varepsilon)$-QPP implies $(\boldsymbol{D}, \varepsilon)$-QPP.*

This follows from the data-processing inequality for the underlying generalized divergence $\boldsymbol{D}$, choosing the channel as the partial trace over the reference system R. It highlights that in certain scenarios where the system of interest is entangled with other reference systems, the general QPP guarantees defined in earlier sections may not be sufficient.

The choice of $\boldsymbol{D}$ could heavily affect the design of useful privacy frameworks with entanglement (recall the example in Remark 2 with the choice $\boldsymbol{D} = D_{\max}$, along with the equivalence of $D_{\max}$ to ε-QPP, as given in (3.167)). Therefore, careful consideration of the appropriate generalized divergence is essential for developing effective and meaningful privacy frameworks that account for entanglement effects.

3.9 Concluding Remarks

In this chapter, we proposed QPP as a flexible privacy framework for quantum systems. We showed that QPP is captured exactly by the DL divergence, endowing the latter with an operational interpretation of the DL divergence. The DL divergence representation was used to study properties of QPP mechanisms and characterize privacy-utility tradeoffs. As a concrete case study, we explored the depolarization QPP mechanisms and characterized the parameter values to achieve privacy. A methodology for auditing quantum privacy (the first of its kind) was also developed. In the longer term, the proposed framework could lay the foundations for privacy-preserving learning in quantum systems.

CHAPTER 4

MEASURED HOCKEY-STICK DIVERGENCE AND ITS APPLICATIONS TO PRIVACY

4.1 Introduction

In classical and quantum information theory, divergences quantify the discrepancy between two probability measures and between two quantum states, respectively. Due to the non-commutative nature of operators, a wide spectrum of quantum divergences generalizes their classical counterparts. One route for introducing a quantum divergence is to perform a measurement on the quantum states and then evaluate the supremum of the classical divergence of the resulting probability measures over all quantum measurements [36, 61, 110]. These are known as measured divergences.

However, in practice, not all measurements can be implemented due to physical restrictions or design considerations. Hence, it is important to study settings in which only limited measurements are allowed. To this end, Refs. [24, 87] analyzed the measured trace distance under restrictive measurements while Ref. [116] analyzed locally measured variants of Rényi divergences.

In this chapter, we focus on measured variants of the hockey-stick divergence under practically relevant classes of measurements. The hockey-stick divergence has recently found application as one of the fundamental quantities characterizing optimal privacy parameters in various classical and quantum privacy frameworks [5, 7, 62, 63, 95]. Quantum differential privacy (QDP) is one such privacy framework, which ensures that it is difficult for an adversary to

distinguish between two distinct quantum states under *all possible measurements* that an adversary can perform [62, 144].

Generalizing QDP, we proposed a flexible privacy framework termed quantum pufferfish privacy (QPP) in Chapter 3. One key feature of the QPP framework is the flexibility to specify the *allowed set of measurements*. It was shown in [95] that the QPP framework for the special case of *all measurements* is equivalent to a constraint on the hockey-stick divergence, by generalizing the connection established for QDP in [62].

Inspired by these connections, we study measured variants of the hockey-stick divergence. These variants can further provide insights into the study of privacy in quantum systems with QPP, taking into consideration practically motivated measurement classes. In particular, this analysis introduces avenues for quantifying optimal privacy parameters in the QPP setting and auditing privacy by computing the measured hockey-stick divergences, which were not addressed in previous works.

4.1.1 Contributions

The main goal of this chapter is to provide a comprehensive analysis of measured hockey-stick divergences under restricted measurements. We first show that, under an additive constraint on the measurement operators (satisfied by all measurement classes that allow classical post-processing), the measured hockey-stick divergence is achieved by two-outcome measurements. We then derive key properties satisfied by the measured hockey-stick divergence, including data processing, triangular inequality, and convexity.

Next, we provide computational methods for estimating the measured divergence. We show that the positive-partial-transpose (PPT) measured hockey-stick divergence can be computed by a semi-definite program (SDP). We arrive at analytical expressions for Werner and isotropic states under restricted measurements. There, we show that they coincide for different classes of measurements, including PPT measurements and measurements that use local operations and classical post-processing.

Furthermore, we show that the measured hockey-stick divergence finds operational meaning in (ε, δ)-QPP as the optimal δ that can be achieved for a fixed ε privacy parameter, extending the connection between hockey-stick divergences and QDP established in [62]. Then, we utilize the derived results to obtain optimal privacy parameters in specific privacy settings.

Lastly, we introduce the measured hockey-stick divergence for channels. To this end, we provide SDP-computable expressions for the channel divergence under PPT measurements and its connection to the divergence between the Choi states of channels for jointly-covariant channels. With that, we establish analytical expressions for the channel divergence for depolarizing channels. We also highlight an application of these channel divergences in ensuring the privacy of channels.

4.2 Measured Hockey-Stick Divergences

Let ρ and σ be states. For $\gamma \geq 0$, the hockey-stick divergence is defined as [99, 119]

$$E_\gamma(\rho\|\sigma) \coloneqq \sup_{0\leq M\leq I} \mathrm{Tr}[M(\rho - \gamma\sigma)] - (1-\gamma)_+, \tag{4.1}$$

where $(x)_+ \coloneqq \max\{0, x\}$. In this case, M and $I - M$ are valid measurement operators that satisfy $0 \leq M \leq I$ and $0 \leq I - M \leq I$. For the classical setting with (discrete) probability distributions P and Q, the hockey-stick divergence for $\gamma \geq 0$ is defined as

$$E_\gamma(P\|Q) \coloneqq \sum_x \max\{0, P(x) - \gamma Q(x)\} - (1-\gamma)_+ \tag{4.2}$$

In this chapter, we are interested in restricted measurement operators, which are inspired by the practical feasibility of only certain measurements. Let $\mathcal{M}$ denote the restricted measurement operator set, and let us define

$$\mathcal{M}_2 \coloneqq \{M : M, I - M \in \mathcal{M}\}. \tag{4.3}$$

Also, note that $\mathcal{M}_2 = \mathcal{M}$ in many cases, including all possible measurements and PPT measurements. With that, we define the measured hockey-stick divergence related to the measurement set $\mathcal{M}$.

Definition 10 (Measured Hockey-Stick Divergence)**.** *Let ρ and σ be states. For $\gamma \geq 0$, we define the measured hockey-stick divergence based on the measurement set $\mathcal{M}$ as follows:*

$$E_\gamma^{\mathcal{M}}(\rho\|\sigma) \coloneqq \sup_{M\in\mathcal{M}_2} \{\mathrm{Tr}[M(\rho - \gamma\sigma)]\} - (1-\gamma)_+, \tag{4.4}$$

where $\mathcal{M}_2$ is defined in (4.3).

By choosing $\gamma = 1$, we arrive at the measured trace distance, as used in [24, 87].

An alternative way to define the measured variant of the hockey-stick divergence is as follows: Let $\mu^{\rho,M}$ correspond to the classical probability distribution with atoms $\{\mathrm{Tr}[M_x\rho]\}_x$, and let $\mu^{\sigma,M}$ correspond to $\{\mathrm{Tr}[M_x\sigma]\}_x$ for a POVM $\{M_x\}_x$.

$$\widehat{E}_\gamma^{\mathcal{M}}(\rho\|\sigma) := \sup_{\{M_x\in\mathcal{M}\}_x} \left\{ E_\gamma\left(\mu^{\rho,M}\|\mu^{\sigma,M}\right) : M_x \geq 0,\ \sum_x M_x = I \right\}, \tag{4.5}$$

where the maximization is over $\{M_x \in \mathcal{M}\}_x$ POVMs and the classical hockey-stick divergence is defined as in (4.2).

4.3 Properties of Measured Hockey-Stick Divergence

In this section, we show that the measured hockey-stick divergence defined in Definition 10 coincides with the alternative definition in (4.5) in many settings, and it satisfies several other properties, including data processing under measurement-compatible channels, the triangular inequality, and convexity. In this chapter, we focus on the case when $\gamma \geq 1$, and note that our results extend to $\gamma \geq 0$ by utilizing (4.4) and applying the equality $E_\gamma^{\mathcal{M}}(\rho\|\sigma) = \gamma E_{1/\gamma}^{\mathcal{M}}(\sigma\|\rho)$ proved in Lemma 6.

Next, we establish conditions under which Definition 10 coincides with the alternative definition in (4.5).

Proposition 12. *Let $\gamma \geq 1$, and let ρ and σ be states. If $M_1 + M_2 \in \mathcal{M}$ holds for all $M_1, M_2 \in \mathcal{M}$ such that $M_1 + M_2 \leq I$ (coarse-graining), then the following equality holds:*

$$E_\gamma^{\mathcal{M}}(\rho\|\sigma) = \widehat{E}_\gamma^{\mathcal{M}}(\rho\|\sigma), \tag{4.6}$$

where $E_\gamma^{\mathcal{M}}$ is defined in Definition 10 and $\widehat{E}_\gamma^{\mathcal{M}}$ is defined in (4.5).

Proof. See Appendix B.2. □

Remark 24 (Measurement Sets Satisfying Equality)**.** *The coarse-graining process in Proposition 12 can be achieved by classical post-processing of the measurement outcome. Therefore, as long as classical post-processing of the measurement outcomes is allowed in any set $\mathcal{M}$, then (4.6) holds. The set of PPT measurements and the set of local measurements with classical post-processing also satisfy the coarse-graining property (see Section 4.4 for details).*

We define a channel $\mathcal{N}$ to be $\mathcal{M}$-compatible under the measurement class $\mathcal{M}$ if

$$M \in \mathcal{M} \implies \mathcal{N}^\dagger(M) \in \mathcal{M}. \tag{4.7}$$

Proposition 13 (Properties of Measured Hockey-Stick Divergence)**.** *Let $\mathcal{M}$ be the set of allowed measurement operators, and suppose that ρ and σ are states. Then $E_\gamma^{\mathcal{M}}$ in Definition 10 satisfies the following propertes:*

1. *Data Processing under $\mathcal{M}$-compatible channels: If $\mathcal{N}$ satisfies (4.7) under $\mathcal{M}$, then*

$$E_\gamma^{\mathcal{M}}(\rho\|\sigma) \geq E_\gamma^{\mathcal{M}}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))\,. \tag{4.8}$$

2. *Triangular inequality: Let $\gamma_1, \gamma_2 \geq 1$, and suppose that τ is a state. Then*

$$E_{\gamma_1\gamma_2}^{\mathcal{M}}(\rho\|\sigma) \leq E_{\gamma_1}^{\mathcal{M}}(\rho\|\tau) + \gamma_1 E_{\gamma_2}^{\mathcal{M}}(\tau\|\sigma). \tag{4.9}$$

3. *Monotonicity in γ: If $\gamma_1 \geq \gamma_2 \geq 1$, then*

$$E_{\gamma_1}^{\mathcal{M}}(\rho\|\sigma) \leq E_{\gamma_2}^{\mathcal{M}}(\rho\|\sigma). \tag{4.10}$$

4. *Convexity: Let* $\rho := \sum_{x\in\mathcal{X}} p_x\rho_x$ *and* $\sigma := \sum_{x\in\mathcal{X}} p_x\sigma_x$*, where* $p \in \mathcal{P}(\mathcal{X})$ *is a probability mass function and* $(\rho_x)_x$ *and* $(\sigma_x)_x$ *are tuples of states. Fix* $\gamma \geq 1$*. Then,*

$$E_\gamma^{\mathcal{M}}(\rho\|\sigma) \leq \sum_{x\in\mathcal{X}} p_x E_\gamma^{\mathcal{M}}(\rho_x\|\sigma_x). \tag{4.11}$$

Proof. See Appendix B.3. □

4.4 Special Classes of Measurements and States

In this section, we consider special classes of measurements, $\mathcal{M}$, and states in Definition 10, such that $E_\gamma^{\mathcal{M}}$ is computable analytically or by using a semi-definite program (SDP).

Let us define several special classes of measurements.

The set of all measurements is defined as

$$\mathcal{M}_{\text{ALL}} := \{M : 0 \leq M \leq I\}. \tag{4.12}$$

The set of local measurements (LO) is defined as

$$\mathcal{M}_{\text{LO}} := \{E_A^x \otimes F_B^y : \{E_A^x\}_x \text{ and } \{F_B^y\}_y \text{ each form a POVM}\}. \tag{4.13}$$

The set of local operations with classical post-processing is defined as

$$\mathcal{M}_{\text{LO}^\star} := \left\{ \begin{array}{c} \sum_{x,y} T(x,y) E_A^x \otimes F_B^y : \\ \{E_A^x\}_x \text{ and } \{F_B^y\}_y \text{ each form a POVM,} \\ 0 \leq T(x,y) \leq 1 \ \forall x,y \end{array} \right\}. \tag{4.14}$$

The set of local operations and one-way classical communication (1W-LOCC) based measurements is defined as

$$\mathcal{M}_{\text{1W-LOCC}} := \left\{ \sum_x E_A^x \otimes F_B^{x,y} : \{E_A^x\}_x \text{ and } \{F_B^{x,y}\}_y \ \forall\ x, \text{ form a POVM} \right\}. \tag{4.15}$$

This set can also be understood as consisting of measurement operators that correspond to applying a 1W-LOCC channel and then a local measurement, followed by classical post-processing. Similarly, the local operations and classical communications (LOCC) measurement set $\mathcal{M}_{\text{LOCC}}$ is comprised of measurement operators generated by applying an LOCC channel followed by a local measurement and classical post-processing. The set of PPT measurements is defined as

$$\mathcal{M}_{\text{PPT}} = \{M_{AB} : 0 \leq M_{AB}, \mathrm{T}_B(M_{AB}) \leq I\}. \tag{4.16}$$

By choosing $\mathcal{M} = \mathcal{M}_{\text{PPT}}$, we also have that $\mathcal{M}_{\text{PPT}} = \mathcal{M}_2$, where $\mathcal{M}_2$ is defined in (4.3). This follows because $M \in \mathcal{M}_{\text{PPT}}$ implies that $I - M \in \mathcal{M}_{\text{PPT}}$.

Proposition 12 implies that $E_\gamma^{\mathcal{M}}(\rho\|\sigma) = \widehat{E}_\gamma^{\mathcal{M}}(\rho\|\sigma)$ for $\mathcal{M} \in \{\mathcal{M}_{\text{LO}^\star}, \mathcal{M}_{\text{1W-LOCC}}, \mathcal{M}_{\text{LOCC}}, \mathcal{M}_{\text{PPT}}\}$. For the choice of PPT measurements, the measured hockey-stick divergence can be formulated as an SDP as follows.

Proposition 14 (PPT measured Hockey-Stick Divergence)**.** *Let $\gamma \geq 1$, and let $\rho, \sigma \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The quantity E_γ^{PPT} can be expressed as the following SDPs:*

$$\sup_{0 \leq M_{AB} \leq I} \{\mathrm{Tr}[M_{AB}(\rho - \gamma\sigma)] : 0 \leq \mathrm{T}_B(M_{AB}) \leq I\} =$$
$$\inf_{Y_i \geq 0} \{\mathrm{Tr}[Y_4 + Y_3] : Y_3 - Y_1 + T_B(Y_4 - Y_2) \geq \rho - \gamma\sigma\}. \tag{4.17}$$

Proof. See Appendix B.4. □

The above SDPs can be used to provide upper bounds on the locally measured hockey-stick divergence when the measurement set is comprised of LO, LO*, 1W-LOCC, and LOCC, because

$$E_\gamma^{\text{LO}}(\rho\|\sigma) \leq E_\gamma^{\text{LO}^\star}(\rho\|\sigma) \leq E_\gamma^{\text{LOCC}}(\rho\|\sigma) \leq E_\gamma^{\text{PPT}}(\rho\|\sigma) \tag{4.18}$$

for all states ρ and σ.

Next, we consider the Werner states [135], for which we characterize the measured hockey-stick divergence analytically. We define a Werner state as follows:

$$\omega^p := p\Theta + (1-p)\Theta^\perp, \tag{4.19}$$

where $p \in [0,1]$, and

$$\Theta := \frac{I+F}{d(d+1)}, \quad \Theta^\perp := \frac{I-F}{d(d-1)}. \tag{4.20}$$

with $F := \sum_{i,j=1}^{d} |i\rangle\langle j| \otimes |j\rangle\langle i|$.

Proposition 15 (Hockey-Stick Divergence for Werner States)**.** *Let $p, q \in [0,1]$. For $\gamma \geq 1$, we have that*

$$E_\gamma(\omega^q \| \omega^p) = \max\{0, q - \gamma p, (1-q) - \gamma(1-p)\}. \tag{4.21}$$

Proof. See Appendix B.5. □

Proposition 16 (Measured Hockey-Stick Divergence for Werner States)**.** *Let $p, q \in [0,1]$. The following holds for $\gamma \geq 1$ and $\mathcal{M} \in \{\mathcal{M}_{\mathrm{LO}^\star}, \mathcal{M}_{\text{1W-LOCC}}, \mathcal{M}_{\mathrm{LOCC}}, \mathcal{M}_{\mathrm{PPT}}\}$:*

$$E_\gamma^{\mathcal{M}}(\omega^q \| \omega^p) = \max\left\{0, \frac{2(q-\gamma p)}{d+1}, 1-\gamma-\frac{2(q-\gamma p)}{(d+1)}\right\}. \tag{4.22}$$

Proof. We prove the inequality "$\geq$" by specific choices of measurement operators that are LO*: $M = 0$, $M = \sum_{i=1}^{d} |i\rangle\langle i| \otimes |i\rangle\langle i|$, and $M = \sum_{i \neq j=1}^{d} |i\rangle\langle j| \otimes |i\rangle\langle j|$. Then to obtain the inequality "$\leq$", we utilize the symmetry of Werner states and the PPT measurement operators. See Appendix B.5 for a complete proof. □

Remark 25 (High Dimensions)**.** *When $d \to \infty$ in Proposition 16, we see that the measured hockey-stick divergence converges to zero, while with all measurements, it*

can still be strictly positive, as shown in Proposition 15. This highlights that Werner states are distinguishable with all measurements, while with limited measurements, that distinction decays with increasing d. We utilize this property in Proposition 18.

Remark 26 (Isotropic States)**.** *We obtain an analytical expression for the measured hockey-stick divergence of isotropic states in Appendix B.6. They also coincide for different classes of measurements, similar to what we obtained for Werner states.*

4.5 Applications to Privacy

In this section, we show how the measured hockey-stick divergence finds use in ensuring privacy for states, where privacy is imposed by quantum pufferfish privacy introduced in Chapter 3. We also utilize the tools developed in Sections 4.3 and 4.4 to study the privacy of quantum systems.

For simplicity, let us consider a special case of QPP. We call this special setting restricted quantum local differential privacy (QLDP). Note that the results below equally apply to the general QPP setting with the appropriate optimizations in Definition 6.

Definition 11 ($(\mathcal{M},\mathcal{S})$-Restricted QLDP for States)**.** *Fix $\varepsilon \geq 0$ and $\delta \in [0,1]$. Let $\mathcal{S}$ be a subset of quantum states (i.e., $\mathcal{S} \subseteq \mathcal{D}(\mathcal{H})$). Let $\mathcal{A}$ be a quantum algorithm (viz., a quantum channel). Then $\mathcal{A}$ is (ε,δ)-local differentially private under $(\mathcal{M},\mathcal{S})$ if the following condition is satisfied:*

$$\mathrm{Tr}[M\mathcal{A}(\rho)] \leq e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)] + \delta, \tag{4.23}$$

for all $\rho,\sigma \in \mathcal{S}$ and all $M \in \mathcal{M}_2$, where $\mathcal{M}_2$ is defined in (4.3) with the chosen $\mathcal{M}$.

Proposition 17 (Equivalent Representation)**.** *(ε, δ)-local differential privacy under $(\mathcal{M}, \mathcal{S})$ of a mechanism $\mathcal{A}$ (as defined in Definition 11) is equivalent to*

$$\sup_{\rho,\sigma\in\mathcal{S}} E_{e^{\varepsilon}}^{\mathcal{M}}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \delta. \tag{4.24}$$

Proof. This follows by rearranging terms in Definition 11 and recalling Definition 10 with $\gamma = e^{\varepsilon} \geq 1$. □

Remark 27 (Operational Interpretation)**.** *Proposition 17 provides an operational interpretation for the measured hockey-stick divergence $E_{e^{\varepsilon}}^{\mathcal{M}}$ as the optimal δ for fixed ε when the adversary is allowed to use measurements belonging to the set $\mathcal{M}$.*

Proposition 18 (QLDP for Werner States)**.** *Let $\mathcal{S} = \{\omega^p : 0 \leq p \leq 1\}$. Then, for $\mathcal{M} = \mathcal{M}_{\text{ALL}}$, the identity channel is not private; i.e., $\delta = 1$ for all $\varepsilon \geq 0$.*

Furthermore, the identity channel is $(\varepsilon, 2/(d+1))$- QLDP under $(\mathcal{M}, \mathcal{S})$ for $\varepsilon \geq 0$ and for $\mathcal{M} \in \{\mathcal{M}_{\text{LO}^\star}, \mathcal{M}_{\text{1W-LOCC}}, \mathcal{M}_{\text{LOCC}}, \mathcal{M}_{\text{PPT}}\}$.

Proof. The proof follows by choosing $\mathcal{A} = \mathcal{I}$ in Proposition 17, using the convexity of hockey-stick divergence in Proposition 13, and applying Propositions 15 and 16. □

Proposition 18 shows explicitly how the measurement class plays an inherent role in ensuring privacy, if we are aware of the measurements that the adversary can perform. In particular, an algorithm will be completely non-private with a certain class of measurements (all measurements in Proposition 18) and almost completely private for another class of measurements (LOCC measurements in Proposition 18 with increasing dimension).

Remark 28 (Privacy Auditing)**.** *Auditing privacy refers to identifying whether an algorithm is private, as demanded by Definition 11. In many practical scenarios, the*

set $\mathcal{S}$ relevant to an $(\mathcal{M}, \mathcal{S})$-restricted QLDP setting contains a finite number of elements, which allows for a computationally feasible method to audit privacy. For a channel $\mathcal{A}$, a measurement set $\mathcal{M} \subseteq \mathcal{M}_{\mathrm{PPT}}$, and a set of states $\mathcal{S}$ with finite number of elements, one can compute $E_{e^\varepsilon}^{\mathrm{PPT}}(\mathcal{A}(\rho), \mathcal{A}(\sigma))$ for each pair $(\rho, \sigma) \in \mathcal{S} \times \mathcal{S}$ via an SDP. If $\max_{\rho,\sigma\in\mathcal{S}} E_{e^\varepsilon}^{\mathrm{PPT}}(\mathcal{A}(\rho), \mathcal{A}(\sigma)) \leq \delta$, then the mechanism $\mathcal{A}$ is guaranteed to be (ε, δ)-restricted QLDP under $(\mathcal{M}, \mathcal{S})$.

Although the method described above provides a computationally feasible way to verify if a mechanism is private, the time complexity of this algorithm is exponential in the number of qubits and polynomial in the cardinality of $\mathcal{S}$. This calls for more efficient algorithms to estimate the measured hockey-stick divergence and its optimization over a given set of states, which we leave for future investigation.

4.6 Measured Channel Divergences

In this section, we extend the study of measured hockey-stick divergence to quantum channels.

Definition 12 (Hockey-Stick Divergence for Channels)**.** *Let $\mathcal{P}_{A\to B}$, $\mathcal{Q}_{A\to B}$ be channels, and let $\mathcal{M}$ be the allowed set of measurements on systems R and B. Then for $\gamma \geq 0$, the channel divergence for the measured hockey-stick divergence is defined as*

$$E_\gamma^{\mathcal{M}}(\mathcal{P}\|\mathcal{Q}) := \sup_{\rho_{RA}} E_\gamma^{\mathcal{M}}(\mathcal{P}(\rho_{RA})\|\mathcal{Q}(\rho_{RA}))\,, \tag{4.25}$$

where the optimization is over all states with an unbounded reference system R.

Using the Schmidt decomposition theorem and the convexity of measured hockey-stick divergence in Proposition 13, the optimization in (4.25) can be restricted to pure states with the R system isomorphic to the A system.

Next, we show how the channel divergence of the measured hockey-stick divergence appears in applications related to ensuring privacy for quantum channels.

Privacy for Channels: Previously, *privacy of states* has been considered under the QDP and QPP frameworks [62, 95, 144]. However, there are also scenarios where one needs to privatize the channel used in the quantum operation. These will find use in quantum reading [111], where encoding different values corresponds to applying different channels. In this setting, we require the channels to remain indistinguishable to the adversary. To accomplish that, we define a privacy framework that ensures *privacy for channels*. This is accomplished via superchannels, which are linear maps that transform one channel to another channel [26, 53].

Definition 13 ($(\mathcal{M}, C)$-QLDP for Channels)**.** *Fix $\varepsilon \geq 0$ and $\delta \in [0, 1]$. Let Θ be a superchannel, and let C be a subset of channels from $\mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$. The algorithm/supermechanism Θ is (ε, δ)-local differentially private under $(\mathcal{M}, C)$ if for all $\mathcal{P}_{A\to B}, \mathcal{Q}_{A\to B} \in C$ the following condition is satisfied for all $\rho_{RA} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_A)$ and for all $M \in \mathcal{M}_2$:*

$$\mathrm{Tr}[M\,(\Theta(\mathcal{P})(\rho_{RA}))] \leq e^{\varepsilon}\mathrm{Tr}[M\,(\Theta(\mathcal{Q})(\rho_{RA}))] + \delta, \tag{4.26}$$

where $\mathcal{M}_2$ is defined in (4.3) with the chosen $\mathcal{M}$.

Similar to Proposition 17, we next show that there is an equivalent representation for Definition 13 using the channel hockey-stick divergence.

Proposition 19 (Equivalent Representation)**.** *(ε, δ)-local differential privacy under $(\mathcal{M}, C)$ of a supermechanism Θ (as defined in Definition 13) is equivalent to*

$$\sup_{\mathcal{P},\mathcal{Q}\in C} E^{\mathcal{M}}_{e^{\varepsilon}}(\Theta(\mathcal{P})\|\Theta(\mathcal{Q})) \leq \delta. \tag{4.27}$$

Proof. This follows by rearranging terms in Definition 13, recalling (4.25), and applying Definition 10 for $\gamma = e^{\varepsilon} \geq 1$. □

Notably, Proposition 19 shows that the channel hockey-stick divergence is the fundamental quantity in the task of guaranteeing privacy of quantum channels, thus providing an operational interpretation for it.

Special Measurements and Channels

Here, we focus on the channel divergence by considering special classes of measurements, including PPT measurements and special classes of channels, namely, jointly-covariant channels.

For the class of all measurements, the channel divergence is SDP computable. For depolarizing channels, we derive analytical expressions for them. See Appendix B.7 for a detailed study. We next show that the channel divergence under PPT measurements is also SDP computable.

Proposition 20 (Channel Divergence under PPT Measurements)**.** *Let $\mathcal{P}_{A\to B}$ and $\mathcal{Q}_{A\to B}$ be two quantum channels, and let $\gamma \geq 1$. Then $E_\gamma^{\mathrm{PPT}}(\mathcal{P}\|\mathcal{Q})$ is equal to*

$$\sup_{\rho_R,\Omega_{RB}\geq 0} \left\{ \begin{array}{c} \mathrm{Tr}\left[\Omega_{RB}(\Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB})\right] : \\ \mathrm{Tr}[\rho_R] = 1,\ \Omega_{RB} \leq \rho_R \otimes I_B \\ \mathrm{T}_B[\Omega_{RB}] \geq 0, \mathrm{T}_B[\Omega_{RB}] \leq \rho_R \otimes I_B \end{array} \right\} = \tag{4.28}$$

$$\inf_{\substack{\mu\geq 0, Z_{RB}\geq 0 \\ L_{RB}, Y_{RB}\geq 0}} \left\{ \mu : \begin{array}{c} \Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB} \leq Z_{RB} + \mathrm{T}_B(Y_{RB} - L_{RB}) \\ \mu I_{RB} \geq \mathrm{Tr}_B[Z_{RB}] + \mathrm{Tr}_B[Y_{RB}] \end{array} \right\}, \tag{4.29}$$

where $\Gamma^{\mathcal{N}}_{RB}$ is the Choi operator of the channel $\mathcal{N}_{A\to B}$.

Proof. See Appendix B.8. □

Let G be a finite group, and for every $g \in G$, let $g \to U_A(g)$ and $g \to V_B(G)$ be unitary representations acting on the input and output spaces of the channel, respectively. A quantum channel $\mathcal{N}_{A\to B}$ is covariant with respect to these representations if the following equality holds for every input density operator $\rho_A \in \mathcal{L}(\mathcal{H}_A)$ and group element $g \in G$: $\mathcal{N}_{A\to B}\left(U_A(g)\rho_A U_A^\dagger(g)\right) = V_B(g)\mathcal{N}_{A\to B}(\rho_A)V_B^\dagger(g)$.

For covariant quantum channels, the input states can be restricted to a specific form in the optimization in (4.25) and can further be characterized by the Choi states of the channels if the representation $\{U_A(g)\}_{g\in G}$ is irreducible.

Proposition 21. *Let G be a finite group with $\{U_A(g)\}_{g\in G}$ and $\{V_B(g)\}_{g\in G}$ unitary representations. Let $\mathcal{P}_{A\to B}$ and $\mathcal{Q}_{A\to B}$ be quantum channels that are covariant with respect to the group G. Then $E_\gamma^{\mathrm{PPT}}(\mathcal{P}\|\mathcal{Q})$ is equal to*

$$\sup_{\substack{\psi_{RA}:\\ \psi_A=\mathcal{T}_G(\psi_A)}} \left\{E_\gamma^{\mathrm{PPT}}(\mathcal{P}_{A\to B}(\psi_{RA})\|\mathcal{Q}_{A\to B}(\psi_{RA}))\right\}, \tag{4.30}$$

where system R is isomorphic to system A, $\psi_A := \mathrm{Tr}_R[\psi_{RA}]$, and $\mathcal{T}_G(\omega_A) := \frac{1}{|G|}\sum_{g\in G} U_A(g)\omega_A U_A^\dagger(g)$.

Furthermore, if the representation $\{U_A(g)\}_{g\in G}$ is irreducible, then an optimal state in (4.30) is the maximally entangled state Φ_{RA}, so that $E_\gamma^{\mathrm{PPT}}(\mathcal{P}\|\mathcal{Q}) = E_\gamma^{\mathrm{PPT}}(\mathcal{P}_{A\to B}(\Phi_{RA})\|\mathcal{Q}_{A\to B}(\Phi_{RA}))$.

Proof. See Appendix B.10. The proof relies on the data-processing of E_γ^{PPT} under PPT-preserving channels and similar techniques used in the proof of [75, Proposition 7.84]. □

Next, by using Proposition 21 and Proposition 38, we derive an analytical expression for E_γ^{PPT} for two depolarizing channels.

Proposition 22 (Depolarizing Channels)**.** *Let $p, q \in [0,1]$ and $\gamma \geq 1$. Then*

$$E_\gamma^{\mathrm{PPT}}\left(\mathcal{A}_{\mathrm{Dep}}^q \middle\| \mathcal{A}_{\mathrm{Dep}}^p\right) = \max\left\{0, 1-q-\gamma(1-p)+\frac{(q-\gamma p)}{d}, \frac{d-1}{d}(q-\gamma p)\right\}, \quad (4.31)$$

where $\mathcal{A}_{\mathrm{Dep}}^p$ is a depolarizing channel with parameter p and d is the dimension of the input space of the channel $\mathcal{A}_{\mathrm{Dep}}^p$.

Proof. The proof follows by applying Proposition 21 together with the analytical expressions for E_γ^{PPT} of isotropic states (Remark 26). See Appendix B.11. □

4.7 Concluding Remarks

In this chapter, we provided a comprehensive study of measured hockey-stick divergences for states and channels. We showed that they satisfy several desirable properties, including data processing, triangular inequality, and convexity. We also showed that, for some classes of measurements and states, they are SDP computable or analytically characterized. We discussed their use in providing privacy for states and channels, with an emphasis on quantum pufferfish privacy. To this end, we introduce a privacy framework for quantum channels, extending the QPP framework for states.

CHAPTER 5

CONTRACTION OF PRIVATE QUANTUM CHANNELS

5.1 Introduction

Statistical privacy models aim to tackle privacy risks in a principled manner [27, 40, 41, 77, 94]. Local differential privacy (LDP) is one such model (Definition 14), where individual data is protected while answering aggregate queries [43]. Studying statistical problems under local privacy constraints is vital for understanding the price that we have to pay to ensure privacy. To this end, the contraction of statistical measures and divergences under privacy constraints is an important technical tool. One such tool is the contraction coefficient of the total-variation distance $\mathrm{TV}(p,q) \coloneqq \frac{1}{2}\|p-q\|_1$ for two probability distributions p and q: In [72], it was shown that, under ε-LDP privacy constraints,

$$\sup_{\substack{M\in\mathcal{B}_c^\varepsilon,\\ \mathrm{TV}(p,q)\neq 0}} \frac{\mathrm{TV}(M(p),M(q))}{\mathrm{TV}(p,q)} = \frac{e^\varepsilon-1}{e^\varepsilon+1} = \tanh(\varepsilon/2)\,, \tag{5.1}$$

where $\mathcal{B}_c^\varepsilon$ denotes all ε-LDP mechanisms and $M(p)$ and $M(q)$ represent the probability distributions resulting from processing p and q by the private mechanism M, respectively. Bounds for the contraction coefficients of other divergences, including chi-square, Hellinger, and Kullback–Leibler divergence, which are obtained by replacing total-variation distance in (5.1) by the respective divergence, have been studied in [8, 38, 39].

With the development of quantum technologies and the generation of quantum data, ensuring the privacy of quantum systems is an important research direction. To this end, statistical privacy frameworks for quantum data have

been developed recently, which are generalizations of classical statistical privacy frameworks [1,62,95,144] as explained extensively in Chapter 3. Quantum local differential privacy (QLDP), which is a generalization of LDP, ensures that two distinct quantum states passed through a quantum channel are hard to distinguish by a measurement [62].

Quantifying the contraction of quantum divergences under quantum privacy constraints enables the study of statistical tasks under privacy constraints (QLDP constraints). Earlier efforts in this direction include the following: entropic inequalities under QLDP constraints on algorithms and measurements were discussed in [6], while upper bounds on contraction of divergences under QLDP algorithms were studied in [62,63]. However, finding an exact characterization of contraction coefficients of quantum divergences under privacy constraints is still largely unexplored.

5.1.1 Contributions

In this chapter, we study the contraction of quantum divergences under privacy constraints imposed by QLDP.

First, we derive an upper bound on the privatized contraction coefficient of the hockey-stick divergence (Theorem 4, Corollary 4). As a result, for the special case of $\gamma = 1$, we find an upper bound on the privatized contraction coefficient of the normalized trace distance. Using this to establish the converse and mechanism that satisfies QLDP (Proposition 23) for achievability, we show that the privatized contraction coefficient for the normalized trace distance under ε-QLDP mechanisms is precisely equal to $(e^{\varepsilon} - 1)/(e^{\varepsilon} + 1)$. Moreover, using

the above results, we provide upper bounds on the contraction of the Bures distance and quantum relative entropy relative to normalized trace distance in Proposition 24 and Proposition 25, respectively.

5.2 Quantum Local Differential Privacy

We review privacy frameworks for both the classical and quantum settings, focusing especially on the local setting in which the privacy of individual entries and quantum states is ensured, respectively. Local differential privacy (LDP) allows for answering queries about aggregate quantities while protecting the individual entries without the need of a central trustworthy data curator [43]. We note that a randomized privacy mechanism A, as mentioned below, is described by a (regular) conditional probability distribution $P_{A|X}$ for its output given the data.

Definition 14 (Classical Local Differential Privacy)**.** *Fix $\varepsilon \geq 0$ and $\delta \in [0,1]$. A randomized mechanism $A : \mathcal{X} \to \mathcal{Y}$ is (ε, δ)-local differentially private if*

$$\mathbb{P}(A(x) \in \mathcal{B}) \leq e^{\varepsilon}\, \mathbb{P}(A(x') \in \mathcal{B}) + \delta, \tag{5.2}$$

for all $x, x' \in \mathcal{X}$ and $\mathcal{B} \subseteq \mathcal{Y}$ measurable. We say that A satisfies ε-LDP if it satisfies $(\varepsilon, 0)$-LDP.

Quantum local differential privacy (QLDP) ensures the privacy of quantum states, which are given as inputs to a private quantum channel [62, 95].

Definition 15 (Quantum Local Differential Privacy)**.** *Fix $\varepsilon \geq 0$ and $\delta \in [0,1]$. Let $\mathcal{A}$ be a quantum algorithm (viz., a quantum channel). The algorithm $\mathcal{A}$ is (ε, δ)-local*

differentially private if

$$\mathrm{Tr}[M\mathcal{A}(\rho)] \leq e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)] + \delta, \qquad \forall \rho, \sigma \in \mathcal{D}(\mathcal{H}), \quad \forall M : 0 \leq M \leq I. \tag{5.3}$$

We say that $\mathcal{A}$ satisfies ε-QLDP if it satisfies $(\varepsilon, 0)$-QLDP.

Remark 29 (Connection to Hockey-stick Divergence)**.** *(ε, δ)-QLDP of a mechanism $\mathcal{A}$ is equivalent to the following condition:*

$$\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} E_{e^{\varepsilon}}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \delta, \tag{5.4}$$

where $E_{\gamma}(\cdot\|\cdot)$ is defined in (2.3) for $\gamma \geq 1$. This was shown for (ε, δ)-QLDP in [62, Eq. (V.1)]. We provide a brief proof here for convenience: For all states ρ and σ and for every measurement operator M (i.e., satisfying $0 \leq M \leq I$), the ε-QLDP criterion in (5.3) can be arranged as $\mathrm{Tr}[M\,(\mathcal{A}(\rho) - e^{\varepsilon}\mathcal{A}(\sigma))] \leq \delta$. Then, we have

$$\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} E_{e^{\varepsilon}}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = \sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} \mathrm{Tr}[(\mathcal{A}(\rho) - e^{\varepsilon}\mathcal{A}(\sigma))_{+}] \tag{5.5}$$

$$= \sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} \sup_{0\leq M\leq I} \mathrm{Tr}[M\,(\mathcal{A}(\rho) - e^{\varepsilon}\mathcal{A}(\sigma))] \tag{5.6}$$

$$\leq \delta, \tag{5.7}$$

concluding the claim. For the case in which $\delta = 0$, ε-QLDP is equivalent to

$$\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} E_{e^{\varepsilon}}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = 0, \tag{5.8}$$

since $\mathrm{Tr}[(\mathcal{A}(\rho) - e^{\varepsilon}\mathcal{A}(\sigma))_{+}] \geq 0$ in (5.5).

Furthermore, it is sufficient to consider the supremum over orthogonal pure states in (5.4). That is, (ε, δ)-QLDP is equivalent to

$$\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} E_{e^{\varepsilon}}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = \sup_{\varphi_1\perp\varphi_2} E_{e^{\varepsilon}}(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)) \leq \delta, \tag{5.9}$$

where φ_1 and φ_2 are orthogonal pure states. This is implied by the proof of [62, Theorem II.2]. We provide an alternative proof for this fact in Appendix C.1.

Remark 30 (Connection to Max-Relative Entropy and Datta–Leditzky Divergence)**.** *ε-QLDP is also equivalent to the following constraint:*

$$\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} D_{\max}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \varepsilon, \tag{5.10}$$

as shown in (3.167) *(see also [62, Lemma III.2] with $\delta = 0$), where the max-relative entropy is defined as [28]*

$$D_{\max}(\rho\|\sigma) := \ln\inf\{\lambda : \rho \leq \lambda\sigma\} \tag{5.11}$$

$$= \ln \sup_{0\leq M\leq I} \frac{\mathrm{Tr}[M\rho]}{\mathrm{Tr}[M\sigma]}. \tag{5.12}$$

The inequality in (5.10) *can also be inferred from the dual formulation of max-relative entropy in* (2.15)*, together with Definition 15.*

By adapting Proposition 1 in Chapter 3, (ε,δ)-QLDP is also equivalent to the following condition:

$$\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} \overline{D}^{\delta}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \varepsilon, \tag{5.13}$$

where

$$\overline{D}^{\delta}(\rho\|\sigma) := \ln\inf\{\lambda \geq 0 : \mathrm{Tr}[(\rho-\lambda\sigma)_+] \leq \varepsilon\}. \tag{5.14}$$

Furthermore, it is sufficient to consider the supremum over orthogonal pure states in (5.4)*. That is, (ε,δ)-QLDP is equivalent to*

$$\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} \overline{D}^{\delta}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = \sup_{\varphi_1\perp\varphi_2} \overline{D}^{\delta}(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)) \leq \varepsilon, \tag{5.15}$$

where φ_1 and φ_2 are orthogonal pure states. The proof of this fact is in Appendix C.2. For $\delta = 0$, this can also be deduced from [54, Theorem 5.7].

5.3 Contraction under Quantum Privacy Constraints

In this section, we first propose a quantum mechanism that satisfies ε-QLDP, and we then proceed to evaluate the contraction of generalized divergences, with an emphasis on the hockey-stick divergence and normalized trace distance under privacy constraints, as imposed by QLDP.

Let $\mathcal{A}^p_{\text{Dep}}$ be the depolarizing channel

$$\mathcal{A}^p_{\text{Dep}}(\rho) \coloneqq (1-p)\rho + \frac{p}{d}I, \tag{5.16}$$

where $p \in [0,1]$ and d is the dimension of the Hilbert space on which ρ acts. A depolarization mechanism achieving standard quantum differential privacy with a neighboring relation declared by the closeness of normalized trace distance was presented in [62, 144]. Considering the worst case scenario where the normalized trace distance constraint evaluates to one, a mechanism achieving ε-QLDP based on a depolarizing channel can be obtained, as shown in Remark 10 in Chapter 3. However, the parameters therein depend on d.

In [5, Lemma 5.2], a general recipe to derive privacy mechanisms by applying a POVM followed by a classical privacy mechanism has been studied. The next proposition derives a QLDP mechanism using a similar technique, where a two-outcome POVM is followed by the application of a depolarizing mechanism. In this mechanism, the parameters are independent of the dimension d of the input states.

Proposition 23 (QLDP Mechanism)**.** *For all $\varepsilon \geq 0$ and $p \in [0,1]$, the mechanism $\mathcal{A}^p_{\text{Dep}} \circ \mathcal{M}$ satisfies ε-QLDP if*

$$p \geq \frac{2}{e^{\varepsilon}+1}, \tag{5.17}$$

where, for a fixed measurement operator M (satisfying $0 \leq M \leq I$) and for an input state ω, the measurement channel $\mathcal{M}$ is defined as

$$\mathcal{M}(\omega) := \text{Tr}[M\omega]|0\rangle\langle 0| + \text{Tr}[(I - M)\omega]|1\rangle\langle 1|. \tag{5.18}$$

Proof. Fix ρ, σ, $d = 2$, and M such that $0 \leq M \leq I$ and consider that

$$\begin{aligned}
&\frac{\text{Tr}\left[M\mathcal{A}^p_{\text{Dep}}(\mathcal{M}(\rho))\right]}{\text{Tr}\left[M\mathcal{A}^p_{\text{Dep}}(\mathcal{M}(\sigma))\right]} - 1 \\
&= \frac{(1-p)\text{Tr}[M\mathcal{M}(\rho)] + \frac{p}{d}\text{Tr}[M]}{(1-p)\text{Tr}[M\mathcal{M}(\sigma)] + \frac{p}{d}\text{Tr}[M]} - 1 && (5.19) \\
&= \frac{(1-p)\text{Tr}[M(\mathcal{M}(\rho) - \mathcal{M}(\sigma))]}{(1-p)\text{Tr}[M\mathcal{M}(\sigma)] + \frac{p}{d}\text{Tr}[M]} && (5.20) \\
&\leq \frac{(1-p)\left|\text{Tr}[M(\mathcal{M}(\rho) - \mathcal{M}(\sigma))]\right|}{\frac{p}{d}\text{Tr}[M]}. && (5.21)
\end{aligned}$$

Consider that

$$\begin{aligned}
\text{Tr}[M(\mathcal{M}(\rho) - \mathcal{M}(\sigma))] &\leq \|M\|_1 \frac{\|\mathcal{M}(\rho) - \mathcal{M}(\sigma)\|_1}{2} && (5.22) \\
&\leq \text{Tr}[M], && (5.23)
\end{aligned}$$

where the last inequality follows because $\|M\|_1 = \text{Tr}[M]$, given that $M \geq 0$, and the normalized trace distance of quantum states is bounded from above by one.

Given the above, together with (5.21), if

$$\varepsilon \geq \ln\left(1 + \frac{d(1-p)}{p}\right), \tag{5.24}$$

then

$$\frac{\text{Tr}\left[M\mathcal{A}^p_{\text{Dep}}(\mathcal{M}(\rho))\right]}{\text{Tr}\left[M\mathcal{A}^p_{\text{Dep}}(\mathcal{M}(\sigma))\right]} \leq e^{\varepsilon}. \tag{5.25}$$

By recalling that $d = 2$, and rearranging terms in (5.24), we arrive at (5.17). □

5.3.1 Contraction Coefficients under Privacy Constraints

Let ρ and σ be quantum states. A generalized divergence, by its definition in (2.1), satisfies data processing under every channel $\mathcal{N}$; i.e.,

$$\boldsymbol{D}(\rho\|\sigma) \geq \boldsymbol{D}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))\,. \tag{5.26}$$

For some cases, this contraction can be characterized more tightly if there exists $\eta_{\boldsymbol{D}} \in (0,1)$ such that the following inequality holds for all ρ and σ:

$$\eta_{\boldsymbol{D}}\, \boldsymbol{D}(\rho\|\sigma) \geq \boldsymbol{D}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))\,. \tag{5.27}$$

Here we define a privatized contraction coefficient for a generalized divergence $\boldsymbol{D}$ as follows:

$$\eta_{\boldsymbol{D}}^{\varepsilon} := \sup_{\substack{\mathcal{N}\in\mathcal{B}^{\varepsilon},\\ \rho,\sigma\in\mathcal{D},\\ \boldsymbol{D}(\rho\|\sigma)\neq 0}} \frac{\boldsymbol{D}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))}{\boldsymbol{D}(\rho\|\sigma)}, \tag{5.28}$$

where $\mathcal{B}^{\varepsilon}$ corresponds to the set of all ε-QLDP mechanisms, defined formally as

$$\mathcal{B}^{\varepsilon} := \left\{ \mathcal{N} \in \mathrm{CPTP} : \sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} E_{e^{\varepsilon}}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) = 0 \right\}. \tag{5.29}$$

In the above, we made use of the previously stated fact that (5.8) is equivalent to ε-QLDP.

Contraction under private channels has been studied in [62] by choosing normalized trace distance as the quantum divergence. In particular, Corollary V.I therein states that, for a channel $\mathcal{N}$ that satisfies ε-QLDP,

$$T(\mathcal{N}(\rho),\mathcal{N}(\sigma)) \leq \eta_T^{\varepsilon} T(\rho,\sigma), \tag{5.30}$$

where

$$\eta_T^{\varepsilon} \leq \frac{e^{\varepsilon}-1}{e^{\varepsilon}}. \tag{5.31}$$

Observe that Theorem 4 below improves this upper bound, and Theorem 5 states that the bound from Theorem 4 is in fact tight.

In this spirit, we define the following privatized contraction coefficient for the hockey-stick divergence E_γ:

$$\eta_{E_\gamma}^{\varepsilon} \coloneqq \sup_{\substack{\mathcal{N}\in\mathcal{B}^{\varepsilon},\\ \rho,\sigma\in\mathcal{D},\\ E_\gamma(\rho\|\sigma)\neq 0}} \frac{E_\gamma(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))}{E_\gamma(\rho\|\sigma)}. \tag{5.32}$$

Observe that $\gamma = 1$ corresponds to the normalized trace distance.

Next, we provide an upper bound on the privatized contraction coefficient defined in (5.32) for $\gamma \in [1, e^{\varepsilon}]$. Note that $\eta_{E_\gamma}^{\varepsilon} = 0$ for $\gamma \geq e^{\varepsilon}$, as a consequence of (5.8) and the fact that E_γ is monotone non-increasing in γ [29, Lemma 4.2]. The privatized contraction coefficient of the hockey-stick divergence in the classical privacy setting (i.e., analogous to Theorem 4) was established in [140, Theorem 1], and our proof below is inspired by their proof.

Theorem 4 (Privatized Contraction Coefficient of Hockey-Stick Divergence)**.** *For $\gamma \in [1, e^{\varepsilon}]$, the privatized contraction coefficient $\eta_{E_\gamma}^{\varepsilon}$ satisfies the following inequality:*

$$\eta_{E_\gamma}^{\varepsilon} \leq \frac{e^{\varepsilon} - \gamma}{e^{\varepsilon} + 1}. \tag{5.33}$$

Proof. Let $\mathcal{N}$ be an arbitrary channel from $\mathcal{L}(\mathcal{H}_A)$ to $\mathcal{L}(\mathcal{H}_B)$. By [62, Theorem II.2], we have that

$$\sup_{\rho,\sigma\in\mathcal{D}} \frac{E_\gamma(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))}{E_\gamma(\rho\|\sigma)} = \sup_{|\phi\rangle\perp|\psi\rangle} E_\gamma(\mathcal{N}(|\phi\rangle\langle\phi|)\|\mathcal{N}(|\psi\rangle\langle\psi|)), \tag{5.34}$$

where $|\phi\rangle \perp |\psi\rangle$ denotes orthogonal state vectors. Then consider that

$$\eta^{\varepsilon}_{E_\gamma} = \sup_{\mathcal{N}\in\mathcal{B}^{\varepsilon}} \sup_{|\phi\rangle\perp|\psi\rangle} E_\gamma(\mathcal{N}(|\phi\rangle\langle\phi|)\|\mathcal{N}(|\psi\rangle\langle\psi|)) \tag{5.35}$$

$$= \sup_{\substack{|\phi\rangle\perp|\psi\rangle,\\ \mathcal{N}\,:\, E_{e^\varepsilon}(\mathcal{N}(|\phi\rangle\langle\phi|)\|\mathcal{N}(|\psi\rangle\langle\psi|))=0,\\ E_{e^\varepsilon}(\mathcal{N}(|\psi\rangle\langle\psi|)\|\mathcal{N}(|\phi\rangle\langle\phi|))=0}} E_\gamma(\mathcal{N}(|\phi\rangle\langle\phi|)\|\mathcal{N}(|\psi\rangle\langle\psi|)) \tag{5.36}$$

$$\leq \sup_{\substack{\rho',\sigma'\in\mathcal{D}(\mathcal{H}_B),d_B\geq 1,\\ E_{e^\varepsilon}(\rho'\|\sigma')=E_{e^\varepsilon}(\sigma'\|\rho')=0}} E_\gamma(\rho'\|\sigma')\,, \tag{5.37}$$

where the second equality follows from the hockey-stick divergence equivalent form of QLDP given in (5.8), and the last inequality because the set being optimized over is larger. Note that the case $d_B = 1$ is trivial since the only one-dimensional state is the number 1, and the hockey-stick divergence is equal to zero for such one-dimensional states.

Next, we show that we can restrict the above optimization to a qubit space $d_B = 2$ as follows: consider the measurement channel

$$\mathcal{M}(\omega) := \mathrm{Tr}[\Pi_+\omega]|0\rangle\langle 0| + \mathrm{Tr}[(I-\Pi_+)\omega]|1\rangle\langle 1|, \tag{5.38}$$

where Π_+ is the projection onto the positive eigenspace of $\rho' - \gamma\sigma'$. With that, one can check that, for all $\gamma \geq 1$,

$$E_\gamma(\rho'\|\sigma') = E_\gamma(\mathcal{M}(\rho')\|\mathcal{M}(\sigma'))\,. \tag{5.39}$$

To this end, by data processing and non-negativity of the hockey-stick divergence, we arrive at

$$0 = E_{e^\varepsilon}(\rho'\|\sigma') \geq E_{e^\varepsilon}(\mathcal{M}(\rho')\|\mathcal{M}(\sigma')) \geq 0, \tag{5.40}$$

implying that $E_{e^\varepsilon}(\mathcal{M}(\rho')\|\mathcal{M}(\sigma')) = 0$. Proceeding with the above ideas, we can rewrite the right-hand side of (5.37) as

$$\sup_{\substack{\rho',\sigma'\in\mathcal{D}(\mathcal{H}_B),d_B\geq 2,\\ E_{e^\varepsilon}(\mathcal{M}(\rho')\|\mathcal{M}(\sigma'))=E_{e^\varepsilon}(\mathcal{M}(\sigma')\|\mathcal{M}(\rho'))=0}} E_\gamma(\mathcal{M}(\rho')\|\mathcal{M}(\sigma'))\,. \tag{5.41}$$

Define the state

$$\omega_p := p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|, \tag{5.42}$$

and observe that the output of the channel $\mathcal{M}$ is a special case of this state with $p = \mathrm{Tr}[\Pi_+ \omega]$. Then, we find an upper bound on (5.41), which is

$$\sup_{\substack{p,q\in[0,1],\\ E_{e^\varepsilon}(\omega_p\|\omega_q)=E_{e^\varepsilon}(\omega_q\|\omega_p)=0}} E_\gamma(\omega_p\|\omega_q). \tag{5.43}$$

Putting this together with the inequality in (5.37), we arrive at

$$\eta^\varepsilon_{E_\gamma} \leq \sup_{\substack{p,q\in[0,1],\\ E_{e^\varepsilon}\left(\omega_p\|\omega_q\right)=E_{e^\varepsilon}\left(\omega_q\|\omega_p\right)=0}} E_\gamma\left(\omega_p\|\omega_q\right) \tag{5.44}$$

$$= \sup_{\substack{p,q\in[0,1],\\ q\leq p\leq \min\{qe^\varepsilon,\ qe^{-\varepsilon}+1-e^{-\varepsilon}\}}} p - \gamma q \tag{5.45}$$

$$= \frac{e^\varepsilon - \gamma}{e^\varepsilon + 1}. \tag{5.46}$$

The fact that (5.44) and (5.46) are equal was stated in the proof of [140, Theorem 1]. Here we provide a short proof for convenience: The first equality above follows from the fact that the states ω_p and ω_q correspond to Bernoulli distributions with parameters p and q, respectively, and the reasoning given below: $E_{e^\varepsilon}\left(\omega_p\|\omega_q\right) = 0$ is equivalent to $p \leq e^\varepsilon q$ and $(1-p) \leq e^\varepsilon(1-q)$, and $E_{e^\varepsilon}\left(\omega_q\|\omega_p\right) = 0$ is equivalent to $q \leq e^\varepsilon p$ and $(1-q) \leq e^\varepsilon(1-p)$. Without loss of generality, suppose that $q \leq p$. Then all of these inequalities can be written in the following compact form: $q \leq p \leq \min\{qe^\varepsilon,\ qe^{-\varepsilon} + 1 - e^{-\varepsilon}\}$. The last equality follows from solving the linear program. Indeed, we can eliminate p by noting that $\min\{qe^\varepsilon,\ qe^{-\varepsilon} + 1 - e^{-\varepsilon}\} \leq 1$ and thus we can maximize its value by picking $p = \min\{qe^\varepsilon,\ qe^{-\varepsilon} + 1 - e^{-\varepsilon}\}$. From there, observe that $qe^\varepsilon \leq qe^{-\varepsilon} + 1 - e^{-\varepsilon}$ is equivalent to $q \leq 1/(e^\varepsilon + 1)$, and $qe^\varepsilon \geq qe^{-\varepsilon} + 1 - e^{-\varepsilon}$ is equivalent to $q \geq 1/(e^\varepsilon + 1)$. With this, $p - \gamma q$ is supremized at $q = 1/(e^\varepsilon + 1)$. □

Lemma 6. *Let $\gamma \geq 0$ and ρ and σ be states. Then,*

$$E_\gamma(\rho\|\sigma) = \gamma E_{\frac{1}{\gamma}}(\sigma\|\rho), \tag{5.47}$$

where $E_\gamma(\cdot\|\cdot)$ is defined in (2.3).

Proof. Let $0 \leq \gamma \leq 1$. By applying (2.3), consider that

$$\begin{aligned} E_\gamma(\rho\|\sigma) &= \sup_{0\leq M\leq I} \operatorname{Tr}[M(\rho - \gamma\sigma)] - (1-\gamma) && (5.48)\\ &= \sup_{0\leq M\leq I} \operatorname{Tr}[(I-M)(\rho - \gamma\sigma)] - (1-\gamma) && (5.49)\\ &= \sup_{0\leq M\leq I} \operatorname{Tr}[M(\gamma\sigma - \rho)] && (5.50)\\ &= \gamma \sup_{0\leq M\leq I} \operatorname{Tr}\left[M\left(\sigma - \frac{1}{\gamma}\rho\right)\right] && (5.51)\\ &= \gamma E_{\frac{1}{\gamma}}(\sigma\|\rho), && (5.52) \end{aligned}$$

where the last inequality follows by applying (2.3) with $1/\gamma \geq 1$ since $\gamma \leq 1$.

For $\gamma' \geq 1$, substituting $\gamma = 1/\gamma'$ in (5.52), we conclude the proof. □

Recall that $\mathcal{A} \in \mathcal{B}^\varepsilon$ is equivalent to

$$\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} E_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = 0. \tag{5.53}$$

Now, using Lemma 6, we also have that

$$\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} E_{e^{-\varepsilon}}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = 0. \tag{5.54}$$

Furthermore, using the monotonicity of $\gamma' \mapsto E_{\gamma'}$ for $\gamma' \geq 1$, we find the following for $\gamma \leq e^{-\varepsilon}$:

$$E_{\frac{1}{\gamma}}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq E_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = 0. \tag{5.55}$$

This then leads to $E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = 0$ for $\gamma \le e^{-\varepsilon}$ and

$$\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = 0. \tag{5.56}$$

Next, we quantify the contraction of the hockey-stick divergence for $\gamma \in [e^{-\varepsilon}, 1]$, similar to Theorem 4 for $\gamma \in [1, e^{\varepsilon}]$.

Corollary 4 (Contraction of Hockey-Stick Divergence under QLDP)**.** *Let $\gamma \in [e^{-\varepsilon}, 1]$, $\mathcal{A} \in \mathcal{B}^\varepsilon$, and ρ and σ states. Then, we have that*

$$E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \le \frac{\gamma e^\varepsilon - 1}{\gamma(e^\varepsilon + 1)} E_\gamma(\rho\|\sigma), \tag{5.57}$$

where $E_\gamma(\cdot\|\cdot)$ is defined in (2.3). Furthermore, for $\gamma \in [e^{-\varepsilon}, 1]$, the privatized contraction coefficient of hockey-stick divergence, $\eta^\varepsilon_{E_\gamma}$, satisfies the following inequality:

$$\eta^\varepsilon_{E_\gamma} \le \frac{\gamma e^\varepsilon - 1}{\gamma(e^\varepsilon + 1)}. \tag{5.58}$$

Proof. Since $\gamma \in [e^{-\varepsilon}, 1]$, it follows that $1/\gamma \in [1, e^\varepsilon]$. By applying Lemma 6, we obtain the following:

$$\begin{aligned} E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) &= \gamma E_{\frac{1}{\gamma}}(\mathcal{A}(\sigma)\|\mathcal{A}(\rho)) && (5.59)\\ &\le \gamma \frac{e^\varepsilon - 1/\gamma}{e^\varepsilon + 1} E_{\frac{1}{\gamma}}(\sigma\|\rho) && (5.60)\\ &= \frac{\gamma e^\varepsilon - 1}{\gamma(e^\varepsilon + 1)} E_\gamma(\rho\|\sigma), && (5.61) \end{aligned}$$

where the first inequality follows from Theorem 4 with $1/\gamma \ge 1$ and the last equality by applying Lemma 6 again.

By dividing both sides by $E_\gamma(\rho\|\sigma)$, and supremizing over all $\mathcal{A} \in \mathcal{B}^\varepsilon$ and all states such that $E_\gamma(\rho\|\sigma) \neq 0$, we arrive at the desired result. □

Theorem 4 and Corollary 4 strengthen the previously known upper bound on the privatized contraction coefficient for the normalized trace distance (i.e.,

η_T^ε by choosing $\gamma = 1$), as recalled in (5.31). Next, Theorem 5 states that the upper bound from Theorem 4, for trace distance, is indeed tight. Under ε-classical local differential privacy (recall Definition 14), Theorem 5 was presented as [72, Corollary 2.9]. Our proof method is different from that presented for [72, Corollary 2.9].

Theorem 5 (Privatized Contraction Coefficient of Trace Distance)**.** *Let $\varepsilon \geq 0$ and ρ and σ be states such that $T(\rho, \sigma) \neq 0$. We have*

$$\sup_{\mathcal{N} \in \mathcal{B}^\varepsilon} \frac{T(\mathcal{N}(\rho), \mathcal{N}(\sigma))}{T(\rho, \sigma)} = \frac{e^\varepsilon - 1}{e^\varepsilon + 1}. \tag{5.62}$$

Consequently, the privatized contraction coefficient for the trace distance under ε-QLDP is given by

$$\eta_T^\varepsilon = \frac{e^\varepsilon - 1}{e^\varepsilon + 1}. \tag{5.63}$$

Proof. In Proposition 23, choose $M = \Pi_+$ for the measurement channel $\mathcal{M}$, which is the projection onto the positive eigenspace of $\rho - \sigma$. We then conclude that

$$T(\mathcal{M}(\rho), \mathcal{M}(\sigma)) = T(\rho, \sigma). \tag{5.64}$$

To this end, consider the composition $\mathcal{A}_{\text{Dep}}^p \circ \mathcal{M}$,

$$T\Big((\mathcal{A}_{\text{Dep}}^p \circ \mathcal{M})(\rho), (\mathcal{A}_{\text{Dep}}^p \circ \mathcal{M})(\sigma)\Big) = (1 - p) T(\rho, \sigma), \tag{5.65}$$

and choose $p = 2/(e^\varepsilon + 1)$ to ensure that $\mathcal{A}_{\text{Dep}}^p \circ \mathcal{M}$ is ε-QLDP.

Then, for the channel $\mathcal{A}_{\text{Dep}}^p \circ \mathcal{M}$ and states ρ and σ, we find that

$$\frac{T\Big((\mathcal{A}_{\text{Dep}}^p \circ \mathcal{M})(\rho), (\mathcal{A}_{\text{Dep}}^p \circ \mathcal{M})(\sigma)\Big)}{T(\rho, \sigma)} = 1 - p = \frac{e^\varepsilon - 1}{e^\varepsilon + 1}. \tag{5.66}$$

In addition, by the definition of η_T^ε, for this choice of channel and states, we get the following inequality:

$$\frac{T\Big((\mathcal{A}_{\text{Dep}}^p \circ \mathcal{M})(\rho), (\mathcal{A}_{\text{Dep}}^p \circ \mathcal{M})(\sigma)\Big)}{T(\rho, \sigma)} = \frac{e^\varepsilon - 1}{e^\varepsilon + 1} \leq \eta_T^\varepsilon. \tag{5.67}$$

By picking $\gamma = 1$ in Theorem 4, we also conclude that

$$\eta^{\varepsilon}_{E_1} = \eta^{\varepsilon}_{T} \leq \frac{e^{\varepsilon} - 1}{e^{\varepsilon} + 1}. \tag{5.68}$$

We conclude the equality in (5.62) by combining (5.67) and (5.68), and the equality in (5.63) by supremizing over ρ and σ such that $T(\rho, \sigma) \neq 0$. □

Next, we apply the findings of Theorem 4 and Theorem 5 to obtain contraction bounds for other quantum divergences under QLDP privacy constraints.

Proposition 24 (Contraction of Bures Distance under ε-QLDP)**.** *Let $\mathcal{A}$ be an ε-QLDP mechanism, and let ρ and σ be states. Then*

$$\left[d_B(\mathcal{A}(\rho), \mathcal{A}(\sigma))\right]^2 \leq 2\frac{(e^{\varepsilon/2} - 1)^2}{e^{\varepsilon} + 1} T(\rho, \sigma). \tag{5.69}$$

Proof. Recall from [63, Eq. (5.50)],

$$\left[d_B(\rho, \sigma)\right]^2 = 2\left(1 - \sqrt{F}(\rho, \sigma)\right) \leq H_{1/2}(\rho \| \sigma)\,, \tag{5.70}$$

where

$$H_{1/2}(\rho \| \sigma) := \frac{1}{2}\int_1^{\infty} \left[E_{\gamma}(\rho \| \sigma) + E_{\gamma}(\sigma \| \rho)\right] \gamma^{-3/2}\, \mathrm{d}\gamma. \tag{5.71}$$

Now, as a consequence of the fact that $\mathcal{A} \in \mathcal{B}^{\varepsilon}$ implies that $E_{e^{\varepsilon}}(\mathcal{A}(\rho) \| \mathcal{A}(\sigma)) = 0$ and $E_{e^{\varepsilon}}(\mathcal{A}(\sigma) \| \mathcal{A}(\rho)) = 0$, consider that

$$\sup_{\mathcal{A} \in \mathcal{B}^{\varepsilon}} H_{1/2}(\mathcal{A}(\rho) \| \mathcal{A}(\sigma)) \leq \sup_{\substack{\mathcal{A}\,:\, E_{e^{\varepsilon}}(\mathcal{A}(\rho) \| \mathcal{A}(\sigma)) = 0, \\ E_{e^{\varepsilon}}(\mathcal{A}(\sigma) \| \mathcal{A}(\rho)) = 0}} H_{1/2}(\mathcal{A}(\rho) \| \mathcal{A}(\sigma))\,. \tag{5.72}$$

We use this to show the following, employing similar techniques used in the proof of [8, Theorem 1]:

$$\sup_{\mathcal{A} \in \mathcal{B}^{\varepsilon}} H_{1/2}(\mathcal{A}(\rho) \| \mathcal{A}(\sigma)) \leq 2\frac{(e^{\varepsilon/2} - 1)^2}{e^{\varepsilon} + 1} T(\rho, \sigma). \tag{5.73}$$

To this end, by definition,

$$\sup_{\substack{\mathcal{A}:\, E_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))=0,\\ E_{e^\varepsilon}(\mathcal{A}(\sigma)\|\mathcal{A}(\rho))=0}} H_{1/2}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))$$

$$= \sup_{\substack{\mathcal{A}:\, E_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))=0,\\ E_{e^\varepsilon}(\mathcal{A}(\sigma)\|\mathcal{A}(\rho))=0}} \frac{1}{2}\int_1^{\infty} \Big[E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) + E_\gamma(\mathcal{A}(\sigma)\|\mathcal{A}(\rho))\Big]\gamma^{-3/2}\,\mathrm{d}\gamma \tag{5.74}$$

$$= \sup_{\substack{\mathcal{A}:\, E_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))=0,\\ E_{e^\varepsilon}(\mathcal{A}(\sigma)\|\mathcal{A}(\rho))=0}} \frac{1}{2}\int_1^{e^\varepsilon} \Big[E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) + E_\gamma(\mathcal{A}(\sigma)\|\mathcal{A}(\rho))\Big]\gamma^{-3/2}\,\mathrm{d}\gamma \tag{5.75}$$

$$\leq \sup_{\substack{\mathcal{A}:\, E_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))=0,\\ E_{e^\varepsilon}(\mathcal{A}(\sigma)\|\mathcal{A}(\rho))=0}} \frac{1}{2}\int_1^{e^\varepsilon} \frac{e^\varepsilon-\gamma}{e^\varepsilon+1}\Big[E_\gamma(\rho\|\sigma) + E_\gamma(\sigma\|\rho)\Big]\gamma^{-3/2}\,\mathrm{d}\gamma \tag{5.76}$$

$$= \frac{1}{2}\int_1^{e^\varepsilon} \frac{e^\varepsilon-\gamma}{e^\varepsilon+1}\Big[E_\gamma(\rho\|\sigma) + E_\gamma(\sigma\|\rho)\Big]\gamma^{-3/2}\,\mathrm{d}\gamma \tag{5.77}$$

$$\leq E_1(\rho\|\sigma)\int_1^{e^\varepsilon} \frac{e^\varepsilon-\gamma}{e^\varepsilon+1}\gamma^{-3/2}\,\mathrm{d}\gamma \tag{5.78}$$

$$= 2\frac{(e^{\varepsilon/2}-1)^2}{(e^\varepsilon+1)}T(\rho,\sigma), \tag{5.79}$$

where the second equality follows from the fact that $E_{e^\varepsilon}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = E_{e^\varepsilon}(\mathcal{A}(\sigma)\|\mathcal{A}(\rho)) = 0$ and $\gamma \mapsto E_\gamma(\cdot\|\cdot)$ is a monotonically decreasing function for $\gamma \geq 1$; first inequality by applying Theorem 4; third equality by $E_\gamma(\rho\|\sigma) \leq E_1(\rho\|\sigma)$ for $\gamma \geq 1$; and finally the last equality by solving the integral and substituting $E_1(\cdot\|\cdot) = T(\cdot,\cdot)$.

With the use of (5.70) and the inequality derived above, we conclude the proof of the upper bound. □

Remark 31 (Another Contraction Bound for Bures Distance)**.** *By [142, Theorem 2.1] we have*

$$1 - \sqrt{F}(\mathcal{A}(\rho),\mathcal{A}(\sigma)) \leq T(\mathcal{A}(\rho),\mathcal{A}(\sigma))\frac{e^{D_{\max}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))/2}}{1+e^{D_{\max}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))/2}}. \tag{5.80}$$

Applying Theorem 5 together with the fact that $\mathcal{A} \in \mathcal{B}^\varepsilon$ is equivalent to

$$\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} D_{\max}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \varepsilon, \tag{5.81}$$

for all $\rho, \sigma \in \mathcal{D}$, we obtain the following:

$$[d_B(\mathcal{A}(\rho), \mathcal{A}(\sigma))]^2 \leq 2\frac{(e^\varepsilon - 1)e^{\varepsilon/2}}{(e^\varepsilon + 1)(e^{\varepsilon/2} + 1)} T(\rho, \sigma). \tag{5.82}$$

However, this bound is weaker than the bound presented in Proposition 24.

Proposition 25 (Contraction of Quantum Relative Entropy under Local Privacy)**.** *Let $\mathcal{A}$ be an ε-QLDP mechanism, and let ρ and σ be quantum states. Then*

$$D(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \varepsilon \left(\frac{e^\varepsilon - 1}{e^\varepsilon + 1}\right) T(\rho, \sigma). \tag{5.83}$$

Proof. Using the integral form of the quantum relative entropy in [63, Eq. (1.6)], we have

$$D(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = \int_1^\infty \frac{1}{\gamma} E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) + \frac{1}{\gamma^2} E_\gamma(\mathcal{A}(\sigma)\|\mathcal{A}(\rho)) \, \mathrm{d}\gamma \tag{5.84}$$

$$= \int_1^{r_1} \frac{1}{\gamma} E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \, \mathrm{d}\gamma + \int_1^{r_2} \frac{1}{\gamma^2} E_\gamma(\mathcal{A}(\sigma)\|\mathcal{A}(\rho)) \, \mathrm{d}\gamma, \tag{5.85}$$

where

$$r_1 \equiv e^{D_{\max}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))}, \qquad r_2 \equiv e^{D_{\max}(\mathcal{A}(\sigma)\|\mathcal{A}(\rho))}, \tag{5.86}$$

and we made use of [63, Eq. (2.3)]. Since $\mathcal{A} \in \mathcal{B}^\varepsilon$, we have $\sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} D_{\max}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \varepsilon$ and we arrive at

$$D(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \int_1^{e^\varepsilon} \frac{1}{\gamma} E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \, \mathrm{d}\gamma + \int_1^{e^\varepsilon} \frac{1}{\gamma^2} E_\gamma(\mathcal{A}(\sigma)\|\mathcal{A}(\rho)) \, \mathrm{d}\gamma \tag{5.87}$$

$$\leq \int_1^{e^\varepsilon} \frac{1}{\gamma} \frac{e^\varepsilon - \gamma}{e^\varepsilon + 1} E_\gamma(\rho\|\sigma) + \frac{1}{\gamma^2} \frac{e^\varepsilon - \gamma}{e^\varepsilon + 1} E_\gamma(\sigma\|\rho) \, \mathrm{d}\gamma \tag{5.88}$$

$$\leq T(\rho, \sigma) \left(\int_1^{e^\varepsilon} \frac{1}{\gamma} \frac{e^\varepsilon - \gamma}{e^\varepsilon + 1} + \frac{1}{\gamma^2} \frac{e^\varepsilon - \gamma}{e^\varepsilon + 1} \, \mathrm{d}\gamma \right) \tag{5.89}$$

$$= \varepsilon \frac{e^\varepsilon - 1}{e^\varepsilon + 1} T(\rho, \sigma), \tag{5.90}$$

where the second inequality follows from Theorem 4; the third inequality from $E_\gamma(\rho\|\sigma) \leq E_1(\rho\|\sigma) = T(\rho, \sigma)$ for all $\gamma \geq 1$ due to monotonicity of the hockey-stick divergence with respect to γ; and finally the last equality by evaluating the integral. □

Note that the inequality in (5.83) can alternatively be obtained by using [63, Proposition 5.3] together with Theorem 5. In that case, it is also important to note that Theorem 4 is a key ingredient in the proof of Theorem 5 to obtain a tight privatized contraction coefficient for the trace distance.

Let $f : (0, \infty) \to \mathbb{R}$ be a convex and twice differentiable function satisfying $f(1) = 0$. Then, for all quantum states ρ and σ, recall the quantum f-divergence defined in [63, Definition 2.3]:

$$D_f(\rho\|\sigma) := \int_1^\infty f''(\gamma) E_\gamma(\rho\|\sigma) + \gamma^{-3} f''(\gamma^{-1}) E_\gamma(\sigma\|\rho)\, \mathrm{d}\gamma. \tag{5.91}$$

Using Lemma 6, we obtain the following equivalent expression for the above quantum f-divergence.

Proposition 26 (Equivalent Expression for f-Divergence)**.** *Let ρ and σ be states. Then, $D_f(\rho\|\sigma)$ defined in (5.91) is equivalent to the following expression:*

$$D_f(\rho\|\sigma) = \int_0^\infty f''(\gamma) E_\gamma(\rho\|\sigma)\, \mathrm{d}\gamma, \tag{5.92}$$

where $E_\gamma(\cdot\|\cdot)$ is defined in (2.3).

Proof. Consider that

$$\begin{aligned}
\int_1^\infty \gamma^{-3} f''(\gamma^{-1}) E_\gamma(\sigma\|\rho)\, \mathrm{d}\gamma &= \int_1^\infty \gamma^{-3} f''(\gamma^{-1}) \gamma E_{\frac{1}{\gamma}}(\rho\|\sigma)\, \mathrm{d}\gamma && (5.93)\\
&= \int_1^\infty \gamma^{-2} f''(\gamma^{-1}) E_{\frac{1}{\gamma}}(\rho\|\sigma)\, \mathrm{d}\gamma && (5.94)\\
&= \int_0^1 f''(\nu) E_\nu(\rho\|\sigma)\, \mathrm{d}\nu && (5.95)
\end{aligned}$$

where the first equality follows from Lemma 6 and the last equality follows by change of variables with the substitution $\nu = \gamma^{-1}$.

Then, we have that

$$D_f(\rho\|\sigma) = \int_1^\infty f''(\gamma)E_\gamma(\rho\|\sigma)\,\mathrm{d}\gamma + \int_0^1 f''(\nu)E_\nu(\rho\|\sigma)\,\mathrm{d}\nu \tag{5.96}$$

$$= \int_0^\infty f''(\gamma)E_\gamma(\rho\|\sigma)\,\mathrm{d}\gamma, \tag{5.97}$$

concluding the proof. □

Similar to previous cases, the privatized contraction of such an f-divergence (defined in (5.91)) relative to the normalized trace distance can be bounded from above as follows:

Proposition 27 (Contraction of f-Divergences under Privacy Constraints)**.** *Let $\mathcal{A}$ be an ε-QDP mechanism. Then for all ρ, σ states we have*

$$D_f(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \frac{f(e^\varepsilon) + e^\varepsilon f(e^{-\varepsilon})}{e^\varepsilon + 1} T(\rho, \sigma), \tag{5.98}$$

where $D_f(\cdot\|\cdot)$ is defined in (5.91)*.*

Proof. Consider that

$$D_f(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \int_1^{e^\varepsilon} f''(\gamma)\frac{e^\varepsilon - \gamma}{e^\varepsilon + 1}E_\gamma(\rho\|\sigma) + \gamma^{-3}f''(\gamma^{-1})\frac{e^\varepsilon - \gamma}{e^\varepsilon + 1}E_\gamma(\sigma\|\rho)\,\mathrm{d}\gamma \tag{5.99}$$

$$\leq \frac{1}{e^\varepsilon + 1}\left(\int_1^{e^\varepsilon} (f''(\gamma) + \gamma^{-3}f''(\gamma^{-1}))(e^\varepsilon - \gamma)\,\mathrm{d}\gamma\right) T(\rho, \sigma) \tag{5.100}$$

$$= \frac{f(e^\varepsilon) + e^\varepsilon f(e^{-\varepsilon})}{e^\varepsilon + 1} T(\rho, \sigma), \tag{5.101}$$

where the first inequality follows from Theorem 4 together with (5.8), the second inequality from $E_\gamma(\rho\|\sigma) \leq T(\rho, \sigma)$ for all $\gamma \geq 1$ and the last inequality by integration by parts with the use of $f(1) = 0$ and $\frac{\mathrm{d}^2}{\mathrm{d}\gamma^2}\gamma f(\gamma^{-1}) = \gamma^{-3}f''(\gamma^{-1})$ (also see the proof of [63, Proposition 5.2]). □

Note that the inequality in (5.98) is tighter than the bounds obtained by combining [63, Proposition 5.2] with the contraction bound previously known for trace distance (i.e., that in (5.31)).

Proposition 28. *For $\varepsilon \geq 0$, the following equality holds:*

$$\sup_{\substack{A \in \mathcal{B}^{\varepsilon}, \\ \rho,\sigma \in \mathcal{D}}} D_f(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = \frac{f(e^{\varepsilon}) + e^{\varepsilon} f(e^{-\varepsilon})}{e^{\varepsilon} + 1}, \tag{5.102}$$

where $D_f(\cdot\|\cdot)$ is defined in (5.91) and the optimization is over all ε-QLDP channels $\mathcal{A}$ and states ρ and σ. The equality is achieved by a pair of orthogonal states and a channel $\mathcal{A}$ with the output dimension two.

Proof. See Appendix C.3. □

Define the chi-square divergence as follows:

$$\chi^2(\rho\|\sigma) := 2 \int_1^{\infty} E_{\gamma}(\rho\|\sigma) + \gamma^{-3} E_{\gamma}(\sigma\|\rho)\, \mathrm{d}\gamma. \tag{5.103}$$

We also mention a useful upper bound on the contraction of the chi-square divergence derived in [23].

Proposition 29 ([23, Lemma 4.7, Eq. (4.17), Eq. (4.19)])**.** *Let $\mathcal{A} \in \mathcal{B}^{\varepsilon}$. Then, for all ρ and σ states, we have that*

$$\chi^2(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq 2e^{-\varepsilon}(e^{\varepsilon} - 1)^2 \left[T(\rho, \sigma)\right]^2. \tag{5.104}$$

5.3.2 Contraction under (ε, δ)-QLDP channels

Until now in this section, we have focused on the privacy constraints imposed by ε-QLDP with $\delta = 0$. Generalizing (5.28), set $\delta \geq 0$, and define privatized

contraction coefficients under (ε, δ)-QLDP privacy constraints as follows:

$$\eta_{\boldsymbol{D}}^{\varepsilon,\delta} := \sup_{\substack{\mathcal{N}\in\mathcal{B}^{\varepsilon,\delta},\\ \rho,\sigma\in\mathcal{D},\\ \boldsymbol{D}(\rho\|\sigma)\neq 0}} \frac{\boldsymbol{D}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))}{\boldsymbol{D}(\rho\|\sigma)}, \tag{5.105}$$

where $\mathcal{B}^{\varepsilon,\delta}$ corresponds to the set of all (ε, δ)-QLDP mechanisms, defined formally as

$$\mathcal{B}^{\varepsilon,\delta} := \left\{ \mathcal{N} \in \text{CPTP} : \sup_{\rho,\sigma\in\mathcal{D}(\mathcal{H})} E_{e^{\varepsilon}}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \leq \delta \right\}. \tag{5.106}$$

Next, we quantify the contraction of the normalized trace distance by instantiating the generalized divergence to be the normalized trace distance. The following extends Theorem 5 and provides an alternative proof for the converse bound in the proof of Theorem 5.

Theorem 6 (Contraction of Trace Distance under (ε, δ)-QLDP Constraints)**.** *Let $\varepsilon, \delta \geq 0$, and let ρ and σ be states such that $T(\rho, \sigma) \neq 0$. Then*

$$\sup_{\mathcal{N}\in\mathcal{B}^{\varepsilon,\delta}} \frac{T(\mathcal{N}(\rho), \mathcal{N}(\sigma))}{T(\rho, \sigma)} = \frac{e^{\varepsilon} - 1 + 2\delta}{e^{\varepsilon} + 1}. \tag{5.107}$$

Consequently, the privatized contraction coefficient for the trace distance under (ε, δ)-QLDP is given by

$$\eta_T^{\varepsilon,\delta} = \frac{e^{\varepsilon} - 1 + 2\delta}{e^{\varepsilon} + 1}. \tag{5.108}$$

Proof. Converse bound: Let ρ and σ be states, let $\gamma \geq 1$, and consider that

$$\begin{aligned}
(\gamma + 1)T(\rho, \sigma) &= (\gamma + 1) \sup_{0\leq M\leq I} \text{Tr}[M(\rho - \sigma)] && (5.109)\\
&= \sup_{0\leq M\leq I} \{\text{Tr}[M(\rho - \gamma\sigma)] + \text{Tr}[M(\gamma\rho - \sigma)]\} && (5.110)\\
&\leq \sup_{0\leq M\leq I} \{\text{Tr}[M(\rho - \gamma\sigma)]\} + \sup_{0\leq M\leq I} \{\text{Tr}[M(\gamma\rho - \sigma)]\} && (5.111)\\
&= E_{\gamma}(\rho\|\sigma) + \gamma \sup_{0\leq M\leq I} \text{Tr}\left[M\left(\rho - \frac{1}{\gamma}\sigma\right)\right] && (5.112)\\
&= E_{\gamma}(\rho\|\sigma) + E_{\gamma}(\sigma\|\rho) + \gamma - 1, && (5.113)
\end{aligned}$$

where the penultimate equality follows from the variational representation of the hockey-stick divergence from [62, Lemma II.1] and the last equality follows from the following reasoning (see also the symmetry property of hockey-stick divergence in [62, Eq. (II.21)]):

$$E_\gamma(\sigma\|\rho) = \sup_{0\le M\le I} \mathrm{Tr}[M(\sigma-\gamma\rho)] \tag{5.114}$$

$$= \sup_{0\le M\le I} \mathrm{Tr}[(I-M)(\sigma-\gamma\rho)] \tag{5.115}$$

$$= \mathrm{Tr}[\sigma-\gamma\rho] + \sup_{0\le M\le I} \mathrm{Tr}[M(\gamma\rho-\sigma)] \tag{5.116}$$

$$= (1-\gamma) + \gamma \sup_{0\le M\le I} \mathrm{Tr}\left[M\left(\rho - \frac{1}{\gamma}\sigma\right)\right]. \tag{5.117}$$

With the above setup, choose $\rho = \mathcal{A}(|\phi\rangle\langle\phi|)$ and $\sigma = \mathcal{A}(|\psi\rangle\langle\psi|)$, where $\mathcal{A} \in \mathcal{B}^{\varepsilon,\delta}$ and $|\phi\rangle\langle\phi|, |\psi\rangle\langle\psi|$ are orthogonal pure states together with $\gamma = e^\varepsilon$. Then, we have

$$\begin{aligned}
&(e^\varepsilon+1)T(\mathcal{A}(|\phi\rangle\langle\phi|), \mathcal{A}(|\psi\rangle\langle\psi|)) \\
&\le E_{e^\varepsilon}(\mathcal{A}(|\phi\rangle\langle\phi|)\|\mathcal{A}(|\psi\rangle\langle\psi|)) + E_{e^\varepsilon}(\mathcal{A}(|\psi\rangle\langle\psi|)\|\mathcal{A}(|\phi\rangle\langle\phi|)) + e^\varepsilon - 1 &(5.118)\\
&\le 2\delta + e^\varepsilon - 1, &(5.119)
\end{aligned}$$

where the last inequality follows due to $\mathcal{A} \in \mathcal{B}^{\varepsilon,\delta}$ and applying the constraints $E_{e^\varepsilon}(\rho\|\sigma) \le \delta$ and $E_{e^\varepsilon}(\sigma\|\rho) \le \delta$. We arrive at

$$T(\mathcal{A}(|\phi\rangle\langle\phi|), \mathcal{A}(|\psi\rangle\langle\psi|)) \le \frac{2\delta + e^\varepsilon - 1}{e^\varepsilon + 1}. \tag{5.120}$$

Next, with the choice $\gamma = 1$ in (5.34), we have for all $\mathcal{A} \in \mathcal{B}^{\varepsilon,\delta}$

$$\begin{aligned}
\sup_{\rho,\sigma\in\mathcal{D}} \frac{T(\mathcal{A}(\rho), \mathcal{A}(\sigma))}{T(\rho,\sigma)} &= \sup_{|\phi\rangle\perp|\psi\rangle} T(\mathcal{A}(|\phi\rangle\langle\phi|), \mathcal{A}(|\psi\rangle\langle\psi|)) &(5.121)\\
&\le \frac{2\delta + e^\varepsilon - 1}{e^\varepsilon + 1}, &(5.122)
\end{aligned}$$

where the last inequality follows from (5.120).

So we conclude that

$$\sup_{\mathcal{A}\in\mathcal{B}^{\varepsilon,\delta}} \frac{T(\mathcal{A}(\rho),\mathcal{A}(\sigma))}{T(\rho,\sigma)} \leq \eta_T^{\varepsilon,\delta} \leq \frac{2\delta + e^{\varepsilon} - 1}{e^{\varepsilon}+1}. \tag{5.123}$$

Achievability bound: Let $\mathcal{A} = \mathcal{A}^p_{\text{Dep}} \circ \mathcal{M}$ with $p = 2(1-\delta)/(e^{\varepsilon}+1)$, where $\mathcal{M}$ corresponds to (5.18). First, we need to show that $\mathcal{A} \in \mathcal{B}^{\varepsilon,\delta}$. To this end, consider that

$$\begin{aligned} E_{e^{\varepsilon}}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) &\leq \max\left\{0, (1-e^{\varepsilon})\frac{p}{2} + (1-p)E_{e^{\varepsilon}}(\mathcal{M}(\rho)\|\mathcal{M}(\sigma))\right\} && (5.124)\\ &\leq \max\left\{0, (1-e^{\varepsilon})\frac{p}{2} + (1-p)\right\} && (5.125)\\ &\leq \max\{0,\delta\} && (5.126)\\ &= \delta, && (5.127) \end{aligned}$$

where the first inequality follows from [62, Lemma IV.1], the second inequality from $E_{e^{\varepsilon}}(\mathcal{M}(\rho)\|\mathcal{M}(\sigma)) \leq 1$, and the last inequality by substituting $p = 2(1-\delta)/(e^{\varepsilon}+1)$. This shows that for all ρ,σ, $E_{e^{\varepsilon}}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \delta$, proving that $\mathcal{A} \in \mathcal{B}^{\varepsilon,\delta}$.

Now choose M in (5.18) to be the projection onto the positive eigenspace of $\rho - \sigma$. With that, we have

$$\begin{aligned} T(\mathcal{A}(\rho),\mathcal{A}(\sigma)) &= (1-p)T(\rho,\sigma) && (5.128)\\ &= \frac{2\delta + e^{\varepsilon} - 1}{e^{\varepsilon}+1} T(\rho,\sigma), && (5.129) \end{aligned}$$

showing that for all ρ,σ such that $T(\rho,\sigma) \neq 0$ with the specific choice of $\mathcal{A} = \mathcal{A}^p_{\text{Dep}} \circ \mathcal{M}$, we achieve the upper bound in (5.123), concluding the proof. □

Corollary 5 (Contraction of f-divergences under (ε,δ)-QLDP)**.** *Let $\varepsilon,\delta \geq 0$ with $\mathcal{A}$ satisfying (ε,δ)-QLDP. For states ρ and σ, we have*

$$D_f(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \frac{e^{\varepsilon} - 1 + 2\delta}{e^{\varepsilon}+1} D_f(\rho\|\sigma), \tag{5.130}$$

where $D_f(\cdot\|\cdot)$ is defined in (5.91).

Proof. Proof follows by applying the fact that the contraction coefficient of the f-divergence is upper bounded by the contraction coefficient of trace distance [63, Lemma 4.1] together with Theorem 6. □

Remark 32 (Improvement Compared to Existing Results)**.** *[62, Corollary V.I] states the following upper bound:*

$$\eta_T^{\varepsilon,\delta} \leq \frac{e^{\varepsilon} - 1 + \delta}{e^{\varepsilon}}. \tag{5.131}$$

Theorem 6 provides a tight characterization of the privatized contraction coefficient under (ε, δ)-QLDP constraints, which can be seen from

$$\frac{e^{\varepsilon} - 1 + \delta}{e^{\varepsilon}} - \frac{e^{\varepsilon} - 1 + 2\delta}{e^{\varepsilon} + 1} = \frac{(1-\delta)(e^{\varepsilon} - 1)}{e^{\varepsilon}(e^{\varepsilon} + 1)} \geq 0. \tag{5.132}$$

Note that Theorem 6 also improves the best known bound for the contraction of total variation distance in the classical setting (See [7, Lemma 1]).

By using Theorem 6 instead of (5.131), it is possible to improve the upper bound on the error exponent in asymmetric hypothesis testing with privatized quantum states ($\mathcal{A}(\rho)$ and $\mathcal{A}(\sigma)$), as presented in [63, Corollary 5.14], as follows: For all $\mathcal{A}$ satisfying (ε, δ)-QLDP we have that

$$\lim_{n\to\infty} -\frac{1}{n} \ln \beta_\nu(\mathcal{A}(\rho)^{\otimes n} \| \mathcal{A}(\sigma)^{\otimes n}) = D(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \frac{e^{\varepsilon} - 1 + 2\delta}{e^{\varepsilon} + 1} D(\rho\|\sigma), \tag{5.133}$$

where $\nu \in (0, 1)$ and

$$\beta_\nu(\mathcal{A}(\rho)^{\otimes n} \| \mathcal{A}(\sigma)^{\otimes n}) := \inf_{\Lambda^{(n)}} \left\{ \begin{array}{c} \mathrm{Tr}[\Lambda^{(n)} \mathcal{A}(\sigma)^{\otimes n}] : \\ \mathrm{Tr}[(I^{\otimes n} - \Lambda^{(n)}) \mathcal{A}(\rho)^{\otimes n}] \leq \nu, \\ 0 \leq \Lambda^{(n)} \leq I^{\otimes n} \end{array} \right\}. \tag{5.134}$$

We study further on the non-asymptotic setting (finite $n \in \mathbb{N}$) of hypothesis testing with privatized states in Chapter 6.

5.4 Concluding Remarks

In this chapter, we derived upper bounds on the contraction of quantum divergences under privacy constraints imposed by channels that satisfy ε-QLDP. Notably, we fully characterized the privatized contraction coefficient of trace distance, proving the converse by deriving an upper bound on the contraction of the hockey-stick divergence, and proving achievability by proposing a novel QLDP mechanism that achieves the bound. In addition, we also characterized the privatized contraction coefficient of trace distance under (ε, δ)-QLDP mechanisms.

We utilize the tools developed in this chapter to analyze the cost of privacy one needs to pay in ensuring privacy in statistical testing and learning tasks in Chapter 6.

CHAPTER 6

PRIVATE HYPOTHESIS TESTING AND LEARNING

6.1 Introduction

In this chapter, we study statistical tasks under quantum privacy tasks by utilizing the tools developed in Chapter 5.

One can reexamine classical and quantum information processing tasks under privacy constraints to quantify the cost one is expected to pay to ensure privacy. One such task is symmetric hypothesis testing, which is a well-studied operational task both in classical and quantum information theory [11, 59, 65]. In this task, a weighted average of the type I and type II error probabilities is minimized, where weightings are based on prior beliefs of the occurrence of the two hypotheses. For distinguishing two probability distributions (classical) and two quantum states (quantum), asymptotically optimal error exponents are characterized by the classical Chernoff bound [25, 64] and its quantum generalization [9, 10, 101], respectively.

In the classical setting, the non-asymptotic regime of symmetric hypothesis testing has been well studied. To this end, the number of samples (i.e., the sample complexity) needed to distinguish two probability distributions p and q up to a fixed non-zero error is characterized by $\Theta(1/H^2(p,q))$, where $H^2(\cdot,\cdot)$ is the square of the Hellinger divergence, defined as

$$H^2(p,q) := 2\left(1 - \sum_{x\in\mathcal{X}} \sqrt{p(x)q(x)}\right), \tag{6.1}$$

for $p, q \in \mathcal{P}(\mathcal{X})$ discrete probability distributions over the domain $\mathcal{X}$ (see also the recent work [107]). The sample complexity of symmetric quantum hypoth-

esis testing was recently studied in [22], showing that it is characterized by $\Theta(1/\left[d_{\mathrm{B}}(\rho,\sigma)\right]^2)$, where $\left[d_{\mathrm{B}}(\rho,\sigma)\right]^2$ is the Bures distance defined in (2.5) (see also Theorem 7).

Symmetric hypothesis testing under LDP constraints (when one has access to privatized data samples in contrast to the data from the original distributions) has been extensively studied in a chain of works [8,16,38,39,105]. In [38], a contraction bound for the Kullback–Leibler divergence relative to the total-variation distance has been derived, which implies that one needs

$$n = \Theta\left(\frac{1}{\varepsilon^2[\mathrm{TV}(p,q)]^2}\right) \tag{6.2}$$

samples to achieve a constant error probability while ensuring ε-LDP, for $\varepsilon \ll 1$. In [8, Lemma 2], the following lower and upper bounds were established to achieve ε-LDP, for $\varepsilon > 0$, and a fixed error probability of $1/10$, with the prior probabilities of the distributions occurring fixed to $1/2$:

$$\frac{4}{35}\max\left\{\frac{e^{\varepsilon}}{2(e^{\varepsilon}-1)^2[\mathrm{TV}(p,q)]^2}, \frac{(e^{\varepsilon}+1)^2}{(e^{\varepsilon}-1)^2H^2(p,q)}\right\} \leq n \leq \frac{2\ln(5)(e^{\varepsilon}+1)^2}{(e^{\varepsilon}-1)^2[\mathrm{TV}(p,q)]^2}. \tag{6.3}$$

In the quantum setting, under local privacy constraints, an upper bound on the error exponent in asymmetric hypothesis testing (minimizing the type II error probability while the type I error probability is fixed) was presented in [63, Corollary 5.14]. However, the study of the non-asymptotic regime of symmetric quantum hypothesis testing under privacy constraints has been left as an open research direction.

6.1.1 Contributions

In this chapter, we study statistical tasks including testing and learning under quantum privacy constraints imposed by QLDP.

First, we consider the non-asymptotic regime of the statistical task of quantum symmetric hypothesis testing when we have access to privatized quantum states. Together with the tools developed in Chapter 5, we provide upper and lower bounds on the sample complexity when symmetric hypothesis testing is carried out with access to privatized quantum states (Theorem 8). For the specific case in which the states are orthogonal, we prove that the sample complexity scales as $\Theta\left([(e^{\varepsilon}+1)/(e^{\varepsilon}-1)]^2\right)$, showcasing the cost of privacy (Corollary 6). Furthermore, for $\varepsilon \in (0,1)$, we establish that the sample complexity is characterized by $\Theta\left((1/[\varepsilon T(\rho,\sigma)])^2\right)$. In these two cases, we show that a channel belonging to QLDP mechanisms proposed in Proposition 23 achieves order optimality.

We also find that under certain low-privacy regimes (higher ε), the private sample complexity behaves similarly to the non-private setting (Proposition 30). We extend our analysis to private sample complexity bounds for multiple hypothesis testing and asymmetric hypothesis testing in Remark 37 and Remark 38, respectively.

Furthermore, we study the task of estimating the expectation of an observable $\mathrm{Tr}[A\rho]$ for an observable A while ensuring privacy of a state ρ in Section 6.3. We obtain lower and upper bounds on the number of samples of data are required to achieve a probabilistic accuracy tolerance of estimating the expectation value in Proposition 31 and Proposition 32. When the rest of the parameters except the privacy parameter ε are treated as constants, we show that asymptotically the number of samples to achieve a fixed error tolerance scales as $\Theta\left([(e^{\varepsilon}+1)/(e^{\varepsilon}-1)]^2\right)$ with high probability in Proposition 33.

Lastly, with the use of the contraction bounds derived in Chapter 5, in Proposition 34 we provide stronger characterizations of fairness through QLDP chan-

nels compared to the known result [95, Proposition 14]. We also address the open question of providing bounds on the Holevo information after applying private channels identified in [18], by formally deriving that those private channels satisfy Holevo information stability (see Proposition 35).

6.2 Applications to Private Quantum Hypothesis Testing

In this section, we first review the problem setup of symmetric hypothesis testing without privacy constraints, and then we formulate the private hypothesis testing setup. With the use of tools developed in Section 5.3, we characterize bounds on the sample complexity of the private variant, both in the general and special settings.

6.2.1 Quantum Hypothesis Testing with No Privacy Constraints

Problem setup: Suppose that there are two states ρ and σ, and $\rho^{\otimes n}$ is selected with probability $p \in (0,1)$ and $\sigma^{\otimes n}$ is selected with probability $q = 1 - p$. The sample complexity is equal to the minimum value of n needed to reach a constant error probability in deciding which state was selected. To define this quantity formally, let us recall that the Helstrom–Holevo theorem [59, 65] states that the optimal error probability $p_e(\rho, \sigma, p, q)$ of symmetric quantum hypothesis testing

is as follows:

$$p_e(\rho,\sigma,p,q) \coloneqq \min_{M_1,M_2\geq 0} \{p\,\mathrm{Tr}[M_2\rho] + q\,\mathrm{Tr}[M_1\sigma] : M_1 + M_2 = I\} \tag{6.4}$$

$$= \frac{1}{2}\left(1 - \|p\rho - q\sigma\|_1\right). \tag{6.5}$$

With this in mind, we are assuming in this paradigm that there is a constant $\alpha \in [0,1]$, and our goal is to determine the minimum value of n such that

$$p_e(\rho^{\otimes n},\sigma^{\otimes n},p,q) \coloneqq \frac{1}{2}\left(1 - \left\|p\rho^{\otimes n} - q\sigma^{\otimes n}\right\|_1\right) \leq \alpha. \tag{6.6}$$

To this end, let us define

$$\mathrm{SC}_{(\rho,\sigma)}(\alpha,p,q) \coloneqq \min\{n\in\mathbb{N} : p_e(\rho^{\otimes n},\sigma^{\otimes n},p,q) \leq \alpha\}. \tag{6.7}$$

We also use the shorthand $\mathrm{SC}_{(\rho,\sigma)}$ to refer to this quantity when (α,p,q) are clear from the context.

Theorem 7 (Sample Complexity of Symmetric Hypothesis Testing (Theorem 7 and Corollary 8 of [22])). *Fix the error probability $\alpha \in (0,pq)$. Then for non-orthogonal states ρ and σ,*

$$\max\left\{\frac{\ln(pq/\alpha)}{-\ln F(\rho,\sigma)}, \frac{1-\frac{\alpha(1-\alpha)}{pq}}{[d_{\mathrm{B}}(\rho,\sigma)]^2}\right\} \leq \mathrm{SC}_{(\rho,\sigma)}(\alpha,p,q) \leq \left\lceil \frac{2\ln\left(\frac{\sqrt{pq}}{\alpha}\right)}{-\ln F(\rho,\sigma)} \right\rceil. \tag{6.8}$$

By using the fact $-\ln(\sqrt{x}) \geq 1 - \sqrt{x}$ for all $x > 0$, we have that $-\ln F(\rho,\sigma) \geq [d_B(\rho,\sigma)]^2$. Then, considering α, p, q to be constants and applying Theorem 7, we arrive at

$$\mathrm{SC}_{(\rho,\sigma)}(\alpha,p,q) = \Theta\left(\frac{1}{[d_B(\rho,\sigma)]^2}\right). \tag{6.9}$$

6.2.2 Private Quantum Hypothesis Testing

In Section 6.2.1, we reviewed the notion of sample complexity of symmetric quantum hypothesis testing and recalled bounds on it. Now, we ask the ques-

tion: how will this be affected if the privacy of the quantum states is required? In particular, when we have access to the states $\mathcal{A}(\rho)$ and $\mathcal{A}(\sigma)$, where $\mathcal{A}$ is a private quantum mechanism, how many samples of an unknown privatized state (i.e., $\mathcal{A}(\rho)$ or $\mathcal{A}(\sigma)$) do we need to achieve a fixed error probability α?

In this chapter, we employ QLDP as our privacy metric, which is defined in Definition 15.

Now, we are ready to formally define the sample complexity of private hypothesis testing. Let $\mathcal{A}$ be an ε-QLDP mechanism. Then

$$\mathrm{SC}^{\mathcal{A}}_{(\rho,\sigma)}(\alpha, p, q) \coloneqq \min\{n \in \mathbb{N} : p_e((\mathcal{A}(\rho))^{\otimes n}, (\mathcal{A}(\sigma))^{\otimes n}, p, q) \leq \alpha\}, \tag{6.10}$$

where $p_e(\cdot)$ is defined in (6.5). The sample complexity under ε-QLDP is defined as follows:

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)}(\alpha, p, q) \coloneqq \inf_{\mathcal{A}\in\mathcal{B}^{\varepsilon}} \mathrm{SC}^{\mathcal{A}}_{(\rho,\sigma)}(\alpha, p, q), \tag{6.11}$$

where $\mathcal{B}^{\varepsilon}$ represents all ε-QLDP mechanisms, as defined in (5.29). We also use the shorthand $\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)}$ to refer to the above quantity when (α, p, q) are clear from the context or when they are treated as constants.

Remark 33 (Worst-Case Sample Complexity)**.** *When defining $\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)}$, we infimize over the set of channels, $\mathcal{B}^{\varepsilon}$. This corresponds to the best-case scenario. However, if we consider the worst-case scenario by instead taking the supremum over $\mathcal{B}^{\varepsilon}$, the resulting sample complexity is unbounded. To see this, let us consider a replacement channel $\mathcal{R}$, where the channel replaces all inputs to the channel with a fixed, arbitrary state ω. Then the channel described is indeed private with $\mathcal{R}(\rho) = \mathcal{R}(\sigma) = \omega$. However, in this construction, we cannot distinguish ρ and σ by having access to the privatized data (i.e., $\mathcal{R}(\rho)$ and $\mathcal{R}(\sigma)$), leading to an unbounded sample complexity.*

Next, we provide upper and lower bounds on $\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)}$.

Theorem 8 (Bounds on Private Sample Complexity). *Let $p \in (0,1)$, set $q := 1-p$, and let ρ and σ be states. Fix the error probability $\alpha \in (0, pq)$. For $\varepsilon > 0$, the following holds:*

$$\max\left\{\mathrm{SC}_{(\rho,\sigma)}(\alpha,p,q), \frac{C_{\varepsilon,p,q,\alpha}}{T(\rho,\sigma)}, \frac{\ln(pq/\alpha)e^{\varepsilon}}{2(e^{\varepsilon}-1)^2\left[T(\rho,\sigma)\right]^2}\right\} \leq \mathrm{SC}^{\varepsilon}_{(\rho,\sigma)}(\alpha,p,q) \leq \left\lceil 2\ln\left(\frac{\sqrt{pq}}{\alpha}\right)\left(\frac{(e^{\varepsilon}+1)}{(e^{\varepsilon}-1)T(\rho,\sigma)}\right)^2\right\rceil, \quad (6.12)$$

where $\mathrm{SC}_{(\rho,\sigma)}(\alpha,p,q)$ is defined in (6.7) and

$$C_{\varepsilon,p,q,\alpha} := \max\left\{\frac{\ln(pq/\alpha)(e^{\varepsilon}+1)}{\varepsilon(e^{\varepsilon}-1)}, \frac{\left(1-\frac{\alpha(1-\alpha)}{pq}\right)(e^{\varepsilon}+1)}{2\,(e^{\varepsilon/2}-1)^2}\right\}. \quad (6.13)$$

Proof. First notice that, by applying a QLDP channel $\mathcal{A}$, for $\varepsilon > 0$ on states ρ and σ, it is guaranteed that the states $\mathcal{A}(\rho)$ and $\mathcal{A}(\sigma)$ are not orthogonal. This is due to the fact that orthogonal states can be perfectly distinguished, which indeed violates the privacy requirement if it happens. With this, the current setup of privatized states satisfies the requirements needed to apply Theorem 7.

Upper bound: Notice that due to $-\ln(x) \geq 1-x$ for $x > 0$

$$\frac{1}{-\ln\sqrt{F}(\rho,\sigma)} \leq \frac{2}{\left[d_{\mathrm{B}}(\rho,\sigma)\right]^2} \leq \frac{2}{\left[T(\rho,\sigma)\right]^2}, \quad (6.14)$$

where the last inequality follows from the Fuchs-van-de-Graaf inequalities [47], as below:

$$T(\rho,\sigma) \leq \sqrt{1-F(\rho,\sigma)} \leq \sqrt{2\left(1-\sqrt{F}(\rho,\sigma)\right)} = d_{\mathrm{B}}(\rho,\sigma). \quad (6.15)$$

By fixing a private channel $\mathcal{A}$, consider the distinguishability of the states $\mathcal{A}(\rho)$ and $\mathcal{A}(\sigma)$. To this end, by applying Theorem 7 and (6.14), we see that the sample complexity $\mathrm{SC}^{\mathcal{A}}_{(\rho,\sigma)}$ satisfies the following:

$$\mathrm{SC}^{\mathcal{A}}_{(\rho,\sigma)} \leq \left\lceil \frac{2\ln\left(\frac{\sqrt{pq}}{\alpha}\right)}{\left[T(\mathcal{A}(\rho),\mathcal{A}(\sigma))\right]^2} \right\rceil, \quad (6.16)$$

With that, we have

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} \leq \inf_{\mathcal{A}\in\mathcal{B}^{\varepsilon}} \left\lceil \frac{2\ln\left(\frac{\sqrt{pq}}{\alpha}\right)}{\left[T(\mathcal{A}(\rho),\mathcal{A}(\sigma))\right]^2} \right\rceil, \tag{6.17}$$

This can further be bounded by choosing a specific channel that satisfies ε-QLDP. To this end choose $\mathcal{A} = \mathcal{A}^{p}_{\mathrm{Dep}} \circ \mathcal{M}$ from Proposition 23 by setting $M = \Pi_{+}$, which is the projection onto the positive eigenspace of $\rho - \sigma$. This channel has also been used in the proof of Theorem 5. From there, apply (5.66) to get

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} \leq \left\lceil \frac{2\ln\left(\frac{\sqrt{pq}}{\alpha}\right)}{\left[T((\mathcal{A}^{p}_{\mathrm{Dep}} \circ \mathcal{M})(\rho),(\mathcal{A}^{p}_{\mathrm{Dep}} \circ \mathcal{M})(\sigma))\right]^2} \right\rceil \tag{6.18}$$

$$= \left\lceil 2\ln\left(\frac{\sqrt{pq}}{\alpha}\right)\left(\frac{(e^{\varepsilon}+1)}{(e^{\varepsilon}-1)T(\rho,\sigma)}\right)^2 \right\rceil, \tag{6.19}$$

concluding the proof of the upper bound in (6.12).

Lower bound: To obtain the first part of the lower bound, observe that with the data processing of the channel $\mathcal{A}^{\otimes n}$ and applying [75, Proposition 5.2] we have

$$p_e(\rho^{\otimes n},\sigma^{\otimes n},p,q) \leq p_e((\mathcal{A}(\rho))^{\otimes n},(\mathcal{A}(\sigma))^{\otimes n},p,q)\,. \tag{6.20}$$

Then using the definitions of sample complexity in (6.10), we have that

$$\mathrm{SC}_{(\rho,\sigma)}(\alpha,p,q) \leq \mathrm{SC}^{\mathcal{A}}_{(\rho,\sigma)}(\alpha,p,q). \tag{6.21}$$

Next, optimizing over $\mathcal{A} \in \mathcal{B}^{\varepsilon}$, we arrive at

$$\mathrm{SC}_{(\rho,\sigma)}(\alpha,p,q) \leq \mathrm{SC}^{\varepsilon}_{(\rho,\sigma)}(\alpha,p,q). \tag{6.22}$$

To obtain the second part of the lower bound in (6.12), notice that

$$-\ln F(\rho,\sigma) = \widetilde{D}_{1/2}(\rho\|\sigma) \leq D(\rho\|\sigma), \tag{6.23}$$

where $\widetilde{D}_{1/2}$ is the sandwiched Rényi relative entropy of order $\alpha = 1/2$ [89, 139], and the last inequality follows from the α-monotonicity of the sandwiched

Rényi relative entropy, as well as the fact that $\lim_{\alpha\to 1}\widetilde{D}_\alpha = D$ (see also Proposition 7.28 and Proposition 7.29 in [75]). This leads to the following set of inequalities:

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} \geq \inf_{\mathcal{A}\in\mathcal{B}^{\varepsilon}} \frac{\ln(pq/\alpha)}{-\ln F(\mathcal{A}(\rho),\mathcal{A}(\sigma))} \tag{6.24}$$

$$\geq \inf_{\mathcal{A}\in\mathcal{B}^{\varepsilon}} \frac{\ln(pq/\alpha)}{D(\mathcal{A}(\rho)\|\mathcal{A}(\sigma))} \tag{6.25}$$

$$\geq \ln\left(\frac{pq}{\alpha}\right)\frac{(e^{\varepsilon}+1)}{\varepsilon(e^{\varepsilon}-1)T(\rho,\sigma)}, \tag{6.26}$$

where the last inequality follows from Proposition 25.

The second part of the lower bound follows by using Proposition 24 and the lower bound from Theorem 7 as follows:

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} \geq \inf_{\mathcal{A}\in\mathcal{B}^{\varepsilon}} \frac{pq-\alpha(1-\alpha)}{pq\left[d_{\mathrm{B}}(\mathcal{A}(\rho),\mathcal{A}(\sigma))\right]^2} \tag{6.27}$$

$$\geq \frac{(pq-\alpha(1-\alpha))\,(e^{\varepsilon}+1)}{2pq\ \left(e^{\varepsilon/2}-1\right)^2 T(\rho,\sigma)}. \tag{6.28}$$

To prove the third part of the lower bound, we use that

$$D(\rho\|\sigma) \leq \chi^2(\rho\|\sigma), \tag{6.29}$$

where

$$\chi^2(\rho\|\sigma) := 2\int_1^{\infty} E_\gamma(\rho\|\sigma) + \gamma^{-3}E_\gamma(\sigma\|\rho)\,\mathrm{d}\gamma. \tag{6.30}$$

From Proposition 29, we have that for $\mathcal{A}\in\mathcal{B}^{\varepsilon}$

$$D(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \chi^2(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \tag{6.31}$$

$$\leq 2\left[T(\rho,\sigma)\right]^2 e^{-\varepsilon}(e^{\varepsilon}-1)^2. \tag{6.32}$$

Then, plugging this inequality in (6.25) gives the desired lower bound.

Combining the three lower bounds completes the proof of the lower bound in (6.12). □

Corollary 6 (Sample Complexity for Distinguishing Orthogonal States)**.** *Let ρ and σ be orthogonal quantum states. For $\varepsilon > 0$ and $\alpha \leq pq$, by choosing α, p, q as constants,*

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} = \Theta\left(\left(\frac{e^{\varepsilon}+1}{e^{\varepsilon}-1}\right)^2\right). \tag{6.33}$$

Furthermore, in the high privacy regime (i.e., when $\varepsilon < 1$), we obtain the following:

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} = \Theta\left(\frac{1}{\varepsilon^2}\right). \tag{6.34}$$

Proof. First notice that

$$\frac{(e^{\varepsilon/2}-1)^2}{e^{\varepsilon}+1} \leq \frac{(e^{\varepsilon/2}-1)^2}{e^{\varepsilon}+1} \times \frac{(e^{\varepsilon/2}+1)^2}{e^{\varepsilon}+1} = \left(\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1}\right)^2, \tag{6.35}$$

with the first inequality following from $(e^{\varepsilon/2}+1)^2 \geq e^{\varepsilon}+1$. Then the lower bound in (6.28) simplifies to

$$\mathrm{SC}^{\varepsilon}_{\rho,\sigma)} \geq \frac{(pq-\alpha(1-\alpha))\,(e^{\varepsilon}+1)^2}{2pq\ (e^{\varepsilon}-1)^2\,T(\rho,\sigma)}. \tag{6.36}$$

Together with the upper bound in Theorem 8 and the substitution $T(\rho,\sigma) = 1$ for orthogonal states, we complete the first part of the proof.

The equality in (6.34) follows from applying (6.33), along with the constraint $\varepsilon < 1$, as follows: Firstly, we have

$$\frac{e^{\varepsilon}+1}{e^{\varepsilon}-1} \geq \frac{1}{\varepsilon}, \tag{6.37}$$

by observing that the left-hand side is equal to $(\tanh(\varepsilon/2))^{-1}$ and applying $\tanh(\varepsilon/2) \leq \varepsilon$ for all $\varepsilon > 0$. Also

$$\frac{e^{\varepsilon}+1}{e^{\varepsilon}-1} \leq \frac{4}{\varepsilon}. \tag{6.38}$$

due to $\tanh(\varepsilon/2) \geq \varepsilon/4$ for $\varepsilon \in [0,1]$. This concludes the proof of (6.34). □

Remark 34 (Order Optimality)**.** *In the setting discussed in Corollary 6 for orthogonal states, the ε-QLDP mechanism presented in Proposition 23 is order optimal (i.e., optimal up to constant factors) for all $\varepsilon > 0$.*

Theorem 9 (Sample Complexity in the High-Privacy Regime)**.** *Let ρ and σ be quantum states. For $\varepsilon \in (0,1)$ and $\alpha \leq pq$, by choosing α, p, q as constants, the following holds:*

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} = \Theta\left(\frac{1}{\varepsilon^2 \left[T(\rho,\sigma)\right]^2}\right). \tag{6.39}$$

Proof. Recall from Theorem 8

$$\frac{\ln(pq/\alpha)e^{\varepsilon}}{2(e^{\varepsilon}-1)^2 \left[T(\rho,\sigma)\right]^2} \leq \mathrm{SC}^{\varepsilon}_{(\rho,\sigma)}(\alpha,p,q) \leq \left\lceil 2\ln\left(\frac{\sqrt{pq}}{\alpha}\right)\left(\frac{(e^{\varepsilon}+1)}{2(e^{\varepsilon}-1)T(\rho,\sigma)}\right)^2\right\rceil. \tag{6.40}$$

By choosing α, p, q as constants, together with (6.38) we obtain that

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} \leq C_1 \frac{1}{\varepsilon^2 \left[T(\rho,\sigma)\right]^2}, \tag{6.41}$$

for a constant C_1.

For the lower bound, consider that for $\varepsilon < 1$

$$\frac{e^{\varepsilon}}{(e^{\varepsilon}-1)^2} \geq \frac{1}{2\varepsilon^2}, \tag{6.42}$$

which follows since the function $g(\varepsilon) \coloneqq 2\varepsilon^2 e^{\varepsilon} - (e^{\varepsilon}-1)^2 > 0$ for $\varepsilon \in (0,1)$. This can be verified since the first derivative $g'(\varepsilon) = 2e^{\varepsilon}\left((\varepsilon+1)^2 - e^{\varepsilon}\right) \geq 0$ for $\varepsilon \in (0,1)$ leading to $g(\varepsilon)$ being a non-decreasing function of ε, so as to arrive at

$$g(\varepsilon) \geq g(0) = 0 \tag{6.43}$$

for $\varepsilon \in (0,1)$. Then, by combining the lower bound in (6.40) along with (6.42) we have

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} \geq C_2 \frac{1}{\varepsilon^2 \left[T(\rho,\sigma)\right]^2}, \tag{6.44}$$

which concludes the proof together with (6.41).

□

Remark 35 (Cost of Privacy)**.** *For orthogonal states, we need only one sample of the unknown state to declare whether it is ρ or σ when we have access to non-privatized samples of ρ or σ. However, to achieve ε-QLDP, the input states are privatized by applying the channel $\mathcal{A}$, so that we have access to $\mathcal{A}(\rho)$ and $\mathcal{A}(\sigma)$ only. To this end, we need samples on the order of $\Theta\Big(((e^{\varepsilon}+1)/(e^{\varepsilon}-1))^2\Big)$ for all $\varepsilon > 0$, which is apparent from Corollary 6.*

For states ρ and σ, when the privacy level is demanded by $\varepsilon \in (0,1)$, we need samples on the order of $\Theta\left(\frac{1}{\varepsilon^2\left[T(\rho,\sigma)\right]^2}\right)$ to achieve a fixed non-zero error in distinguishing between the two states, which can be deduced from Theorem 9.

In addition, by [95, Proposition 10], for an ε-QLDP mechanism $\mathcal{A}$ and for all states ρ and σ, we have

$$D(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \min\left\{\varepsilon, \frac{\varepsilon^2}{2}\right\}. \tag{6.45}$$

Applying that in the proof of the lower bound of Theorem 8 (in particular, in (6.25)), it follows that, for all distinct states ρ and σ, we require

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} = \Omega\left(\frac{1}{\min\{\varepsilon, \varepsilon^2\}}\right), \tag{6.46}$$

in order to ensure ε-QLDP, where $\varepsilon > 0$.

Remark 35 discusses the cost of privacy when we have access to privatized data samples instead of original data. Next, we show that when certain conditions are met, the impact of privacy on the sample complexity is not significant and is comparable to the non-private sample complexity. This also provides a sense of how large the privacy parameter ε should be in order to achieve a similar sample complexity as in the non-private case, where we do not need to pay an extra cost to ensure privacy, $d_B(\rho,\sigma) > 1$.

Proposition 30 (Sample Complexity in the Low-Privacy Regime)**.** *Let ρ and σ be qubit states, and let $\varepsilon > 0$. If*

$$\left(\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1}\right)^2 \geq \frac{1}{\left[d_B(\rho,\sigma)\right]^2}, \tag{6.47}$$

then

$$\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} = \Theta\left(\frac{1}{\left[d_B(\rho,\sigma)\right]^2}\right). \tag{6.48}$$

Furthermore, for ρ and σ general quantum states, if

$$\left(\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1}\right)^2 \geq \frac{1}{\left[d_B(\rho,\sigma)\right]^2}, \tag{6.49}$$

then

$$\Omega\left(\frac{1}{\left[d_B(\rho,\sigma)\right]^2}\right) \leq \mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} \leq O\left(\frac{L_{k,k'}}{\left[d_B(\rho,\sigma)\right]^2}\right), \tag{6.50}$$

where

$$L_{k,k'} := \left(\max\left\{1, \frac{\min\{k,k'\}}{2}\right\}\right)^2, \tag{6.51}$$

$k' := \ln\left(4/\left[d_B(\rho,\sigma)\right]^2\right)$, and k is the number of distinct eigenvalues of the operator $\rho\#\sigma^{-1}$, as defined in Eq. (C.84).

Proof. See Appendix C.4. □

Note that from Proposition 30, when ρ and σ are qubits, the order optimality in sample complexity is achieved by the privatization mechanism presented in Proposition 23 by choosing the measurement channel to correspond to the measurement in the eigenbasis of $\rho\#\sigma^{-1}$.

Remark 36 (Different Private Mechanisms)**.** *Previously, we considered the setting in which the same private mechanism $\mathcal{A} \in \mathcal{B}^{\varepsilon}$ is applied to each sample of ρ or σ. We can also define another variant in which different mechanisms are applied at each stage;*

namely, $\mathcal{A}_i \in \mathcal{B}^\varepsilon$ for all $i \in \{1, \dots, n\}$. Then, we define the associated sample complexity as follows:

$$\widetilde{\mathrm{SC}}^{\varepsilon}_{(\rho,\sigma)} := \inf_{\substack{\mathcal{A}_i \in \mathcal{B}^\varepsilon \\ \forall i \in \{1,\dots,n\}}} \min\{n \in \mathbb{N} : p_e(\mathcal{A}_1(\rho) \otimes \cdots \otimes \mathcal{A}_n(\rho), \mathcal{A}_1(\sigma) \otimes \cdots \otimes \mathcal{A}_n(\sigma), p, q) \leq \alpha\}, \tag{6.52}$$

By choosing each $\mathcal{A}_i$ to be the same mechanism, we conclude the following inequality:

$$\widetilde{\mathrm{SC}}^{\varepsilon}_{(\rho,\sigma)} \leq \mathrm{SC}^{\varepsilon}_{(\rho,\sigma)}. \tag{6.53}$$

With that, the upper bound given in Theorem 8 still holds. However, the lower bound may not hold. To this end, we prove the following: for $\alpha \leq \min\{p, q\}$

$$\widetilde{\mathrm{SC}}^{\varepsilon}_{(\rho,\sigma)} \geq 2\left(1 - \frac{\alpha}{\min\{p,q\}}\right)^2 \frac{(e^\varepsilon + 1)}{\varepsilon(e^\varepsilon - 1)T(\rho,\sigma)}. \tag{6.54}$$

The proof is presented in Appendix C.5.[1]

Under this setting, also for orthogonal states ρ and σ and choosing α, p, q as constants, when $\varepsilon < 1$, we arrive at

$$\widetilde{\mathrm{SC}}^{\varepsilon}_{(\rho,\sigma)} = \Theta\left(\frac{1}{\varepsilon^2}\right), \tag{6.55}$$

implying that having different mechanisms may not improve the sample complexity for this case.

Above, we considered $\mathcal{A}_i$ chosen independently of $\mathcal{A}_j$, where $i \neq j$. It is an interesting question for future work to understand whether adaptive strategies can decrease the sample complexity or number of copies of the privatized states required to minimize the cost of privacy.

Remark 37 (Private Multiple Hypothesis Testing)**.** *The sample complexity of private multiple hypothesis testing for a tuple of states $(\rho_m)_{m=1}^M$ with the prior probabilities*

[1] Note that this lower bound also holds in the case where we consider the same private channel for every copy of the state.

$(p_m)_{m=1}^M$ *to achieve at most* α *error probability is defined as*

$$\mathrm{SC}^{\varepsilon}_{(\rho_i)_{i=1}^M} := \inf_{\mathcal{A}\in\mathcal{B}^{\varepsilon}} \left\{ n\in\mathbb{N} : \inf_{(\Lambda_1^{(n)},\ldots,\Lambda_M^{(n)})} \sum_{m=1}^M p_m \mathrm{Tr}\left[(I^{\otimes n} - \Lambda_m^{(n)}) \left(\mathcal{A}(\rho_m)\right)^{\otimes n} \right] \leq \alpha \right\}, \tag{6.56}$$

where $\Lambda_1^{(n)},\ldots,\Lambda_M^{(n)} \geq 0$ *and* $\sum_{m=1}^M \Lambda_m^{(n)} = I^{\otimes n}$*. Then, with the use of tools from the proof of Theorem 8 and [22, Theorem 11], we arrive at*

$$\max_{m\neq\tilde{m}} \frac{\ln\left(\frac{p_m p_{\tilde{m}}}{(p_m+p_{\tilde{m}})\alpha}\right)(e^{\varepsilon}+1)}{\varepsilon(e^{\varepsilon}-1)T(\rho_m,\rho_{\tilde{m}})} \leq$$
$$\mathrm{SC}^{\varepsilon}_{(\rho_i)_{i=1}^M} \leq \left\lceil \max_{m\neq\tilde{m}} 2\ln\left(\frac{M(M-1)\sqrt{p_m}\sqrt{p_{\tilde{m}}}}{2\alpha}\right)\left(\frac{(e^{\varepsilon}+1)}{(e^{\varepsilon}-1)T(\rho_m,\rho_{\tilde{m}})}\right)^2 \right\rceil. \tag{6.57}$$

Remark 38 (Private Asymmetric Hypothesis Testing)**.** *We define the private asymmetric hypothesis testing sample complexity as follows: Let* ρ *and* σ *be states with* $\alpha_1,\alpha_2\in(0,1)$ *and* $\varepsilon\geq 0$

$$n^*(\rho,\sigma,\varepsilon,\alpha_1,\alpha_2) = \inf_{\mathcal{A}\in\mathcal{B}^{\varepsilon}} \inf_{\Lambda^{(n)}} \left\{ \begin{array}{l} n\in\mathbb{N} : \mathrm{Tr}[(I^{\otimes n}-\Lambda^{(n)})\mathcal{A}(\rho)^{\otimes n}] \leq \alpha_1, \\ \mathrm{Tr}[\Lambda^{(n)}\mathcal{A}(\sigma)^{\otimes n}] \leq \alpha_2,\ 0\leq\Lambda^{(n)}\leq I^{\otimes n} \end{array} \right\}, \tag{6.58}$$

where $\mathcal{B}^{\varepsilon}$ *is defined in* (5.29)*. A lower bound on* $n^*(\rho,\sigma,\varepsilon,\alpha_1,\alpha_2)$ *is as follows:*

$$n^*(\rho,\sigma,\varepsilon,\alpha_1,\alpha_2) \geq \max\left\{ \sup_{\beta>1}\left(\frac{\ln\left(\frac{(1-\alpha_1)^{\beta'}}{\alpha_2}\right)}{\min\{\varepsilon,\varepsilon^2\beta/2\}}\right), \sup_{\beta>1}\left(\frac{\ln\left(\frac{(1-\alpha_2)^{\beta'}}{\alpha_1}\right)}{\min\{\varepsilon,\varepsilon^2\beta/2\}}\right)\right\}, \tag{6.59}$$

where $\beta' := \frac{\beta}{\beta-1}$*. The proof follows by applying [95, Proposition 9] together with the lower bound presented in [22, Theorem 9]. In particular, we use the following inequality: Let* $\mathcal{A}\in\mathcal{B}^{\varepsilon}$ *and* $\beta>1$

$$\widetilde{D}_{\beta}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \min\left\{\varepsilon, \frac{\varepsilon^2\beta}{2}\right\} < +\infty, \tag{6.60}$$

where $\widetilde{D}_{\beta}(\omega\|\tau) := \frac{1}{\beta-1}\ln\mathrm{Tr}[(\tau^{(1-\beta)/2\beta}\omega\tau^{(1-\beta)/2\beta})^{\beta}]$ *is the sandwiched Rényi relative entropy of order* β *[89, 139].*

Remark 39 (Private Quantum Hypothesis Testing under (ε,δ)-QLDP)**.** *In Theorem 8, we provide bounds on the sample complexity of private quantum hypothesis*

testing where privacy constraints consist of ε-QLDP with $\delta = 0$. For the case where $\delta \geq 0$, the upper bound therein can be generalized as follows:

$$\inf_{\mathcal{A}\in\mathcal{B}^{\varepsilon,\delta}} \mathrm{SC}^{\mathcal{A}}_{(\rho,\sigma)}(\alpha, p, q) \leq \left\lceil 2\ln\left(\frac{\sqrt{pq}}{\alpha}\right)\left(\frac{(e^{\varepsilon}+1)}{(e^{\varepsilon}-1+2\delta)T(\rho,\sigma)}\right)^2 \right\rceil. \tag{6.61}$$

This follows by the use of the mechanism introduced in the achievability part of the proof of Theorem 6, together with similar techniques used in the proof of the upper bound in Theorem 8.

We leave the general characterization of sample complexity of hypothesis testing when privacy is imposed by (ε, δ)-QLDP for future work. The main technical tools needed to obtain the lower bounds as given in Theorem 8 are the characterization of the contraction of quantum relative entropy and Bures distance under the privacy constraints imposed by (ε, δ)-QLDP. In Appendix C.6, we provide some tools that may be of relevance to accomplish this goal. In particular, we prove several connections between ε-QLDP and (ε, δ)-QLDP mechanisms, which may potentially be used to infer about the impact of (ε, δ)-QLDP constraints using the formal guarantees derived for ε-QLDP in Theorem 8.

6.3 Learning with Privacy

In this section, we study how well one can do learning and inference of properties of quantum data while ensuring the privacy of the data.

Here we focus on the following setup where one needs to estimate $\mathrm{Tr}[O\rho]$ for an observable $O \in \mathcal{O}$ and a quantum state ρ. However, we also need to ensure the privacy of the quantum state. With privatization, we have access to $\mathcal{A}(\rho)$, where $\mathcal{A}$ is a private channel that satisfies Definition 15. The task is to estimate

$\mathrm{Tr}[O\rho]$ with a minimum number of samples from the privatized state $\mathcal{A}(\rho)$.

Next, we formally define the task as follows: Let $\mathcal{A} \in \mathcal{B}^{\varepsilon}$. Let $\hat{E}(\mathcal{A}, \rho, O, n)$ be the estimate for $\mathrm{Tr}[O\rho]$ obtained by a procedure $\mathcal{P}_O$ that takes $(\mathcal{A}(\rho))^{\otimes n}$ as input such that the following condition is satisfied:

$$\left|\mathrm{Tr}[O\rho] - \hat{E}(\mathcal{A}, \rho, O, n)\right| \leq \beta \tag{6.62}$$

with probability at least $1 - \delta$, for $\beta > 0$ and $\delta \in (0, 1)$. We define the optimal sample complexity for estimating the expectation value while ensuring ε-QLDP as follows: Let $O \in \mathcal{O}$ be an observable and let $\mathcal{S}$ be the set of quantum states of interest. Then, the optimal sample complexity is defined as

$$n^*(\mathcal{S}, O, \beta, \delta, \varepsilon) \coloneqq \inf_{\mathcal{A} \in \mathcal{B}^{\varepsilon}, \mathcal{P}_O} \left\{ n \in \mathbb{N} : \left|\mathrm{Tr}[O\rho] - \hat{E}(\mathcal{A}, \rho, O, n)\right| \leq \beta \text{ w.p. } 1 - \delta \; \forall \rho \in \mathcal{S} \right\}. \tag{6.63}$$

We next prove upper and lower bounds on the sample complexity of inference by utilizing the contraction bounds on quantum divergences under privacy constraints developed in Chapter 5. To this end, we consider the following assumptions on the set of observables $\mathcal{O}$ and the set of states $\mathcal{S}$: We choose $O \in \mathcal{O}$ to be a measurement operator such that $0 \leq O \leq I$ to arrive at the upper bound. For the lower bound, we assume that there exist two states $\rho, \sigma \in \mathcal{S}$ such that those two states can be differentiated by the outcomes of at least one of the observables in the set $\mathcal{O}$. We leave the expansion of the study to a more general set of observables and states for future work.

Proposition 31 (Sample Complexity Lower Bound with QLDP)**.** *let $\varepsilon, \beta > 0$ and $\delta \in (0, 1/2)$. If $\exists \rho, \sigma \in \mathcal{S}$ such that $\max_{O \in \mathcal{O}} |\mathrm{Tr}[O(\rho - \sigma)]| \geq 2\beta$, then, we have that*

$$n^*(\mathcal{S}, O, \beta, \delta, \varepsilon) \geq \max\left\{ \frac{(1 - 2\delta)(e^{\varepsilon} + 1)}{(e^{\varepsilon} - 1)}, \frac{(1 - 2\delta)^2 (e^{\varepsilon} + 1)}{2(e^{\varepsilon/2} - 1)^2} \right\}, \tag{6.64}$$

where $n^\star$ is defined in (6.63).

Proof. To prove lower bounds, we first use the same strategy as used in the proof of [123, Theorem 1]. Fix ρ, σ and $O \in \mathcal{O}$ such that $\mathrm{Tr}[O(\rho - \sigma)]| \geq 2\beta$. Let $\mathcal{A} \in \mathcal{B}^{\varepsilon}$. Also identify that $\mathcal{P}_O$ in this case is a quantum channel (quantum to classical) that takes n samples of the privatized state($\mathcal{A}(\rho)$ or $\mathcal{A}(\sigma)$) and outputs a classical value that refers to the random estimate $\hat{E}(\mathcal{A}, \omega, O, n)$ with $\omega \in \{\rho, \sigma\}$ with the underlying probability distribution being $f_{O,\omega}$. This means that we have

$$\mathcal{P}_O((\mathcal{A}(\omega))^{\otimes n}) = \int_a f_{O,\omega}(a)|a\rangle\langle a| \, \mathrm{d}a, \tag{6.65}$$

where $\{a\}_a$ refers to the classical estimates for $\mathrm{Tr}[O\omega]$. With this setup and recalling the condition in (6.62) together with [123, Eq. (A4)] (note that ε therein refers to δ in our setup), we have that

$$T(\mathcal{P}_O((\mathcal{A}(\rho))^{\otimes n}), \mathcal{P}_O((\mathcal{A}(\sigma))^{\otimes n})) \geq 1 - 2\delta. \tag{6.66}$$

Next, by applying the data-processing property of trace distance, we have

$$\begin{aligned} 1 - 2\delta &\leq T((\mathcal{A}(\rho))^{\otimes n}, (\mathcal{A}(\sigma))^{\otimes n}) && (6.67)\\ &\leq nT(\mathcal{A}(\rho), \mathcal{A}(\sigma)) && (6.68)\\ &\leq n\left(\frac{e^{\varepsilon} - 1}{e^{\varepsilon} + 1}\right) T(\rho, \sigma) && (6.69)\\ &\leq n\left(\frac{e^{\varepsilon} - 1}{e^{\varepsilon} + 1}\right), && (6.70) \end{aligned}$$

where the first inequality follows from the sub-additivity of trace distance; the second inequality from $\mathcal{A} \in \mathcal{B}^{\varepsilon}$ and Theorem 5; and finally the last inequality from $T(\rho, \sigma) \leq 1$ for two states ρ and σ.

Since (6.70) holds as long as there exists $\rho, \sigma \in \mathcal{S}$ such that $\max_{O \in \mathcal{O}} |\mathrm{Tr}[O(\rho - \sigma)]| \geq 2\beta$, by rearranging the terms of (6.70), we arrive at the first lower bound.

For the second lower bound, we consider

$$1 - 2\delta \leq T((\mathcal{A}(\rho))^{\otimes n}, (\mathcal{A}(\sigma))^{\otimes n}) \tag{6.71}$$

together with

$$[T(\rho,\sigma)]^2 \leq 1 - F(\rho,\sigma) \tag{6.72}$$

$$= \left(1 - \sqrt{F(\rho,\sigma)}\right)\left(1 + \sqrt{F(\rho,\sigma)}\right) \tag{6.73}$$

$$\leq 2\left(1 - \sqrt{F(\rho,\sigma)}\right) \tag{6.74}$$

$$= [d_B(\rho,\sigma)]^2, \tag{6.75}$$

where $F(\rho,\sigma)$ is defined in (2.4) and $[d_B(\rho,\sigma)]^2$ in (2.5).

This leads to

$$(1-2\delta)^2 \leq [T((\mathcal{A}(\rho))^{\otimes n}, (\mathcal{A}(\sigma))^{\otimes n})]^2 \tag{6.76}$$

$$\leq [d_B((\mathcal{A}(\rho))^{\otimes n}, (\mathcal{A}(\sigma))^{\otimes n})]^2 \tag{6.77}$$

$$= 2\frac{[d_B((\mathcal{A}(\rho))^{\otimes n}, (\mathcal{A}(\sigma))^{\otimes n})]^2}{2} \tag{6.78}$$

$$= 2\left(1 - \left(1 - \frac{[d_B((\mathcal{A}(\rho))^{\otimes n}, (\mathcal{A}(\sigma))^{\otimes n})]^2}{2}\right)\right) \tag{6.79}$$

$$= 2\left(1 - \left(1 - \frac{[d_{\mathrm{B}}(\mathcal{A}(\rho), \mathcal{A}(\sigma))]^2}{2}\right)^n\right) \tag{6.80}$$

$$\leq n\,[d_{\mathrm{B}}(\mathcal{A}(\rho), \mathcal{A}(\sigma))]^2 \tag{6.81}$$

$$\leq 2n\frac{(e^{\varepsilon/2}-1)^2}{e^{\varepsilon}+1}, \tag{6.82}$$

where the third equality holds due to multiplicativity of fidelity; penultimate inequality since $(1-x)^n \geq 1 - nx$ for $x \in [0,1]$ for $n \in \mathbb{N}$; and the last inequality by Proposition 24 for $\mathcal{A} \in \mathcal{B}^{\varepsilon}$ together with $T(\rho,\sigma) \leq 1$.

By rearranging the terms, we conclude the proof of the second lower bound.

□

Next, we obtain an upper bound on the sample complexity to estimate the expected value of an observable with respect to a quantum state $\rho \in \mathcal{S}$ by devising a protocol with the use of a depolarization mechanism. To this end, we

consider the scenario where observable O satisfies $0 \leq O \leq I$ with the assumption that we have access to $\mathrm{Tr}[O]$. This is the special case of the observable being a measurement operator. Furthermore, for a Pauli operator P one can choose $O = (I + P)/2$ so that the condition $0 \leq O \leq I$ holds.

Proposition 32 (Sample Complexity Upper Bound with QLDP)**.** *let $\varepsilon, \beta > 0$ and $\delta \in (0, 1)$. Then, for $O \in \mathcal{O}$ observable such that $0 \leq O \leq I$, we have that*

$$n^*(\mathcal{S}, O, \beta, \delta, \varepsilon) \leq \left\lceil \frac{1}{2\beta^2} \ln\left(\frac{2}{\delta}\right)\left(\frac{e^\varepsilon - 1 + d}{e^\varepsilon - 1}\right)^2 \right\rceil, \tag{6.83}$$

where d is the dimension of the Hilbert space where states of $\mathcal{S}$ belong to and $n^\star$ is defined in (6.63).

Proof. We obtain the desired upper bound by showing that there exists a protocol $\mathcal{P}_O$ to estimate $\mathrm{Tr}[O\rho]$ for all $\rho \in \mathcal{S}$ given access to n copies of $\mathcal{A}(\rho)$ for a particular ε-QLDP channel $\mathcal{A}$.

We fix $\mathcal{A} = \mathcal{A}^p_{\mathrm{dep}}$ to be the depolarizing mechanism in (3.84) with

$$p \coloneqq \frac{d}{e^\varepsilon + d - 1}. \tag{6.84}$$

By (3.103), we have $\mathcal{A} \in \mathcal{B}^\varepsilon$. With that choice, consider

$$\mathrm{Tr}[O\mathcal{A}(\rho)] = (1 - p)\mathrm{Tr}[O\rho] + \frac{p}{d}\mathrm{Tr}[O] \tag{6.85}$$

$$= (1 - p)\mathrm{Tr}[O\rho] + \frac{p}{d}\mathrm{Tr}[O]. \tag{6.86}$$

Now, consider the following protocol where each copy of $\mathcal{A}(\rho)$ passes through the measurement channel formed by O such that $0 \leq O \leq I$. Note that we do not consider the complexity of implementing the measurement channel. Measurement channel refers to

$$\mathcal{M}_O(\omega) \coloneqq \mathrm{Tr}[O\omega]\, |1\rangle\langle 1| + \mathrm{Tr}[(I - O)\omega]\, |0\rangle\langle 0|. \tag{6.87}$$

We pass all copies of the state $\mathcal{A}(\rho)$ independently through $\mathcal{M}_O$ and record the classical outcomes $\{X_i\}_{i=1}^n$ (note that these outcomes are i.i.d. random variables due to the randomness of the measurement channel and the channel being applied independently on each copy). We use the following as the estimator

$$\hat{E}(\mathcal{A}, \rho, O, n) = \frac{1}{(1-p)} \frac{1}{n} \sum_{i=1}^{n} X_i - \frac{p}{(1-p)d} \mathrm{Tr}[O]. \tag{6.88}$$

Also, observe that $\mathbb{E}[X_i] = (1-p)\mathrm{Tr}[O\rho] + \frac{p}{d}\mathrm{Tr}[O]$.

With this estimator, let us consider the following:

$$\begin{aligned}
&\Pr\Big(\Big|\hat{E}(\mathcal{A}, \rho, O, n) - \mathrm{Tr}[O\rho]\Big| \geq \beta\Big) \\
&= \Pr\Big(\Big|(1-p)\hat{E}(\mathcal{A}, \rho, O, n) - (1-p)\mathrm{Tr}[O\rho]\Big| \geq (1-p)\beta\Big) && (6.89) \\
&= \Pr\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - (1-p)\mathrm{Tr}[O\rho] - \frac{p}{d}\mathrm{Tr}[O]\right| \geq (1-p)\beta\right) && (6.90) \\
&= \Pr\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}[X_i]\right| \geq (1-p)\beta\right) && (6.91) \\
&\leq 2\exp\Big(-2n(1-p)^2\beta^2\Big), && (6.92)
\end{aligned}$$

where the last inequality follows from Hoeffding's inequality by using $0 \leq X_i \leq 1$ since $X_i \in \{0, 1\}$.

Then, choosing n such that $2\exp\Big(-2n(1-p)^2\beta^2\Big) \leq \delta$, we obtain $\Big|\hat{E}(\mathcal{A}, \rho, O, n) - \mathrm{Tr}[O\rho]\Big| \leq \beta$ with probability at least $1-\delta$. Thus, we need

$$n \geq \frac{1}{2\beta^2} \ln\left(\frac{2}{\delta}\right)\left(\frac{e^{\varepsilon} - 1 + d}{e^{\varepsilon} - 1}\right)^2. \tag{6.93}$$

To this end, by choosing

$$n = \left\lceil \frac{1}{2\beta^2} \ln\left(\frac{2}{\delta}\right)\left(\frac{e^{\varepsilon} - 1 + d}{e^{\varepsilon} - 1}\right)^2 \right\rceil, \tag{6.94}$$

we arrive at the desired success criterion, concluding the proof. □

Proposition 33 (Cost of Privacy)**.** *Let $\varepsilon > 0$. Let $O \in \mathcal{O}$ be an observable such that $0 \leq O \leq I$. Also assume that there exists $\rho, \sigma \in \mathcal{S}$ such that $\max_{O \in \mathcal{O}} |\mathrm{Tr}[O(\rho - \sigma)]| \geq 2\beta$. Then, by considering β, δ, and d the dimension of the respective Hilbert space where states in $\mathcal{S}$ lie as constants, we have that*

$$n^*(\mathcal{S}, O, \beta, \delta, \varepsilon) = \Theta\left(\left(\frac{e^\varepsilon + 1}{e^\varepsilon - 1}\right)^2\right), \tag{6.95}$$

where $n^\star$ is defined in (6.63).

Proof. For the lower bound, consider

$$\frac{(e^{\varepsilon/2} - 1)^2}{e^\varepsilon + 1} \leq \frac{(e^{\varepsilon/2} - 1)^2}{e^\varepsilon + 1} \times \frac{(e^{\varepsilon/2} + 1)^2}{e^\varepsilon + 1} = \left(\frac{e^\varepsilon - 1}{e^\varepsilon + 1}\right)^2, \tag{6.96}$$

with the first inequality following from $(e^{\varepsilon/2} + 1)^2 \geq e^\varepsilon + 1$. Then, with the second lower bound in Proposition 31, we obtain that there is a constant C_1 depending on δ such that

$$n^*(\mathcal{S}, O, \beta, \delta, \varepsilon) \geq C_1 \left(\frac{e^\varepsilon + 1}{e^\varepsilon - 1}\right)^2. \tag{6.97}$$

For the upper bound, consider that

$$\left(\frac{e^\varepsilon - 1 + d}{e^\varepsilon - 1}\right)^2 = \left(\frac{e^\varepsilon + 1}{e^\varepsilon - 1} + \frac{d - 2}{e^\varepsilon - 1}\right)^2 \tag{6.98}$$

$$\leq \left(\frac{e^\varepsilon + 1}{e^\varepsilon - 1} + \frac{(d - 2)(e^\varepsilon + 1)}{e^\varepsilon - 1}\right)^2 \tag{6.99}$$

$$= (d - 1)^2 \left(\frac{e^\varepsilon + 1}{e^\varepsilon - 1}\right)^2. \tag{6.100}$$

With the above together with Proposition 32, we obtain that there exists a constant C_2 depending on d, β, δ such that

$$n^*(\mathcal{S}, O, \beta, \delta, \varepsilon) \leq C_2 \left(\frac{e^\varepsilon + 1}{e^\varepsilon - 1}\right)^2. \tag{6.101}$$

Combining (6.97) and (6.101), we conclude the proof. □

6.4 Other Applications

In this section, we show how the results presented in Section 5.3 find use in several applications, including quantum fairness and learning settings, in which we provide formal guarantees on the level of fairness and stability of learning algorithms.

6.4.1 Quantum Fairness through QLDP

Classical decision models, including machine learning models, are prone to discriminating against individuals based on different characteristics, for example, skin color or gender [45]. This has even led to legal mandates of ensuring fairness. With the introduction of quantum machine learning models, there is also a risk of whether fairness will be ensured for both classical and quantum data fed into these algorithms. In [55], it was shown that noisy quantum algorithms can improve fairness. Since fairness and privacy are related domains, it is vital to understand the impact of private algorithms on ensuring fairness as well.

Formally, quantum fairness aims to treat all input states equally, meaning that all pairs of input states that are close in some distance metric (e.g., close in normalized trace distance) should yield similar outcomes when processed by a quantum channel [55]. Define $\mathcal{A} \coloneqq \mathcal{M} \circ \mathcal{E}$, which is a quantum-to-classical channel where a quantum channel $\mathcal{E}$ is followed by a measurement channel comprised of a POVM $\{M_i\}_{i \in O}$. With that, quantum fairness is defined in [55] as follows.

Definition 16 ((α, β)-Fairness [55])**.** *Let $\mathcal{A} = \mathcal{M} \circ \mathcal{E}$, and let $\hat{D}(\cdot\|\cdot)$ and $d(\cdot\|\cdot)$ be*

distance metrics on $\mathcal{D}(\mathcal{H})$ and $\mathcal{D}(\mathcal{O})$, respectively. Fix $0 < \alpha, \beta \leq 1$. Then the decision model $\mathcal{A}$ is (α, β) fair if for all $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ such that $\hat{D}(\rho \| \sigma) \leq \alpha$,

$$d(\mathcal{A}(\rho) \| \mathcal{A}(\sigma)) \leq \beta. \tag{6.102}$$

By choosing $\hat{D}$ to be the normalized trace distance and $d(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = \frac{1}{2}\sum_i \left|\mathrm{Tr}[M_i \mathcal{E}(\rho - \sigma)]\right|$, Proposition 34 states that $\mathcal{A}$ satisfying ε-QPP implies that it is also $(\alpha, \sqrt{\varepsilon'/2})$-fair, where $\varepsilon' = \min\{\varepsilon, \varepsilon^2/2\}$. Note that for the case $\varepsilon > 1$, the above bound could be weak. With Theorem 5, we improve the existing bound on achievable fairness through ε-QLDP mechanisms (for a special class of QPP), which is applicable for all $\varepsilon \geq 0$, and we also extend them to cases for which $\delta \geq 0$.

Proposition 34 (Fairness Guarantee from (ε, δ)-QLDP)**.** *Suppose that $\hat{D}(\rho\|\sigma) = \frac{1}{2}\|\rho - \sigma\|_1$ and $d(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = \frac{1}{2}\sum_i \left|\mathrm{Tr}[M_i \mathcal{E}(\rho - \sigma)]\right|$. If $\mathcal{E}$ satisfies (ε, δ)-QLDP, then $\mathcal{A} = \{\mathcal{E}, \{M_i\}_{i \in O}\}$ is $(\alpha, \varepsilon'(\alpha))$-fair for all $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ such that $\hat{D}(\rho\|\sigma) \leq \alpha$, where*

$$\varepsilon'(\alpha) \coloneqq \alpha \left(\frac{e^\varepsilon - 1 + 2\delta}{e^\varepsilon + 1} \right). \tag{6.103}$$

Proof. From Theorem 6, the channel $\mathcal{E}$ being (ε, δ)-QLDP implies that

$$\frac{1}{2}\|\mathcal{E}(\rho) - \mathcal{E}(\sigma)\|_1 \leq \frac{e^\varepsilon - 1 + 2\delta}{e^\varepsilon + 1} \left(\frac{1}{2}\|\rho - \sigma\|_1 \right) \tag{6.104}$$

$$\leq \alpha \frac{e^\varepsilon - 1 + 2\delta}{e^\varepsilon + 1}, \tag{6.105}$$

where the last inequality follows from the assumption $\hat{D}(\rho\|\sigma) \leq \alpha$.

Then consider the measurement channel that performs the following transformation:

$$\mathcal{E}(\rho) \to \sum_{i \in O} \mathrm{Tr}[M_i \mathcal{E}(\rho)] \, |i\rangle\langle i|. \tag{6.106}$$

It follows from the data-processing inequality for the trace distance that

$$\frac{1}{2}\left\|\sum_{i\in O}\left(\mathrm{Tr}[M_i\mathcal{E}(\rho)]-\mathrm{Tr}[M_i\mathcal{E}(\sigma)]\right)|i\rangle\langle i|\right\|_1 \leq \alpha\frac{e^{\varepsilon}-1+2\delta}{e^{\varepsilon}+1}. \tag{6.107}$$

This leads to

$$d(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = \frac{1}{2}\sum_i\left|\mathrm{Tr}[M_i\mathcal{E}(\rho-\sigma)]\right| \leq \alpha\frac{e^{\varepsilon}-1+2\delta}{e^{\varepsilon}+1}, \tag{6.108}$$

concluding the proof. □

6.4.2 Stability for Quantum Learning through Private Channels

Designing learning algorithms that also ensure privacy for input data is of importance, and it has been widely studied in the classical setting (for example, see [73]). A learning algorithm is known to be stable if its output does not depend too much on any individual training data [14, 113]. It is also known that stability and generalization of a classical learning algorithm to new inputs are closely related, where stability implies generalization using information-theoretic tools [113]. Classical differentially private learners were proved to generalize well [113, Theorem 5]. In [94, Proposition 6], it was shown that learners satisfying a mutual-information-based variant of differential privacy also satisfy algorithmic stability, and hence generalize well for new data.

It is a natural research question to explore whether quantum private learners also provide stability and generalization in quantum learning settings. In [18], it was shown that the generalization error of classical-quantum learners is bounded from above by a function of mutual information between input space

and output space, and a Holevo information term (see Theorem 1 therein for an example). To this end, the authors of [18] showed in Eqs. (5.7.1)–(5.7.6) therein that the mutual information terms decay fast for learners comprised of ε-QLDP channels and left the analysis of the Holevo information term as an open question. In this section, we show that ε-QLDP learners also provide Holevo information stability, and we do so by bounding the Holevo information from above by a function of the QLDP privacy parameter ε.

Holevo Information Stability from QLDP: Let $X \sim P_X$ be a random variable, which can take values in an alphabet $\mathcal{X}$. Depending on X, the state ρ^X is chosen from the set $\{\rho^1, \ldots, \rho^{|\mathcal{X}|}\}$. Then the state ρ^X is sent through a quantum channel $\mathcal{A}_{A\to B}$ satisfying QLDP. Afterwards, the goal is to identify X by performing a measurement described by the POVM $\{M_y\}_{y\in\mathcal{Y}}$, which realizes the output Y.

Here, we focus on how much information about X can be learned from the output of the quantum privacy mechanism $\mathcal{A}(\rho^X)$ with an emphasis on the Holevo information $\mathsf{I}(X;B)_\sigma$. Here, we define the classical–quantum state

$$\sigma_{XB} \coloneqq \sum_{x\in\mathcal{X}} P_X(x)\, |x\rangle\langle x| \otimes \mathcal{A}(\rho^x), \tag{6.109}$$

and the Holevo information of σ_{XB} as

$$\mathsf{I}(X;B)_\sigma \coloneqq D(\sigma_{XB}\|\sigma_X \otimes \sigma_B), \tag{6.110}$$

with $\sigma_X = \mathrm{Tr}_B[\sigma_{XB}]$ and $\sigma_B = \mathrm{Tr}_X[\sigma_{XB}]$.

We show that Holevo information stability, i.e., $I(X;B) \leq \beta$, can be achieved by ε-QLDP mechanisms. To this end, we establish a bound improving upon the existing bound $\beta = \min\{\varepsilon, \varepsilon^2/2\}$ given in [95, Proposition 15], with the improvement following because $\varepsilon\left(\frac{e^\varepsilon-1}{e^\varepsilon+1}\right) \leq \min\{\varepsilon, \varepsilon^2/2\}$ for all $\varepsilon \geq 0$.

Proposition 35 (Holevo Information Stability from QLDP)**.** *Let $\mathcal{A}_{A\to B}$ be a quantum channel. If $\mathcal{A}$ satisfies (ε,δ)-QLDP, then the Holevo information has the following upper bound:*

$$\mathsf{I}(X;B)_\sigma \leq \varepsilon\left(\frac{e^\varepsilon-1}{e^\varepsilon+1}\right). \tag{6.111}$$

Proof. Consider that

$$\begin{aligned}
\mathsf{I}(X;B)_\sigma &= \sum_{x\in\mathcal{X}} P_X(x)\, D\left(\mathcal{A}(\rho^x)\middle\|\sum_{x'\in\mathcal{X}} P_X(x')\mathcal{A}(\rho^{x'})\right) && (6.112)\\
&\leq \sum_{x\in\mathcal{X}}\sum_{x'\in\mathcal{X}} P_X(x)P_X(x')\, D\Big(\mathcal{A}(\rho^x)\Big\|\mathcal{A}(\rho^{x'})\Big) && (6.113)\\
&\leq \varepsilon\left(\frac{e^\varepsilon-1}{e^\varepsilon+1}\right)\sum_{x\in\mathcal{X}}\sum_{x'\in\mathcal{X}} P_X(x)P_X(x')\, T(\rho_x,\rho'_x) && (6.114)\\
&\leq \varepsilon\left(\frac{e^\varepsilon-1}{e^\varepsilon+1}\right), && (6.115)
\end{aligned}$$

where the first inequality follows from the joint convexity of quantum relative entropy [82]. The second inequality follows from Proposition 25 and the last inequality from $T(\rho,\sigma)\leq 1$ for states ρ and σ, concluding the proof. □

6.5 Concluding Remarks

In this chapter, we studied quantum private hypothesis testing and provided bounds on the sample complexity when we have access to privatized samples. We showcased the cost of privacy on this statistical task while providing tight sample complexity bounds on special sets of states (i.e., orthogonal states), and special classes of private channels (i.e., high-private channels with $\varepsilon < 1$). Next, we studied the impact of privacy on the estimation of the expectation of an observable with high probability. We provided upper and lower bounds on the

number of samples needed to achieve a fixed error tolerance with high probability. Finally, we explored how the contraction of divergences derived in Chapter 5 can be applied in other applications, including ensuring fairness and providing formal guarantees on stability and generalization for quantum learning settings by addressing an open question posed in [18].

CHAPTER 7

CONCLUSION AND FUTURE DIRECTIONS

In this thesis, we introduced a theoretical foundation to flexible privacy frameworks for quantum systems (quantum pufferfish privacy), while showcasing that they satisfy desired properties, such as post-processing, composition, that are important in private data analysis. We also discussed several distinctions that may arise in the quantum setting, with the possibility of degrading the level of privacy when several quantum systems are jointly doing the inference. We designed the first formal methods to audit the privacy of black-box quantum mechanisms and verify whether they indeed satisfy the privacy level as demanded. Furthermore, we presented several variants of QPP while establishing information-theoretic connections between them.

Notably, we provided the first operational interpretation of Datta–Leditzky information-spectrum divergence [29] when the QPP framework consists of the scenario where an adversary is allowed to perform all possible measurements. Extending this connection to general QPP mechanisms, we studied measured-hockey-stick divergences under practically relevant measurement classes, which quantify optimal privacy parameters.

Moreover, we studied how the contraction coefficients of quantum divergences behave with privacy constraints imposed by quantum local differential privacy (QLDP). In particular, we completely characterized the contraction coefficient of trace distance under (ε, δ)-QLDP. The tools developed here enabled new fields of study, including studying statistical tasks under privacy constraints, which was largely unexplored before this. To this end, we studied how the sample complexity of quantum hypothesis testing and the learning

expectation of observables scale with the required privacy level when we have access to privatized quantum states in contrast to the original quantum states.

Next, we present several future directions that arise with the developments made in this thesis towards ensuring privacy in a quantum world.

Composition of Private Mechanisms: In private data analysis with quantum data, it is vital to understand how privacy degrades when multiple private mechanisms are combined. This process is what we refer to as the composition of private mechanisms. We studied the consequences of sequential, parallel, and adaptive composition of QPP mechanisms in Section 3.4. However, it is not known whether the composition of two QPP mechanisms (when the measurement set comprises of all measurements) with parameters $(\varepsilon_i, \delta_i)$ for $i \in \{1, 2\}$ provides a privacy level demanded by $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$ when $\delta_i \neq 0$. One possible way to provide an answer to this question is to study the sub-additivity of quantum divergences, including the Datta–Leditzky divergence ((2.11)) and the hockey-stick divergence ((2.3)). In particular, for the setting of parallel composition, it is about showing whether

$$\overline{D}^{\delta'}(\rho_1 \otimes \rho_2 \| \sigma_1 \otimes \sigma_2) \overset{?}{\leq} \overline{D}^{\delta_1}(\rho_1 \| \sigma_1) + \overline{D}^{\delta_2}(\rho_2 \| \sigma_2) \tag{7.1}$$

holds or not for $\delta' \in [0, 1)$. Proving the above will provide insights on how we should go about the private data analysis with provably tighter privacy guarantees, improving Theorem 2. If one disproves this inequality, this provides a clear distinction between classical and quantum systems when ensuring privacy, with the possibility of joint measurements inferring more information, degrading the level of privacy.

Privacy-Utility Analysis: To solely achieve privacy, one can fully perturb the

quantum state using a replacement channel. However, then we will not be able to learn anything from that privatized state, making the entire process useless. Therefore, it is crucial to evaluate the utility retained by a privatization mechanism under privacy constraints and to characterize the tradeoffs between privacy and utility. In Section 3.5, we introduced a utility metric grounded in an operational approach where this metric facilitated an in-depth analysis of privacy-utility tradeoffs, particularly for the depolarization mechanism. Recently, in [54], the optimal privacy-utility tradeoffs of unital private mechanisms are studied when the quantum local differential privacy (a special case of QPP) is used as the privacy metric. However, the utility metrics in these settings reflect a worst-case scenario and may impose stricter accuracy demands than required for specific applications. Consequently, developing application-specific utility metrics, in conjunction with the flexible QPP privacy framework, is essential. This approach would open new research paradigms to explore optimal privacy-utility tradeoffs for various information-processing tasks, tailored to operationally relevant privacy and utility metrics.

Auditing Quantum Privacy: Auditing for privacy aims to detect violations in privacy guarantees and reject incorrect algorithms. We proposed the first hypothesis-testing-based auditing pipeline for the setting of quantum differential privacy in Section 3.6 and analyzed its type-I error. Analyzing the type-II error of our quantum privacy auditing pipeline is an interesting future research direction since it provides insights on the power of the test to avoid incorrectly accepting algorithms. We also highlighted how one can use classical methods, including semi-definite programs (SDPs), to do the same task. However, as we elaborated, those methods become infeasible with the increasing number of qubits in the quantum systems. Another avenue to address this is to use

quantum algorithms and methods to estimate quantum divergences, such as Datta–Leditzky divergence and hockey-stick divergence. Since these quantities are SDP computable, it is an interesting direction to explore on how quantum methods to solve SDPs can benefit for this task. For example, it is possible to use the Qslack method introduced in [20] as a means to estimate quantum divergences that will in return, provide guarantees on the privacy of specific quantum algorithms.

For flexible privacy frameworks with practically motivated classes of measurements, we provided methods for auditing using measured hockey-stick divergences in Chapter 4. Towards this, for the class of local measurements with classical post-processing, and one-way local measurements with classical communications, it is interesting to explore how the extendible measurements proposed in [120] will provide efficiently computable bounds on measured divergences that can be used for auditing. Furthermore, these studies will provide insights towards designing efficient quantum algorithms that achieve QPP through the examples found through auditing.

Statistical Learning under Privacy Constraints: Studying statistical problems under privacy constraints is vital to understanding the price that we have to pay to ensure privacy. To this end, the contraction of statistical measures and divergences under privacy constraints is an important technical tool. However, in the quantum setting, this area of research is largely unexplored even for fundamental statistical tasks. We characterized the contraction of quantum divergences under a local variant of quantum privacy (QLDP in Definition 15), which enables this field of study. There are quite interesting questions left to be answered in terms of the QLDP setting and also for general privacy frameworks.

First, it is an interesting future direction to precisely characterize privatized contraction coefficients of other quantities (e.g., quantum relative entropy), similar to classically known results in [8, Theorem 1] for commuting states. Specifically, for the classical setting with the privacy constraints imposed by ε-local differential privacy, the privatized contraction coefficient is characterized as follows:

$$\eta_D^{\varepsilon} = \left(\frac{e^{\varepsilon} - 1}{e^{\varepsilon} + 1}\right)^2. \tag{7.2}$$

It is an interesting question to see whether this characterization is true for the quantum setting with the privacy constraints imposed by QLDP.

With the same goal in mind, we see that Bures distance has appeared in the analysis of a hypothesis testing task through [22]. To this end, finding contraction coefficients of the Bures distance, denoted as η_B^{ε}, would potentially give a lower bound for the sample complexity as follows: for some constant $C > 0$,

$$\mathrm{SC}_{(\rho,\sigma)}^{\varepsilon} \geq C \frac{1}{\eta_B^{\varepsilon} \left[d_B(\rho, \sigma)\right]^2}. \tag{7.3}$$

Future work also includes exploring tight upper and lower bounds related to Theorem 8 in the general setting with $\delta > 0$ and analyzing the cost of privacy and the impact of adaptive strategies on sample complexity. Another interesting research question is characterizing the sample complexity of asymmetric binary quantum hypothesis testing. Here, one could make use of the non-private sample-complexity bounds presented in [22, Theorem 9]. Preliminary results in this direction are found in Remark 38.

Furthermore, the study of estimating expected values of observables with respect to some input states while ensuring privacy of the state also requires further exploration. In Section 6.3, we studied several special cases of observ-

ables satisfying certain properties (e.g., O being a measurement operator to derive achievability protocols). To extend these ideas to a broader class of observables, the tools developed in relation to classical shadow tomography will be useful [68,78]. The main goal of classical shadow tomography is to reconstruct and estimate key properties of a quantum system without needing to fully determine its quantum state, which is a computationally expensive task. In particular, developing private variants of classical shadows will be a potential direction to make progress in this study.

With the framework proposed for ensuring privacy of channels in Definition 13, it is a promising direction to explore the impact on the task of discriminating between two channels when one has access to privatized channels. Towards this, it is important to derive contraction coefficients for channel divergences under privacy constraints. With that, one can also utilize the non-private query complexity of channel discrimination (i.e., number of channel accesses required to achieve a fixed tolerance of guessing the channel correctly) derived in [100] as the baseline for the comparison and cost evaluation.

Moreover, our work opens up new research directions in studying information-constrained statistical problems. The main technical tools for this kind of analysis are the contraction coefficients under the relevant information constraints. These constraints can be access to noisy channels instead of noiseless channel access, communication constraints, memory constraints, network constraints (distributed parties in a network), and many more. Toward this direction, in a recent work [50], the sample complexity of noisy quantum hypothesis testing is studied using the Doeblin coefficient defined as follows:

$$\alpha(\mathcal{N}) \coloneqq \sup_{X_B \in \mathrm{Herm}} \left\{ \mathrm{Tr}[X_B] : I_A \otimes X_B \leq \Gamma^{\mathcal{N}}_{AB} \right\} \tag{7.4}$$

as an assisting tool to provide efficiently computable bounds on the contraction coefficients of general noisy channels.

In the longer term, the tools and frameworks developed in this thesis will provide a theoretical foundation for private data analysis with classical and quantum data in quantum systems and networks.

APPENDIX A

CHAPTER 3 OF APPENDIX

A.1 Proof of Lemma 2

The proof given below is closely related to the proof of [126, Lemma 6.21], but there are some subtle differences and so we provide it here for completeness.

Let

$$\Sigma := (\rho - \lambda\sigma)_+ , \tag{A.1}$$

$$G := (\lambda\sigma)^{1/2} (\lambda\sigma + \Sigma)^{-1/2} . \tag{A.2}$$

Note that

$$0 \leq G^\dagger G \tag{A.3}$$

$$= (\lambda\sigma + \Sigma)^{-1/2} (\lambda\sigma) (\lambda\sigma + \Sigma)^{-1/2} \tag{A.4}$$

$$\leq (\lambda\sigma + \Sigma)^{-1/2} (\lambda\sigma + \Sigma) (\lambda\sigma + \Sigma)^{-1/2} \tag{A.5}$$

$$\leq I. \tag{A.6}$$

From the fact that $\rho - \lambda\sigma \leq \Sigma$, it follows that

$$\rho \leq \lambda\sigma + \Sigma. \tag{A.7}$$

Define the following state:

$$\widetilde{\rho} := \frac{G\rho G^\dagger}{\mathrm{Tr}[G^\dagger G\rho]}. \tag{A.8}$$

Consider that

$$
\begin{aligned}
&1 - \mathrm{Tr}[G\rho G^\dagger] \\
&= \mathrm{Tr}[(I - G^\dagger G)\rho] && \text{(A.9)} \\
&\leq \mathrm{Tr}[(I - G^\dagger G)\,(\lambda\sigma + \Sigma)] && \text{(A.10)} \\
&= \mathrm{Tr}[(I - G^\dagger G)\,(\lambda\sigma + \Sigma)] && \text{(A.11)} \\
&= \mathrm{Tr}[\lambda\sigma + \Sigma] \\
&\quad - \mathrm{Tr}\Big[(\lambda\sigma + \Sigma)^{-1/2}\,(\lambda\sigma)\,(\lambda\sigma + \Sigma)^{-1/2}\,(\lambda\sigma + \Sigma)\Big] && \text{(A.12)} \\
&= \mathrm{Tr}[\lambda\sigma + \Sigma] - \mathrm{Tr}[\lambda\sigma] && \text{(A.13)} \\
&= \mathrm{Tr}[\Sigma] && \text{(A.14)} \\
&= \delta && \text{(A.15)}
\end{aligned}
$$

This implies that

$$\mathrm{Tr}[G\rho G^\dagger] \geq 1 - \delta. \tag{A.16}$$

Then it follows that

$$
\begin{aligned}
\widetilde{\rho} &= \frac{G\rho G^\dagger}{\mathrm{Tr}[G^\dagger G\rho]} && \text{(A.17)} \\
&\leq \frac{G\,(\lambda\sigma + \Sigma)\,G^\dagger}{\mathrm{Tr}[G^\dagger G\rho]} && \text{(A.18)} \\
&= \frac{\lambda\sigma}{\mathrm{Tr}[G^\dagger G\rho]} && \text{(A.19)} \\
&\leq \frac{\lambda\sigma}{1 - \delta}. && \text{(A.20)}
\end{aligned}
$$

Let $\psi_{RA} = \sqrt{\rho_A}\Gamma_{RA}\sqrt{\rho_A}$ be the canonical purification of ρ_A, with $\Gamma_{RA} \coloneqq \sum_{i,j} |i\rangle\langle j|_R \otimes$

$|i\rangle\langle j|_A$, and let $\widetilde{\psi}_{RA} = \frac{G_A \psi_{RA} G_A^\dagger}{\mathrm{Tr}[G^\dagger G \rho]}$ purify $\widetilde{\rho}$. Then

$$\sqrt{F(\rho, \widetilde{\rho})} \geq \frac{1}{\mathrm{Tr}[G^\dagger G \rho]} |\langle \psi|_{RA} I_R \otimes G_A |\psi\rangle_{RA}| \tag{A.21}$$

$$\geq |\langle \psi|_{RA} I_R \otimes G_A |\psi\rangle_{RA}| \tag{A.22}$$

$$= \left|\langle \Gamma|_{RA} I_R \otimes \sqrt{\rho_A} G_A \sqrt{\rho_A} |\Gamma\rangle_{RA}\right| \tag{A.23}$$

$$= |\mathrm{Tr}[G\rho]| \tag{A.24}$$

$$\geq \mathrm{Re}[\mathrm{Tr}[G\rho]] \tag{A.25}$$

$$= \mathrm{Tr}[\overline{G}\rho] \tag{A.26}$$

$$= 1 - \mathrm{Tr}[\left(I - \overline{G}\right)\rho], \tag{A.27}$$

where

$$\overline{G} := \frac{G + G^\dagger}{2}. \tag{A.28}$$

The first inequality follows from Uhlmann's theorem for fidelity, and the second follows because $\mathrm{Tr}[G^\dagger G \rho] \leq 1$. Observe that $\overline{G} \leq I$ because $\|G\|_\infty \leq 1$ and by

applying the triangle inequality. So this means that $I - \overline{G} \geq 0$. Now consider that

$$
\begin{aligned}
&\mathrm{Tr}[\left(I - \overline{G}\right)\rho] \\
&\leq \mathrm{Tr}[\left(I - \overline{G}\right)(\lambda\sigma + \Sigma)] && \text{(A.29)} \\
&= \mathrm{Tr}[\lambda\sigma + \Sigma] - \mathrm{Tr}[\overline{G}\,(\lambda\sigma + \Sigma)] && \text{(A.30)} \\
&= \mathrm{Tr}[\lambda\sigma + \Sigma] - \frac{1}{2}\,\mathrm{Tr}\left[\left((\lambda\sigma)^{1/2}\,(\lambda\sigma + \Sigma)^{-1/2} + (\lambda\sigma + \Sigma)^{-1/2}\,(\lambda\sigma)^{1/2}\right)(\lambda\sigma + \Sigma)\right] && \text{(A.31)} \\
&= \mathrm{Tr}[\lambda\sigma + \Sigma] - \frac{1}{2}\,\mathrm{Tr}\left[(\lambda\sigma)^{1/2}\,(\lambda\sigma + \Sigma)^{-1/2}\,(\lambda\sigma + \Sigma)\right] \\
&\qquad - \frac{1}{2}\,\mathrm{Tr}\left[(\lambda\sigma + \Sigma)^{-1/2}\,(\lambda\sigma)^{1/2}\,(\lambda\sigma + \Sigma)\right] && \text{(A.32)} \\
&= \mathrm{Tr}[\lambda\sigma + \Sigma] - \frac{1}{2}\,\mathrm{Tr}\left[(\lambda\sigma)^{1/2}\,(\lambda\sigma + \Sigma)^{1/2}\right] - \frac{1}{2}\,\mathrm{Tr}\left[(\lambda\sigma)^{1/2}\,(\lambda\sigma + \Sigma)^{1/2}\right] && \text{(A.33)} \\
&= \mathrm{Tr}[\lambda\sigma + \Sigma] - \mathrm{Tr}\left[(\lambda\sigma)^{1/2}\,(\lambda\sigma + \Sigma)^{1/2}\right] && \text{(A.34)} \\
&\leq \mathrm{Tr}[\lambda\sigma + \Sigma] - \mathrm{Tr}\left[(\lambda\sigma)^{1/2}\,(\lambda\sigma)^{1/2}\right] && \text{(A.35)} \\
&= \mathrm{Tr}[\lambda\sigma + \Sigma] - \mathrm{Tr}[\lambda\sigma] && \text{(A.36)} \\
&= \mathrm{Tr}[\Sigma] && \text{(A.37)} \\
&= \delta. && \text{(A.38)}
\end{aligned}
$$

So all of this implies that

$$\sqrt{F(\rho, \widetilde{\rho})} \geq 1 - \delta, \tag{A.39}$$

and in turn that

$$F(\rho, \widetilde{\rho}) \geq (1 - \delta)^2\,. \tag{A.40}$$

By applying the inequality

$$\frac{1}{2}\,\|\rho - \widetilde{\rho}\|_1 \leq \sqrt{1 - F(\rho, \widetilde{\rho})}, \tag{A.41}$$

we conclude that

$$\frac{1}{2}\left\|\rho-\widetilde{\rho}\right\|_1 \leq \sqrt{1-(1-\delta)^2} \tag{A.42}$$

$$= \sqrt{1-(1-2\delta+\delta^2)} \tag{A.43}$$

$$= \sqrt{2\delta-\delta^2} \tag{A.44}$$

$$= \sqrt{\delta\,(2-\delta)}. \tag{A.45}$$

Putting everything together, we see that $\widetilde{\rho}$ is a quantum state satisfying

$$\frac{1}{2}\left\|\rho-\widetilde{\rho}\right\|_1 \leq \sqrt{\delta\,(2-\delta)}, \tag{A.46}$$

$$\widetilde{\rho} \leq \frac{\lambda\sigma}{1-\delta}. \tag{A.47}$$

This means that $\widetilde{\rho}$ and $\frac{\lambda}{1-\delta}$ are feasible for $D_{\max}^{\sqrt{\delta(2-\delta)}}(\rho\|\sigma)$, and so it follows that

$$D_{\max}^{\sqrt{\delta(2-\delta)}}(\rho\|\sigma) \leq \ln\left(\frac{\lambda}{1-\delta}\right) \tag{A.48}$$

$$= \ln\lambda + \ln\left(\frac{1}{1-\delta}\right). \tag{A.49}$$

This concludes the proof.

A.2 Subadditivity of Smooth Max-Relative Entropy

Lemma 7. *Given $\delta_1, \delta_2 \in [0,1]$, states ρ_1 and ρ_2, and PSD operators σ_1 and σ_2, the following subadditivity relation holds*

$$D_{\max}^{\delta_1+\delta_2}(\rho_1\otimes\rho_2\|\sigma_1\otimes\sigma_2) \leq D_{\max}^{\delta_1}(\rho_1\|\sigma_1) + D_{\max}^{\delta_2}(\rho_2\|\sigma_2). \tag{A.50}$$

Proof. Let $\overline{\rho}_i$ and λ_i be optimal choices for $D_{\max}^{\delta_i}(\rho_i\|\sigma_i)$, for $i \in \{1,2\}$. Then, consider that

$$\overline{\rho}_1\otimes\overline{\rho}_2 \leq \lambda_1\sigma_1\otimes\overline{\rho}_2 \leq \lambda_1\sigma_1\otimes\lambda_2\sigma_2 = \lambda_1\lambda_2\sigma_1\otimes\sigma_2. \tag{A.51}$$

Furthermore, consider that

$$\frac{1}{2}\left\|\overline{\rho}_1\otimes\overline{\rho}_2-\rho_1\otimes\rho_2\right\|_1 = \frac{1}{2}\left\|\overline{\rho}_1\otimes\overline{\rho}_2-\overline{\rho}_1\otimes\rho_2+\overline{\rho}_1\otimes\rho_2-\rho_1\otimes\rho_2\right\|_1 \tag{A.52}$$

$$= \frac{1}{2}\left\|\overline{\rho}_1\otimes(\overline{\rho}_2-\rho_2)+(\overline{\rho}_1-\rho_1)\otimes\rho_2\right\|_1 \tag{A.53}$$

$$\leq \frac{1}{2}\left\|\overline{\rho}_1\otimes(\overline{\rho}_2-\rho_2)\right\|_1+\frac{1}{2}\left\|(\overline{\rho}_1-\rho_1)\otimes\rho_2\right\|_1 \tag{A.54}$$

$$= \frac{1}{2}\left\|\overline{\rho}_1\right\|_1\left\|\overline{\rho}_2-\rho_2\right\|_1+\frac{1}{2}\left\|\overline{\rho}_1-\rho_1\right\|_1\|\rho_2\|_1 \tag{A.55}$$

$$= \frac{1}{2}\left\|\overline{\rho}_2-\rho_2\right\|_1+\frac{1}{2}\left\|\overline{\rho}_1-\rho_1\right\|_1 \tag{A.56}$$

$$\leq \delta_1+\delta_2, \tag{A.57}$$

where (A.52) follows from the triangular inequality for the trace norm and the final inequality from the assumption that $\overline{\rho}_i$ are the optimizers for $D_{\max}^{\delta_i}(\rho_i\|\sigma_i)$, for $i\in\{1,2\}$.

Finally we have shown that $\overline{\rho}_1\otimes\overline{\rho}_2$ and $\lambda_1\lambda_2$ are candidates for the optimization for $D_{\max}^{\delta_1+\delta_2}(\rho_1\otimes\rho_2\|\sigma_1\otimes\sigma_2)$, thus concluding the proof. □

A.3 Proof of Theorem 1

For (1): Fix $(\rho^{\mathcal{R}},\rho^{\mathcal{T}})$ in (3.4), $M\in\mathcal{M}$. Consider

$$\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{R}})\Big] = \mathrm{Tr}\left[M\sum_{i=1}^{k}p_i\mathcal{A}_i(\rho^{\mathcal{R}})\right] \tag{A.58}$$

$$= \sum_{i=1}^{k}p_i\mathrm{Tr}\Big[M\mathcal{A}_i(\rho^{\mathcal{R}})\Big] \tag{A.59}$$

$$\overset{(a)}{\leq} \sum_{i=1}^{k}p_i\left(e^{\varepsilon}\mathrm{Tr}\Big[M\mathcal{A}_i(\rho^{\mathcal{T}})\Big]+\delta\right) \tag{A.60}$$

$$= \sum_{i=1}^{k}p_ie^{\varepsilon}\mathrm{Tr}\Big[M\mathcal{A}_i(\rho^{\mathcal{T}})\Big]+\delta \tag{A.61}$$

$$= e^{\varepsilon}\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big]+\delta, \tag{A.62}$$

where (a) follows due to each $\mathcal{A}_i$ being (ε, δ)-QPP. This relation is true for every $(\rho^{\mathcal{R}}, \rho^{\mathcal{T}})$, and so it is true for all such pairs generated from $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$.

For (2): Fix $0 \leq M' \leq I$ such that $M' \in \mathcal{M}'$ as stated in the property. With that assumption, there exists $M \in \mathcal{M}$ such that $M = \mathcal{N}^\dagger(M')$. Consider that

$$\begin{aligned} \mathrm{Tr}\Big[M'\mathcal{N}\Big(\mathcal{A}(\rho^{\mathcal{R}})\Big)\Big] &= \mathrm{Tr}\Big[\mathcal{N}^\dagger(M')\mathcal{A}(\rho^{\mathcal{R}})\Big] && \text{(A.63)} \\ &= \mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{R}})\Big], && \text{(A.64)} \end{aligned}$$

where $\mathcal{N}^\dagger$ is the adjoint of $\mathcal{N}$, implying that

$$0 \leq \mathcal{N}^\dagger(M') = M \leq I \tag{A.65}$$

because $\mathcal{N}^\dagger$ is positive and unital by the assumption $\mathcal{N}$ is a quantum channel. Similarly, we have that $\mathrm{Tr}\Big[M'\mathcal{N}\left(\mathcal{A}(\rho^{\mathcal{T}})\right)\Big] = \mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big]$, and we conclude that the processed mechanism satisfies (ε, δ)-QPP with the choice of $\mathcal{M}' \subseteq \{M' : \mathcal{N}^\dagger(M') \in \mathcal{M}\}$.

For (3): Fix $(\mathcal{R}_i, \mathcal{T}_i) \in \mathcal{Q}$ for $i \in \{1, \ldots, k\}$, and $\bigotimes_{i=1}^k M_i \in \bigotimes_{i=1}^k \mathcal{M}_i$. Denote

$$\mathcal{A}^{(k)}(\rho^{\mathcal{R}^{(k)}}) \coloneqq \mathcal{A}_1(\rho^{\mathcal{R}_1}) \otimes \mathcal{A}_2(\rho^{\mathcal{R}_2}) \otimes \ldots \mathcal{A}_k(\rho^{\mathcal{R}_k}) \tag{A.66}$$

and $\mathcal{A}^{\otimes k}(\rho^{\mathcal{T}^{(k)}})$ similarly by replacing $\mathcal{R}$ with $\mathcal{T}$.

Fix $i \in \{1, \ldots, k\}$. Consider that $\mathrm{Tr}\Big[M_i\mathcal{A}_i(\rho^{\mathcal{R}_i})\Big] \leq 1 \leq 1 + \delta_i$ because $\mathrm{Tr}\Big[M_i\mathcal{A}_i(\rho^{\mathcal{R}_i})\Big]$ is a probability. Combining with the inequality $\mathrm{Tr}\Big[M_i\mathcal{A}_i(\rho^{\mathcal{R}_i})\Big] \leq e^{\varepsilon_i}\mathrm{Tr}\Big[M_i\mathcal{A}_i(\rho^{\mathcal{T}_i})\Big] + \delta_i$, which holds from the assumption that QPP holds, we conclude that

$$\mathrm{Tr}\Big[M_i\mathcal{A}_i(\rho^{\mathcal{R}_i})\Big] \leq \min\Big\{e^{\varepsilon_i}\mathrm{Tr}\Big[M_i\mathcal{A}_i(\rho^{\mathcal{T}_i})\Big], 1\Big\} + \delta_i. \tag{A.67}$$

Consider that

$$\prod_{i=1}^{k} \operatorname{Tr}\left[M_i \mathcal{A}_i(\rho^{\mathcal{R}_i})\right]$$

$$\leq \left(\min\left\{1, e^{\varepsilon_1} \operatorname{Tr}\left[M_1 \mathcal{A}_1(\rho^{\mathcal{R}_1})\right]\right\} + \delta_1\right) \prod_{i=2}^{k} \operatorname{Tr}\left[M_i \mathcal{A}_i(\rho^{\mathcal{R}_i})\right] \tag{A.68}$$

$$\leq \min\left\{1, e^{\varepsilon_1} \operatorname{Tr}\left[M_1 \mathcal{A}_1(\rho^{\mathcal{R}_1})\right]\right\} \prod_{i=2}^{k} \operatorname{Tr}\left[M_i \mathcal{A}_i(\rho^{\mathcal{R}_i})\right] + \delta_1 \tag{A.69}$$

$$\leq \min\left\{1, e^{\varepsilon_1} \operatorname{Tr}\left[M_1 \mathcal{A}_1(\rho^{\mathcal{R}_1})\right]\right\} \times$$

$$\left(\min\left\{1, e^{\varepsilon_2} \operatorname{Tr}\left[M_2 \mathcal{A}_2(\rho^{\mathcal{T}_2})\right]\right\} + \delta_2\right) \prod_{i=3}^{k} \operatorname{Tr}\left[M_i \mathcal{A}_i(\rho^{\mathcal{R}_i})\right] + \delta_1 \tag{A.70}$$

$$\leq \prod_{j=1}^{2} \min\left\{e^{\varepsilon_j} \operatorname{Tr}\left[M_j \mathcal{A}_j(\rho^{\mathcal{T}_j})\right], 1\right\} \prod_{i=3}^{k} \operatorname{Tr}\left[M_i \mathcal{A}_i(\rho^{\mathcal{R}_i})\right] + \delta_1 + \delta_2 \tag{A.71}$$

$$\leq e^{\sum_{i=1}^{k} \varepsilon_i} \prod_{i=1}^{k} \operatorname{Tr}\left[M_i \mathcal{A}_i(\rho^{\mathcal{T}_i})\right] + \sum_{i=1}^{k} \delta_i, \tag{A.72}$$

where the last inequality follows by proceeding with similar expansions for each remaining term of the product as carried out in the first three steps.

A.4 Proof of Proposition 3

Fix $M', M \in \bar{\mathcal{M}}$, $P_X \in \Theta$, and $(\mathcal{R}_1, \mathcal{T}_1), (\mathcal{R}_2, \mathcal{T}_2) \in \mathcal{Q}$. Let $M_y := (\langle y| \otimes I) M' (|y\rangle \otimes I)$ and note that M_y is a measurement operator in $\bar{\mathcal{M}}$. Recall the definition of the channel

$$\overline{\mathcal{E}} := \sum_{y \in \mathcal{Y}} \mathcal{E}^y. \tag{A.73}$$

Consider that

$$\begin{aligned}
&\operatorname{Tr}\left[(M\otimes M')\left(\sum_{y\in\mathcal{Y}}\mathcal{E}^y(\mathcal{A}_1(\rho^{\mathcal{R}_1}))\otimes|y\rangle\langle y|\otimes\mathcal{A}_2^y(\rho^{\mathcal{R}_2})\right)\right]\\
&=\sum_{y\in\mathcal{Y}}\operatorname{Tr}\left[M\mathcal{E}^y(\mathcal{A}_1(\rho^{\mathcal{R}_1}))\otimes M'\left(|y\rangle\langle y|\otimes\mathcal{A}_2^y(\rho^{\mathcal{R}_2})\right)\right] &\text{(A.74)}\\
&=\sum_{y\in\mathcal{Y}}\operatorname{Tr}[M\mathcal{E}^y(\mathcal{A}_1(\rho^{\mathcal{R}_1}))]\operatorname{Tr}[M'\left(|y\rangle\langle y|\otimes\mathcal{A}_2^y(\rho^{\mathcal{R}_2})\right)] &\text{(A.75)}\\
&=\sum_{y\in\mathcal{Y}}\operatorname{Tr}[M\mathcal{E}^y(\mathcal{A}_1(\rho^{\mathcal{R}_1}))]\operatorname{Tr}[M'_y\mathcal{A}_2^y(\rho^{\mathcal{R}_2})] &\text{(A.76)}\\
&\overset{(a)}{\leq}\sum_{y\in\mathcal{Y}}\operatorname{Tr}[M\mathcal{E}^y(\mathcal{A}_1(\rho^{\mathcal{R}_1}))]\left(\min\left\{1,e^{\varepsilon_2}\operatorname{Tr}[M'_y\mathcal{A}_2^y(\rho^{\mathcal{T}_2})]\right\}+\delta_2\right) &\text{(A.77)}\\
&=\sum_{y\in\mathcal{Y}}\operatorname{Tr}[\mathcal{E}^{y\dagger}(M)\mathcal{A}_1(\rho^{\mathcal{R}_1})]\left(\min\left\{1,e^{\varepsilon_2}\operatorname{Tr}[M'_y\mathcal{A}_2^y(\rho^{\mathcal{T}_2})]\right\}+\delta_2\right) &\text{(A.78)}\\
&\overset{(b)}{=}\operatorname{Tr}[M\overline{\mathcal{E}}(\mathcal{A}_1(\rho^{\mathcal{R}_1}))]\delta_2+\sum_{y\in\mathcal{Y}}\left(\operatorname{Tr}[\mathcal{E}^{y\dagger}(M)\mathcal{A}_1(\rho^{\mathcal{R}_1})]\min\left\{1,e^{\varepsilon_2}\operatorname{Tr}[M'_y\mathcal{A}_2^y(\rho^{\mathcal{T}_2})]\right\}\right) &\text{(A.79)}\\
&\overset{(c)}{\leq}\delta_2+\sum_{y\in\mathcal{Y}}\operatorname{Tr}[\mathcal{E}^{y\dagger}(M)\mathcal{A}_1(\rho^{\mathcal{R}_1})]\min\left\{1,e^{\varepsilon_2}\operatorname{Tr}[M'_y\mathcal{A}_2^y(\rho^{\mathcal{T}_2})]\right\} &\text{(A.80)}\\
&\overset{(d)}{\leq}\delta_2+\sum_{y\in\mathcal{Y}}\left(e^{\varepsilon_1}\operatorname{Tr}[\mathcal{E}^{y\dagger}(M)\mathcal{A}_1(\rho^{\mathcal{T}_1})]+\delta_1\right)\min\left\{1,e^{\varepsilon_2}\operatorname{Tr}[M'_y\mathcal{A}_2^y(\rho^{\mathcal{T}_2})]\right\} &\text{(A.81)}\\
&\leq\sum_{y\in\mathcal{Y}}\left(e^{\varepsilon_1}\operatorname{Tr}[\mathcal{E}^{y\dagger}(M)\mathcal{A}_1(\rho^{\mathcal{T}_1})]e^{\varepsilon_2}\operatorname{Tr}[M'_y\mathcal{A}_2^y(\rho^{\mathcal{T}_2})]+\delta_1\right)+\delta_2 &\text{(A.82)}\\
&=\sum_{y\in\mathcal{Y}}\left(e^{\varepsilon_2}e^{\varepsilon_1}\operatorname{Tr}[\mathcal{E}^{y\dagger}(M)\mathcal{A}_1(\rho^{\mathcal{T}_1})]\operatorname{Tr}[M'_y\mathcal{A}_2^y(\rho^{\mathcal{T}_2})]+\delta_1\right)+\delta_2 &\text{(A.83)}\\
&=e^{\varepsilon'}\operatorname{Tr}\left[(M\otimes M')\left(\sum_{y\in\mathcal{Y}}\mathcal{E}^y(\mathcal{A}_1(\rho^{\mathcal{T}_1}))\otimes|y\rangle\langle y|\otimes\mathcal{A}_2^y(\rho^{\mathcal{T}_2})\right)\right]+\delta_2+\delta_1|\mathcal{Y}|, &\text{(A.84)}
\end{aligned}$$

where: (a) from $\operatorname{Tr}[M'_y\mathcal{A}_2^y(\rho^{\mathcal{R}_2})]\leq 1\leq 1+\delta_2$ and $\mathcal{A}_2^y$ being (ε_2,δ_2)-QPP; (b) and (c) from

$$\begin{aligned}
\operatorname{Tr}\left[\sum_{y\in\mathcal{Y}}\mathcal{E}^{y\dagger}(M)\mathcal{A}_1(\rho^{\mathcal{T}_1})\right]&=\operatorname{Tr}\left[M\sum_{y\in\mathcal{Y}}\mathcal{E}^y\mathcal{A}_1(\rho^{\mathcal{T}_1})\right] &\text{(A.85)}\\
&=\operatorname{Tr}[M\overline{\mathcal{E}}(\mathcal{A}_1(\rho^{\mathcal{T}_1}))] &\text{(A.86)}\\
&\leq 1; &\text{(A.87)}
\end{aligned}$$

and (d) from $\mathcal{A}_1$ being $(\varepsilon_1, \delta_1)$-QPP and the fact that $\mathcal{E}^{y\dagger}(M)$ is a measurement operator in $\bar{\mathcal{M}}$.

A.5 Composability with Classically Correlated States

In Property 3 of Theorem 1 and Proposition 3, we considered the case in which two mechanisms, composed either in parallel or adaptively, receive independent inputs (i.e., the input being $\rho^{X_1} \otimes \rho^{X_2}$ where $X_i \sim P_X \in \Theta$ for $i = \{1, 2\}$, which are chosen independently). We now focus on the setting in which the inputs are classically correlated. The input is chosen as a separable state of the form

$$\sigma_I \coloneqq \sum_{z \in \mathcal{Z}} q(z)\, \omega^z \otimes \tau^z, \tag{A.88}$$

where q represents a probability distribution with $q(z) \geq 0$ and $\sum_{z \in \mathcal{Z}} q(z) = 1$, and ω^z and τ^z are quantum states for all $z \in \mathcal{Z}$[1]. One special case of interest is as follows:

$$\sigma_I \coloneqq \sum_{x \in \mathcal{X}} P_X(x)\, \rho^x \otimes \rho^x. \tag{A.89}$$

In this setting, QDP ensures the indistinguishability of the input states

$$\sigma_I^1 \coloneqq \sum_{z \in \mathcal{Z}} q(z)\, \omega_1^z \otimes \tau_1^z \quad \text{and} \quad \sigma_I^2 \coloneqq \sum_{z \in \mathcal{Z}} q(z)\, \omega_2^z \otimes \tau_2^z, \tag{A.90}$$

where $\omega_1^z \sim \omega_2^z$ and $\tau_1^z \sim \tau_2^z$ are neighbors for all $z \in \mathcal{Z}$.

We consider an instance of the QPP framework, called flexible QDP, where $(\mathcal{S}, \mathcal{Q}, \Theta, \mathcal{M})$ is such that Θ and $\mathcal{M}$ are chosen based on user needs, while the other parameters are as given in Section 3.2.3. Flexible QDP then satisfies the following composability properties.

[1]Note that (A.88) covers the case of having input states of the form $\sigma_I \coloneqq \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} q(x, y)\, \omega^x \otimes \tau^y$ where for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ ω^x and τ^y are states, by considering z to be an index for multiple variables, i.e., setting $z = (x, y)$.

Corollary 7 (Composability of Flexible QDP)**.** *Let the initial input to the two mechanisms $\mathcal{A}_1$ and $\mathcal{A}_2$ be of the form $\sum_{z\in\mathcal{Z}} q(z)\,\omega^z\otimes\tau^z$. The following composability properties hold for the QDP framework.*

Parallel composability: Consider the parallel composed mechanism $\sum_{z\in\mathcal{Z}} q(z)\,\mathcal{A}_1(\omega^z)\otimes\mathcal{A}_2(\tau^z)$.

1. *If $\mathcal{A}_i$ is (ε_i,δ_i)-QDP in the framework $(\mathcal{S},\mathcal{Q},\Theta,\mathcal{M}_i)$, for $i\in\{1,2\}$, then the composed mechanism satisfies $(\varepsilon_1+\varepsilon_2,\delta_1+\delta_2)$-QDP in $\left(\mathcal{S},\mathcal{Q}^{(2)},\Theta,\bigotimes_{i=1}^{2}\mathcal{M}_i\right)$*
2. *If $\mathcal{A}_i$ is (ε_i,δ_i)-QDP in the framework $(\mathcal{S},\mathcal{Q},\Theta,\bar{\mathcal{M}})$, for $i\in\{1,2\}$, then the composed mechanism satisfies (ε',δ')-QDP in $\left(\mathcal{S},\mathcal{Q}^{(2)},\Theta,\bar{\mathcal{M}}^2\right)$ with*

$$\varepsilon' := \varepsilon_1+\varepsilon_2+\ln\left(\frac{1}{(1-\delta_1)(1-\delta_2)}\right), \tag{A.91}$$

$$\delta' := \sqrt{\delta_1(2-\delta_1)}+\sqrt{\delta_2(2-\delta_2)}. \tag{A.92}$$

and also satisfies $(\varepsilon_1+\varepsilon_2,\delta)$ in the same framework with $\delta := \min\{\delta_1+e^{\varepsilon_1}\delta_2,\delta_2+e^{\varepsilon_2}\delta_1\}$.

Adaptive composability: Suppose that $\mathcal{A}_1$ satisfies (ε_1,δ_1)-QDP and $\mathcal{A}_2$ chosen adaptively satisfies (ε_2,δ_2)-QDP, as in (3.82). Then, the composed mechanism in Fig. 3.3 with σ_I in (A.88) satisfies $(\varepsilon_1+\varepsilon_2,\delta_2+\delta_1|\mathcal{Y}|)$ in the framework $(\mathcal{S},\mathcal{Q}\times\mathcal{Q},\Theta,\bar{\mathcal{M}}\otimes\bar{\mathcal{M}})$.

Proof. Item 1 in the parallel composability part follows by a similar argument as given in the proof of Property 3 from Theorem 1. For the proof of Item 2, first, we use quasi-convexity of the DL divergence (property 2 in Proposition 2) and then adapt Item 3 of Theorem 2. The adaptive composition result follows along the same lines as the proof of Proposition 3 for fixed z, and then averaging over all $z\in\mathcal{Z}$ gives the desired result. □

Remark 40 (Extensions Beyond Flexible QDP)**.** *Corollary 7 does not hold for the general QPP framework. Indeed, it fails to hold, for instance, for the classical PP framework [77, Theorem 9.1]. Nevertheless, Corollary 7 can be extended to account for input states* $\sum_{x\in\mathcal{X}} P_X(x)\rho^x \otimes \rho^x$ *subjected to additional structural assumptions on the class of admissible distributions:*

$$\Theta \subseteq \left\{ P_X \in \mathcal{P}(\mathcal{X}) : \begin{array}{c} \forall\ (\mathcal{R},\mathcal{T}) \in \mathcal{Q},\ \exists\ x, x' \in \mathcal{X} \\ \text{s.t. } q_{\mathcal{R}}(x) = q_{\mathcal{T}}(x') = 1 \end{array} \right\} \tag{A.93}$$

where $q_{\mathcal{R}}$ *and* $q_{\mathcal{T}}$ *are defined as in Definition 6. The classical version of this condition for PP is known as "universally composable scenarios" [77, Corollary 9.4].*

A.6 Characterizing Optimal Privacy-Utility Tradeoff

In this Appendix, we focus on identifying the optimal utility that can be obtained by applying an (ε,δ)-QPP mechanism. Here, we first focus on the setting in which $\mathcal{Q} = \{(\mathcal{R}_1,\mathcal{R}_2),(\mathcal{R}_2,\mathcal{R}_1)\}$, $\bar{\mathcal{M}} = \{M : 0 \leq M \leq I\}$, and $\Theta = \{P_X\}$, but the following ideas can be extended to the case when $\mathcal{Q}$ is an arbitrary finite set and Θ includes a finite number of probability distributions. However, the computational complexity involved in identifying the optimal utility increases with the cardinality of the set $\mathcal{Q}$ and Θ, due to the addition of more constraints to the optimization problem.

To incorporate privacy requirements, we use the equivalent formulation of QPP via the DL divergence presented in Proposition 1. To this end, first, we employ the SDP formulated in Lemma 1 to compute the relevant DL divergence and then use that in the optimization of utility. We showcase the use of this SDP in characterizing optimal utility next.

Proposition 36 (Optimal Utility for Fixed Privacy Constraints)**.** *The optimal utility, as quantified by the γ-utility metric, for every privacy mechanism that is (ε, δ)-QPP in the $(\mathcal{S}, \mathcal{Q}, \Theta, \bar{\mathcal{M}})$ framework, where $\mathcal{Q} = \{(\mathcal{R}_1, \mathcal{R}_2), (\mathcal{R}_2, \mathcal{R}_1)\}$, is given by the following:*

$$\mathsf{U}(\varepsilon, \delta, \mathcal{R}_1, \mathcal{R}_2) \quad := \quad 1 \; - \; \inf_{\substack{\mu \geq 0 \\ Z_{AD} \geq 0 \\ \Gamma^{\mathcal{B}}_{CD} \geq 0 \\ \Gamma^{\mathcal{A}}_{AC} \geq 0 \\ \lambda_1 \geq 0, Y_1 \geq 0 \\ \lambda_2 \geq 0, Y_2 \geq 0}} \left\{ \begin{array}{c} \mu : \\ Z_{AD} \geq \Gamma_{AD} - \mathrm{Tr}_C\Big[\Gamma^{\mathcal{B}}_{CD} \mathrm{T}_C(\Gamma^{\mathcal{A}}_{AC})\Big], \\ \mu I_A \geq \mathrm{Tr}_D[Z_{AD}], \\ \mathrm{Tr}_D\Big[\Gamma^{\mathcal{B}}_{CD}\Big] = I_C, \\ \mathrm{Tr}_D\Big[\Gamma^{\mathcal{A}}_{AC}\Big] = I_A, \\ \ln(\lambda_1) \leq \varepsilon, \\ \mathrm{Tr}[Y_1] \leq \delta, \\ Y_1 \geq \mathrm{Tr}_A\Big[(\mathrm{T}(\rho^{\mathcal{R}_1}) \otimes I_C)\Gamma^{\mathcal{A}}_{AC}\Big] \\ -\lambda_1 \mathrm{Tr}_A\Big[(\mathrm{T}(\rho^{\mathcal{R}_2}) \otimes I_C)\Gamma^{\mathcal{A}}_{AC}\Big], \\ \ln(\lambda_2) \leq \varepsilon, \\ \mathrm{Tr}[Y_2] \leq \delta, \\ Y_2 \geq \mathrm{Tr}_A\Big[(\mathrm{T}(\rho^{\mathcal{R}_2}) \otimes I_C)\Gamma^{\mathcal{A}}_{AC}\Big] \\ -\lambda_2 \mathrm{Tr}_A\Big[(\mathrm{T}(\rho^{\mathcal{R}_1}) \otimes I_C)\Gamma^{\mathcal{A}}_{AC}\Big] \end{array} \right\}. \tag{A.94}$$

Proof. The proof follows from the SDP formulation of the γ-utility given in Proposition 6, and the privacy constraints
(i.e., $\max\{\overline{D}^{\delta}\Big(\mathcal{A}(\rho^{\mathcal{R}_1}) \| \mathcal{A}(\rho^{\mathcal{R}_2})\Big), \overline{D}^{\delta}\Big(\mathcal{A}(\rho^{\mathcal{R}_2}) \| \mathcal{A}(\rho^{\mathcal{R}_1})\Big)\} \leq \varepsilon$) imposed through the SDP formulation of DL divergence presented in Lemma 1. We also used the fact that for a superoperator $\mathcal{A}$ from system A to C, the following equality holds [75, Eq. (3.2.14)]

$$\mathcal{A}(\rho^{\mathcal{R}_1}) = \mathrm{Tr}_A\Big[(\mathrm{T}(\rho^{\mathcal{R}_1}) \otimes I_C)\Gamma^{\mathcal{A}}_{AC}\Big]. \tag{A.95}$$

□

Remark 41 (Privacy Constraints via Equivalent Formulation through Hockey–

Stick Divergence). *Instead of using the DL divergence, we can also encode the privacy constraints through the equivalent formulation in Remark 5. To this end, the dual formulation of the hockey-stick divergence (can be obtained by (3.28)), as*

$$E_\lambda(\rho\|\sigma) = \inf_{Z\geq 0} \{\mathrm{Tr}[Z] : Z \geq \rho - \lambda\sigma\}, \tag{A.96}$$

can also be incorporated to compute the optimum utility.

Remark 42 (Optimal Privacy Parameters for Fixed Utility). *To find out the optimal (minimal) privacy parameter $\varepsilon^\star$ for a given mechanism $\mathcal{A}$ with the utility constraint γ, and fixed tolerance δ, first we compute the following quantity:*

$$\lambda_1^\star(\mathcal{A},\gamma,\delta) := \inf_{\substack{\lambda\geq 0\\ Z_{AD}\geq 0\\ \Gamma^{\mathcal{B}}_{CD}\geq 0\\ Y_1\geq 0}} \left\{ \begin{array}{c} \lambda : Z_{AD} \geq \Gamma_{AD} - \mathrm{Tr}_C\left[\Gamma^{\mathcal{B}}_{CD}\mathrm{T}_C(\Gamma^{\mathcal{A}}_{AC})\right], \\ (1-\gamma)I_A \geq \mathrm{Tr}_D[Z_{AD}], \\ \mathrm{Tr}_D\left[\Gamma^{\mathcal{B}}_{CD}\right] = I_C, \\ \mathrm{Tr}[Y_1] \leq \delta, \\ Y_1 \geq \mathrm{Tr}_A\left[(\mathrm{T}(\rho^{\mathcal{R}_1})\otimes I_C)\Gamma^{\mathcal{A}}_{AC}\right] \\ -\lambda\mathrm{Tr}_A\left[(\mathrm{T}(\rho^{\mathcal{R}_2})\otimes I_C)\Gamma^{\mathcal{A}}_{AC}\right] \end{array} \right\}. \tag{A.97}$$

Similarly $\lambda_2^\star$ can be obtained by exchanging $\rho^{\mathcal{R}_1}$ and $\rho^{\mathcal{R}_2}$. Then the optimal value is given by

$$\varepsilon^\star(\mathcal{A},\gamma,\delta) := \ln(\max\{\lambda_1^\star(\mathcal{A},\gamma,\delta), \lambda_2^\star(\mathcal{A},\gamma,\delta)\}). \tag{A.98}$$

The optimal (minimal) δ for a fixed ε with a utility constraint can be obtained by encoding the privacy constraint through the dual form of hockey-stick divergence, as given in Remark 41.

A.7 Proof of Lemma 4

The proof follows analogously to the classical version of this bound in [15, Proposition 3.3], along with the upper bound for an arbitrary $D_\alpha(\cdot\|\cdot)$ satisfy-

ing data processing. Set $\alpha > 1$. Then, for such $D_\alpha(\cdot\|\cdot)$, from [126, Equation 4.34], which is obtained by choosing a specific preparation channel, we have that

$$D_\alpha(\rho\|\sigma) \leq \frac{1}{\alpha - 1} \log \mathrm{Tr}\left[\sigma^{1/2}\left(\sigma^{-1/2}\rho\sigma^{-1/2}\right)^\alpha \sigma^{1/2}\right]. \tag{A.99}$$

Let us use the following substitution:

$$D_T(\rho\|\sigma) = \varepsilon. \tag{A.100}$$

Then we have $D_{\max}(\rho\|\sigma) \leq \varepsilon$ and $D_{\max}(\sigma\|\rho) \leq \varepsilon$. Moreover, with the definition of $D_{\max}(\cdot\|\cdot)$ in (2.14), we have $\rho \leq e^\varepsilon\sigma$ and $\sigma \leq e^\varepsilon\rho$. Then we find that

$$e^{-\varepsilon} I \leq \sigma^{-1/2}\rho\sigma^{-1/2} \leq e^\varepsilon I. \tag{A.101}$$

Suppose that $\sigma^{-1/2}\rho\sigma^{-1/2}$ has the following spectral decomposition $\sum_i t_i|\phi_i\rangle\langle\phi_i|$. Then $e^{-\varepsilon} \leq t_i \leq e^\varepsilon$, and so for all i, $\exists\lambda_i \in [0,1]$ such that

$$t_i = \lambda_i e^\varepsilon + (1 - \lambda_i)e^{-\varepsilon}. \tag{A.102}$$

Consider that

$$\begin{aligned}
e^{(\alpha-1)D_\alpha(\rho\|\sigma)} &\leq \mathrm{Tr}\left[\sigma^{1/2}\left(\sigma^{-1/2}\rho\sigma^{-1/2}\right)^\alpha \sigma^{1/2}\right] && \text{(A.103)}\\
&= \mathrm{Tr}\left[\sigma^{1/2}\sum_i (\lambda_i e^\varepsilon + (1-\lambda_i)e^{-\varepsilon})^\alpha |\phi_i\rangle\langle\phi_i| \sigma^{1/2}\right] && \text{(A.104)}\\
&= \sum_i (\lambda_i e^\varepsilon + (1-\lambda_i)e^{-\varepsilon})^\alpha \mathrm{Tr}[\sigma|\phi_i\rangle\langle\phi_i|] && \text{(A.105)}\\
&\leq \sum_i (\lambda_i e^{\varepsilon\alpha} + (1-\lambda_i)e^{-\varepsilon\alpha}) \mathrm{Tr}[\sigma|\phi_i\rangle\langle\phi_i|] && \text{(A.106)}\\
&= e^{\varepsilon\alpha}c_1 + e^{-\varepsilon\alpha}c_2 && \text{(A.107)}
\end{aligned}$$

where the first inequality follows from the inequality in (A.99), the second from the convexity of the function x^α for $\alpha > 1$, and the definitions

$$c_1 := \sum_i \lambda_i \mathrm{Tr}[\sigma|\phi_i\rangle\langle\phi_i|], \tag{A.108}$$

$$c_2 := \sum_i (1-\lambda_i)\mathrm{Tr}[\sigma|\phi_i\rangle\langle\phi_i|]. \tag{A.109}$$

In (A.107), we arrive at a function of α (i.e., $e^{\varepsilon\alpha}c_1 + e^{-\varepsilon\alpha}c_2$). Observing that $c_1 + c_2 = 1$, we can find c_1 and c_2 by evaluating this function of α at $\alpha = 1$, which turns out to be equal to one because

$$\sum_i t_i \mathrm{Tr}[\sigma|\phi_i\rangle\langle\phi_i|] = \mathrm{Tr}[\rho] = 1, \tag{A.110}$$

where t_i is given in (A.102). Proceeding with this we get $c_1 = \frac{1-e^{-\varepsilon}}{e^{\varepsilon}-e^{-\varepsilon}}$. Then, collecting all these relations and simplifying, we obtain,

$$e^{(\alpha-1)D_\alpha(\rho\|\sigma)} \leq \frac{\sinh(\alpha\varepsilon) - \sinh((\alpha-1)\varepsilon)}{\sinh(\varepsilon)}. \tag{A.111}$$

Together with [15, Lemma B.1], we can further bound (A.111) from above by $e^{\alpha(\alpha-1)\varepsilon^2/2}$. With the substitution in (A.100) we conclude the proof.

A.8 Proof of Lemma 5

Fix $P_X \in \Theta$ and $(\mathcal{R}, \mathcal{T}) \in \mathcal{Q}$. By assumption, we have that

$$\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{R}})\Big] \leq e^{\varepsilon}\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big] + \delta. \tag{A.112}$$

With the choice for δ' as in the Lemma statement (i.e., (3.175)), we have $\delta - \delta' = (1-\delta)(e^{\varepsilon'} - e^{\varepsilon})/(e^{\varepsilon}+1)$. Plugging this in, we find that

$$\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{R}})\Big] \leq e^{\varepsilon'}\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big] + \delta' + (\delta - \delta') + (e^{\varepsilon} - e^{\varepsilon'})\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big] \tag{A.113}$$

$$= e^{\varepsilon'}\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big] + \delta' + (e^{\varepsilon} - e^{\varepsilon'})\left(\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big] - \frac{1-\delta}{e^{\varepsilon}+1}\right). \tag{A.114}$$

Since $\varepsilon' \leq \varepsilon$, we get the desired inequality if $\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big] \leq \frac{1-\delta}{e^{\varepsilon}+1}$.

By choosing the measurement operator $I - M$, we also have

$$\mathrm{Tr}\Big[(I-M)\mathcal{A}(\rho^{\mathcal{T}})\Big] \leq e^{\varepsilon}\mathrm{Tr}\Big[(I-M)\mathcal{A}(\rho^{\mathcal{R}})\Big] + \delta. \tag{A.115}$$

Rewriting (A.115), we arrive at

$$\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{R}})\Big] \leq 1 - e^{-\varepsilon}(1-\delta) + e^{-\varepsilon}\,\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big]. \tag{A.116}$$

Similar to the previous manipulations, we get

$$\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{R}})\Big] \leq e^{\varepsilon'}\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big] + \delta' + (e^{\varepsilon'} - e^{-\varepsilon})\left(-\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big] + \frac{1-\delta}{e^{\varepsilon}+1}\right). \tag{A.117}$$

Since $\varepsilon' \leq \varepsilon$, we arrive at the desired inequality when $\mathrm{Tr}\Big[M\mathcal{A}(\rho^{\mathcal{T}})\Big] \geq \frac{1-\delta}{e^{\varepsilon}+1}$. By these two arguments, the desired inequality holds for either of the cases, proving its validity.

A similar inequality holds for every $(\mathcal{R},\mathcal{T}) \in \mathcal{Q}$ and $P_X \in \Theta$. Thus, the desired implication has been proved.

APPENDIX B

CHAPTER 4 OF APPENDIX

B.1 Other Properties of Measured Hockey-Stick Divergence

Lemma 8. *Let $\gamma \geq 0$, and let ρ and σ be states. The following equality holds:*

$$E_\gamma^{\mathcal{M}}(\rho\|\sigma) = \gamma E_{\frac{1}{\gamma}}^{\mathcal{M}}(\sigma\|\rho), \tag{B.1}$$

where $E_\gamma^{\mathcal{M}}(\cdot\|\cdot)$ is defined in (4.5).

Proof. Let $0 \leq \gamma \leq 1$. By applying (4.4), consider that

$$\begin{aligned}
E_\gamma^{\mathcal{M}}(\rho\|\sigma) &= \sup_{M\in\mathcal{M}_2} \operatorname{Tr}[M(\rho - \gamma\sigma)] - (1-\gamma) && \text{(B.2)}\\
&= \sup_{M\in\mathcal{M}_2} \operatorname{Tr}[(I-M)(\rho - \gamma\sigma)] - (1-\gamma) && \text{(B.3)}\\
&= \sup_{M\in\mathcal{M}_2} \operatorname{Tr}[M(\gamma\sigma - \rho)] && \text{(B.4)}\\
&= \gamma \sup_{M\in\mathcal{M}_2} \operatorname{Tr}\left[M\left(\sigma - \frac{1}{\gamma}\rho\right)\right] && \text{(B.5)}\\
&= \gamma E_{\frac{1}{\gamma}}^{\mathcal{M}}(\sigma\|\rho), && \text{(B.6)}
\end{aligned}$$

where the last inequality follows by applying (4.4) with $1/\gamma \geq 1$ since $\gamma \leq 1$.

For $\gamma' \geq 1$, by substituting $\gamma = 1/\gamma'$ in (B.6), we conclude the proof. □

Lemma 9. *Let ρ and σ be states, and let $\gamma \geq 1$. The following inequality holds:*

$$(\gamma+1)E_1^{\mathcal{M}}(\rho\|\sigma) \leq E_\gamma^{\mathcal{M}}(\rho\|\sigma) + E_\gamma^{\mathcal{M}}(\sigma\|\rho) + \gamma - 1. \tag{B.7}$$

Proof. By applying (4.4), consider that

$$(\gamma+1)E_1^{\mathcal{M}}(\rho\|\sigma) = (\gamma+1)\sup_{M\in\mathcal{M}_2}\operatorname{Tr}[M(\rho-\sigma)] \tag{B.8}$$

$$= \sup_{M\in\mathcal{M}_2}\{\operatorname{Tr}[M(\rho-\gamma\sigma)] + \operatorname{Tr}[M(\gamma\rho-\sigma)]\} \tag{B.9}$$

$$\leq \sup_{M\in\mathcal{M}_2}\{\operatorname{Tr}[M(\rho-\gamma\sigma)]\} + \sup_{M\in\mathcal{M}_2}\{\operatorname{Tr}[M(\gamma\rho-\sigma)]\} \tag{B.10}$$

$$= E_\gamma^{\mathcal{M}}(\rho\|\sigma) + \gamma\sup_{M\in\mathcal{M}_2}\operatorname{Tr}\left[M\left(\rho-\frac{1}{\gamma}\sigma\right)\right] \tag{B.11}$$

$$= E_\gamma^{\mathcal{M}}(\rho\|\sigma) + \gamma\left(E_{\frac{1}{\gamma}}^{\mathcal{M}}(\rho\|\sigma) + 1 - \frac{1}{\gamma}\right) \tag{B.12}$$

$$= E_\gamma^{\mathcal{M}}(\rho\|\sigma) + E_\gamma^{\mathcal{M}}(\sigma\|\rho) + \gamma - 1, \tag{B.13}$$

where the penultimate equality follows from (4.4) with $1/\gamma \leq 1$ and the last equality follows by some algebraic manipulations, along with Lemma 8.

□

B.2 Proof of Proposition 12

Consider an arbitrary POVM $\{M, I-M\} \in \mathcal{M}_2$, where $\mathcal{M}_2$ is defined in (4.3). Then the following inequality holds for every $\gamma \geq 1$:

$$\operatorname{Tr}[M(\rho-\gamma\sigma)] \leq \max\{0, \operatorname{Tr}[M(\rho-\gamma\sigma)]\} \tag{B.14}$$

$$\leq \max\{0, \operatorname{Tr}[M(\rho-\gamma\sigma)]\} + \max\{0, \operatorname{Tr}[(I-M)(\rho-\gamma\sigma)]\} \tag{B.15}$$

$$\leq \sup_{\{M_x\in\mathcal{M}\}_x}\sum_x \max\{0, \operatorname{Tr}[M_x(\rho-\gamma\sigma)]\}, \tag{B.16}$$

where the last inequality follows because $\{M, I-M\}$ is a POVM in $\mathcal{M}$. By optimizing the left-hand side over all $M \in \mathcal{M}_2$, we arrive at

$$E_\gamma^{\mathcal{M}}(\rho\|\sigma) \leq \widehat{E}_\gamma^{\mathcal{M}}(\rho\|\sigma). \tag{B.17}$$

To prove the reverse inequality, we consider a POVM $\{M_x \in \mathcal{M}\}_{x\in\mathcal{X}}$ that obeys the coarse-graining condition; that is, $M_x + M_{x'} \in \mathcal{M}$ for all $x, x' \in \mathcal{X}$. Consider the following equality:

$$\sum_x \max\{0, \mathrm{Tr}[M_x(\rho - \gamma\sigma)]\} = \sum_{x:\mathrm{Tr}\left[M_x(\rho-\gamma\sigma)\right]\geq 0} \mathrm{Tr}[M_x(\rho - \gamma\sigma)] \tag{B.18}$$

$$= \mathrm{Tr}[M_+(\rho - \gamma\sigma)] \tag{B.19}$$

$$\leq \sup_{M\in\mathcal{M}_2} \mathrm{Tr}[M(\rho - \gamma\sigma)], \tag{B.20}$$

where the second equality follows from defining

$$M_+ := \sum_{x:\mathrm{Tr}\left[M_x(\rho-\gamma\sigma)\right]\geq 0} M_x \tag{B.21}$$

and the last inequality follows because $M_+, I - M_+ \in \mathcal{M}$ with the coarse-graining assumption and

$$I - M_+ = \sum_{x:\mathrm{Tr}\left[M_x(\rho-\gamma\sigma)\right]<0} M_x. \tag{B.22}$$

Finally, optimizing over all POVMs $\{M_x \in \mathcal{M}\}_x$, we obtain the inequality

$$\widehat{E}_\gamma^{\mathcal{M}}(\rho\|\sigma) \leq E_\gamma^{\mathcal{M}}(\rho\|\sigma), \tag{B.23}$$

and together with (B.17) we conclude the proof.

B.3 Proof of Proposition 13

Data Processing: Under the assumption that the channel $\mathcal{N}$ is $\mathcal{M}$-compatible, it follows that $\mathcal{N}^\dagger(M) \in \mathcal{M}$ for all $M \in \mathcal{M}$. Since the adjoint of a trace-preserving map is unital, it also follows that $\mathcal{N}^\dagger(I) = I$, leading to $\mathcal{N}^\dagger(I-M) = I-\mathcal{N}^\dagger(M) \in \mathcal{M}$.

With that and fixing $M \in \mathcal{M}_2$ such that $M, I - M \in \mathcal{M}$, consider that

$$\begin{aligned} \mathrm{Tr}[M\,(\mathcal{N}(\rho) - \gamma\mathcal{N}(\sigma))] &= \mathrm{Tr}\left[\mathcal{N}^\dagger(M)\,(\rho - \gamma\sigma)\right] && \text{(B.24)} \\ &\leq \sup_{M' \in \mathcal{M}_2} \mathrm{Tr}[M'\,(\rho - \gamma\sigma)] && \text{(B.25)} \\ &= E_\gamma^{\mathcal{M}}(\rho\|\sigma). && \text{(B.26)} \end{aligned}$$

We arrive at the desired inequality by optimizing the left-hand side over all $M \in \mathcal{M}$ such that $I - M \in \mathcal{M}$.

Triangular Inequality: Let $\gamma_1, \gamma_2 \geq 1$. Then

$$\begin{aligned} E_{\gamma_1\gamma_2}^{\mathcal{M}}(\rho\|\sigma) &= \sup_{M \in \mathcal{M}_2} \mathrm{Tr}[M\,(\rho - \gamma_1\gamma_2\sigma)] && \text{(B.27)} \\ &= \sup_{M \in \mathcal{M}_2} \mathrm{Tr}[M\,(\rho - \gamma_1\tau + \gamma_1\tau - \gamma_1\gamma_2\sigma)] && \text{(B.28)} \\ &\leq \sup_{M \in \mathcal{M}_2} \mathrm{Tr}[M\,(\rho - \gamma_1\tau)] + \sup_{M \in \mathcal{M}_2} \mathrm{Tr}[M\,(\gamma_1\tau - \gamma_1\gamma_2\sigma)] && \text{(B.29)} \\ &= E_{\gamma_1}^{\mathcal{M}}(\rho\|\tau) + \gamma_1 E_{\gamma_2}^{\mathcal{M}}(\tau\|\sigma). && \text{(B.30)} \end{aligned}$$

Monotonicity: Let $\gamma_1 \geq \gamma_2 \geq 1$. Then $\gamma_1\sigma \geq \gamma_2\sigma$ and $\rho - \gamma_1\sigma \leq \rho - \gamma_2\sigma$. Since $M \geq 0$, we have that for all $M \in \mathcal{M}_2$

$$\mathrm{Tr}[M\,(\rho - \gamma_1\sigma)] \leq \mathrm{Tr}[M\,(\rho - \gamma_2\sigma)]\,. \tag{B.31}$$

Supremizing over all $M \in \mathcal{M}_2$, we obtain the desired inequality.

Convexity: Let $\{p_x\}_{x\in\mathcal{X}}$ be a probability distribution, and let $\{\rho_x\}_{x\in\mathcal{X}}$ and $\{\sigma_x\}_{x\in\mathcal{X}}$ be sets of quantum states. Then for every $M \in \mathcal{M}_2$ and $\gamma \geq 1$, the following

equality holds:

$$\mathrm{Tr}[M(\rho - \gamma\sigma)] = \sum_x p_x \mathrm{Tr}[M(\rho_x - \gamma\sigma_x)] \tag{B.32}$$

$$\leq \sum_x p_x \sup_{M \in \mathcal{M}_2} \mathrm{Tr}[M(\rho_x - \gamma\sigma_x)] \tag{B.33}$$

$$= \sum_x p_x E_\gamma^{\mathcal{M}}(\rho_x \| \sigma_x). \tag{B.34}$$

By optimizing over $M \in \mathcal{M}_2$ on the left-hand side, we conclude the proof.

B.4 Proof of Proposition 14

Recall the standard form of primal and dual SDPs, as characterized by the Hermitian matrices A and B and a Hermiticity-preserving superoperator Φ [75, Definition 2.26]:

$$\sup_{X \geq 0} \{\mathrm{Tr}[AX] : \Phi(X) \leq B\}, \tag{B.35}$$

$$\inf_{Y \geq 0} \left\{\mathrm{Tr}[BY] : \Phi^\dagger(Y) \geq A\right\}. \tag{B.36}$$

The SDP for the PPT measured hockey-stick divergence can be written in the standard form as follows:

$$A = \rho - \lambda\sigma, \quad X = M, \tag{B.37}$$

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_{AB} & 0 \\ 0 & 0 & 0 & I_{AB} \end{bmatrix}, \tag{B.38}$$

$$\Phi(X) = \begin{bmatrix} -M & 0 & 0 & 0 \\ 0 & -\mathrm{T}_B(M) & 0 & 0 \\ 0 & 0 & M & 0 \\ 0 & 0 & 0 & \mathrm{T}_B(M) \end{bmatrix}. \tag{B.39}$$

Setting

$$Y = \begin{bmatrix} Y_1 & 0 & 0 & 0 \\ 0 & Y_2 & 0 & 0 \\ 0 & 0 & Y_3 & 0 \\ 0 & 0 & 0 & Y_4 \end{bmatrix}, \tag{B.40}$$

we find that

$$\mathrm{Tr}[\Phi(X)Y] = \mathrm{Tr}[-MY_1 - \mathrm{T}_B(M)Y_2 + MY_3 + T_B(M)Y_4] \tag{B.41}$$

$$= \mathrm{Tr}[M(Y_3 - Y_1) + \mathrm{T}_B(M)(Y_4 - Y_2)] \tag{B.42}$$

$$= \mathrm{Tr}[M\left(Y_3 - Y_1 + \mathrm{T}_B(Y_4 - Y_2)\right)], \tag{B.43}$$

where the last equality holds due to $\mathrm{Tr}[\mathrm{T}_B(L_{AB})S_{AB}] = \mathrm{Tr}[L_{AB}\mathrm{T}_B(S_{AB})]$ with the adjoint of the partial transpose superoperator being the partial transpose superoperator. This leads to

$$\Phi^\dagger(Y) = Y_3 - Y_1 + \mathrm{T}_B(Y_4 - Y_2). \tag{B.44}$$

Strong duality holds due to Slater's conditions by the following choice of feasible and strictly feasible solutions: choose $Y_3 = (\rho - \gamma\sigma)_+$, $Y_i = 0$ for all $i \in \{1, 2, 4\}$ as a feasible solution for the dual SDP and choose $M = (1 - \delta)I_{AB}$ with $\delta \in (0, 1)$ as a strictly feasible solution to the primal SDP. Together with the strong duality, we conclude the proof.

B.5 Werner States

Here, we analyze the measured hockey-stick divergence between two Werner states. We first obtain a simpler expression for the measured hockey-stick divergence between two Werner states using the symmetries of Werner states, which we state in Lemma 10. We then use the statement of Lemma 10 to prove Propositions 15 and 16.

Let Π_{AB}^{sym} and Π_{AB}^{asym} be the projections onto the symmetric and antisymmetric subspaces, respectively. These projections can be written in terms of the identity and swap operators as follows:

$$\Pi_{AB}^{\text{sym}} := \frac{1}{2}(I_{AB} + F_{AB}), \tag{B.45}$$

$$\Pi_{AB}^{\text{asym}} := \frac{1}{2}(I_{AB} - F_{AB}). \tag{B.46}$$

Note that the states Θ_{AB} and $\Theta_{AB}^{\perp}$ can be obtained by normalizing the aforementioned projectors. That is,

$$\Theta_{AB} = \frac{\Pi_{AB}^{\text{sym}}}{\text{Tr}\left[\Pi_{AB}^{\text{sym}}\right]} = \frac{2}{d(d+1)}\Pi_{AB}^{\text{sym}}, \tag{B.47}$$

$$\Theta_{AB}^{\perp} = \frac{\Pi_{AB}^{\text{asym}}}{\text{Tr}\left[\Pi_{AB}^{\text{asym}}\right]} = \frac{2}{d(d-1)}\Pi_{AB}^{\text{asym}}. \tag{B.48}$$

Lemma 10. *Fix $p, q \in [0,1]$. For $\gamma \geq 1$, the measured hockey-stick divergence between two Werner states, ω_{AB}^{q} and ω_{AB}^{p}, is equal to the following:*

$$E_{\gamma}^{\mathcal{M}}\left(\omega_{AB}^{q} \| \omega_{AB}^{p}\right) = \sup_{M \in \mathcal{M}_2} \left[\frac{\operatorname{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right]}{\operatorname{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]}(q - \gamma p) + \frac{\operatorname{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\operatorname{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}(1 - q - \gamma(1-p)) \right], \tag{B.49}$$

where Π_{AB}^{sym} and Π_{AB}^{asym} are the projections onto the symmetric and antisymmetric subspaces, respectively.

Proof. Consider the following twirling channel:

$$\mathcal{T}_{AB}(\cdot) \coloneqq \int dU \, (U_A \otimes U_B)(\cdot)\,(U_A \otimes U_B)^{\dagger}. \tag{B.50}$$

The action of this channel on an operator $X_{AB} \in \mathcal{L}(\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}})$ results in an operator of the following form [135]:

$$\mathcal{T}_{AB}(X_{AB}) = \frac{\operatorname{Tr}\left[X_{AB}\Pi_{AB}^{\mathrm{sym}}\right]}{\operatorname{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]}\Pi_{AB}^{\mathrm{sym}} + \frac{\operatorname{Tr}\left[X_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\operatorname{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}\Pi_{AB}^{\mathrm{asym}}. \tag{B.51}$$

It is easily verified that the Werner states are invariant under the action of $\mathcal{T}_{AB}$. Also, note that the Hilbert–Schmidt adjoint of $\mathcal{T}_{AB}$ is $\mathcal{T}_{AB}$ itself.

The measured hockey-stick divergence between two Werner states, ω_{AB}^{q} and

ω_{AB}^{p}, can be written as follows:

$$
\begin{aligned}
&E_{\gamma}^{\mathcal{M}}\left(\omega_{AB}^{q} \| \omega_{AB}^{p}\right) \\
&= \sup_{M \in \mathcal{M}_2} \operatorname{Tr}\left[M_{AB}\left(\omega_{AB}^{q} - \gamma \omega_{AB}^{p}\right)\right] && \text{(B.52)} \\
&= \sup_{M \in \mathcal{M}_2} \operatorname{Tr}\left[M_{AB}\left(\mathcal{T}_{AB}\left(\omega_{AB}^{q}\right) - \gamma \mathcal{T}_{AB}\left(\omega_{AB}^{p}\right)\right)\right] && \text{(B.53)} \\
&= \sup_{M \in \mathcal{M}_2} \operatorname{Tr}\left[\mathcal{T}_{AB}(M_{AB})\left(\omega_{AB}^{q} - \gamma \omega_{AB}^{p}\right)\right], && \text{(B.54)} \\
&= \sup_{M \in \mathcal{M}_2} \frac{\operatorname{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right]}{\operatorname{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]} \operatorname{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\left(\omega_{AB}^{q} - \gamma \omega_{AB}^{p}\right)\right] + \frac{\operatorname{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\operatorname{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]} \operatorname{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\left(\omega_{AB}^{q} - \gamma \omega_{AB}^{p}\right)\right] \\
& && \text{(B.55)} \\
&= \sup_{M \in \mathcal{M}_2} \frac{\operatorname{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right]}{\operatorname{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]}(q - \gamma p) + \frac{\operatorname{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\operatorname{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}(1 - q - \gamma(1-p)), && \text{(B.56)}
\end{aligned}
$$

where the second equality follows from the fact that Werner states are invariant under the action of the twirling channel $\mathcal{T}_{AB}$, the third equality follows from the fact that $\mathcal{T}_{AB}$ is self-adjoint, the penultimate equality follows from (B.51), and the ultimate equality follows from the definition of Werner states along with (B.47) and (B.48). □

Proof of Proposition 15: The statement of Proposition 15 can be obtained from Lemma 10 by considering $\mathcal{M}$ to be the set of all measurements.

For $\gamma \geq 1$, at most one of the two quantities can be positive: $q - \gamma p$ or $1 - q - \gamma(1-p)$. Choosing $\mathcal{M}$ to be the set of all measurements and ignoring the

negative term in (B.49), we arrive at the following:

$$E_\gamma\left(\omega_{AB}^q \| \omega_{AB}^p\right) = \sup_{0\leq M\leq I}\left[\frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]}(q-\gamma p) + \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}(1-q-\gamma(1-p))\right] \tag{B.57}$$

$$\leq \sup_{0\leq M\leq I}\max\left\{\frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]}(q-\gamma p), \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}(1-q-\gamma(1-p))\right\} \tag{B.58}$$

$$\leq \max\left\{0, q-\gamma p, 1-q-\gamma(1-p)\right\}, \tag{B.59}$$

where the last inequality follows from the fact that $M \leq I$ and that $M = 0$ is a valid choice of measurement.

To show that the inequality in (B.59) is saturated, consider Π_{AB}^{sym} to be a specific choice for the measurement operator in (B.57), which implies $E_\gamma\left(\omega_{AB}^q \| \omega_{AB}^p\right) \geq q-\gamma p$. Furthermore, choosing $M_{AB} = \Pi_{AB}^{\mathrm{asym}}$ leads to $E_\gamma\left(\omega_{AB}^q \| \omega_{AB}^p\right) \geq 1-q-\gamma(1-p)$, and choosing $M_{AB} = 0$ leads to $E_\gamma\left(\omega_{AB}^q \| \omega_{AB}^p\right) \geq 0$. Combining the three inequalities, we arrive at the following inequality:

$$E_\gamma\left(\omega_{AB}^q \| \omega_{AB}^p\right) \geq \max\left\{0, q-\gamma p, 1-q-\gamma(1-p)\right\}, \tag{B.60}$$

which completes the proof.

Proof of Proposition 16: Lower bound: Recall that for $\gamma \geq 1$

$$E_\gamma^{\mathrm{LO}^\star}(\rho\|\sigma) \coloneqq \sup_{M\in\mathrm{LO}^\star} \mathrm{Tr}[M(\rho-\gamma\sigma)]. \tag{B.61}$$

Since the measurement operators $M = 0$ and $I - M = I$, trivially belong to the set of LO* operators, we have

$$E_\gamma^{\mathrm{LO}^\star}(\omega^q\|\omega^p) \geq 0. \tag{B.62}$$

Observe that $M = \sum_{i=1}^{d} |i\rangle\langle i| \otimes |i\rangle\langle i|$ and $I - M = \sum_{i \neq j} |i\rangle\langle i| \otimes |j\rangle\langle j| = I - \sum_{i=1}^{d} |i\rangle\langle i| \otimes |i\rangle\langle i|$ belong to the set of LO* operators as we can see by the following argument: This measurement can be implemented by Alice and Bob first measuring their local systems in the computational basis. Then they perform a classical post-processing of their measurement outcomes by accepting if the outcomes match, which corresponds to the measurement operator $\sum_{i=1}^{d} |i\rangle\langle i|_A \otimes |i\rangle\langle i|_A$. and rejecting if the outcomes do not match, which corresponds to the measurement operator $I_{AB} - \sum_{i=1}^{d} |i\rangle\langle i|_A \otimes |i\rangle\langle i|_A$. With the former, we obtain the inequality

$$E_\gamma^{\mathrm{LO}^\star}(\omega^q \| \omega^p) \geq \mathrm{Tr}\left[\sum_{i=1}^{d} |i\rangle\langle i| \otimes |i\rangle\langle i| \left((q - \gamma p)\Theta + ((1-q) - \gamma(1-p))\,\Theta^\perp\right)\right] \tag{B.63}$$

$$= (q - \gamma p) \times \frac{2}{d+1} + ((1-q) - \gamma(1-p)) \times 0 \tag{B.64}$$

$$= \frac{2(q - \gamma p)}{d+1}, \tag{B.65}$$

where the first equality follows by $\mathrm{Tr}[M\Theta] = 2/(d+1)$ and $\mathrm{Tr}[M\Theta^\perp] = 0$.

Furthermore, with the choice $I - M = \sum_{i \neq j} |i\rangle\langle i| \otimes |j\rangle\langle j|$, we arrive at another lower bound as follows:

$$E_\gamma^{\mathrm{LO}^\star}(\omega^q \| \omega^p) \geq \mathrm{Tr}\left[\left(I - \sum_{i=1}^{d} |i\rangle\langle i| \otimes |i\rangle\langle i|\right)\left((q - \gamma p)\Theta + ((1-q) - \gamma(1-p))\,\Theta^\perp\right)\right] \tag{B.66}$$

$$= 1 - \gamma - \frac{2(q - \gamma p)}{d+1}, \tag{B.67}$$

where the equality follows from (B.65).

Combining (B.62), (B.65), and (B.67), we obtain the following inequality:

$$E_\gamma^{\mathrm{LO}^\star}(\omega^q \| \omega^p) \geq \max\left\{0, \frac{2(q - \gamma p)}{(d+1)}, (1-\gamma) - \frac{2(q - \gamma p)}{(d+1)}\right\}. \tag{B.68}$$

Upper bound: Now we will show that the lower bound on $E_\gamma^{\mathrm{LO}^\star}(\omega^q \| \omega^p)$ obtained in (B.68) is also an upper bound on $E_\gamma^{\mathrm{PPT}}(\omega^q \| \omega^p)$.

Let M_{AB} be a PPT measurement operator; that is, $0 \leq M_{AB} \leq I_{AB}$ and $0 \leq \mathrm{T}_B(M_{AB}) \leq I_{AB}$. From Lemma 10, we know that

$$E_\gamma^{\mathrm{PPT}}\left(\omega_{AB}^q \| \omega_{AB}^p\right) = \sup_{\substack{0 \leq M \leq I \\ 0 \leq \mathrm{T}_B(M) \leq I}} \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]}(q - \gamma p) + \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}(1 - q - \gamma(1-p)). \tag{B.69}$$

The condition $0 \leq M_{AB} \leq I_{AB}$ implies that

$$0 \leq \mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right] \leq \mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right], \tag{B.70}$$

$$0 \leq \mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right] \leq \mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]. \tag{B.71}$$

The partial transpose of the swap operator is the unnormalized maximally entangled vector. That is, $\mathrm{T}_B(F_{AB}) = |\Gamma\rangle\langle\Gamma|_{AB}$, where

$$|\Gamma\rangle_{AB} \coloneqq \sum_{i=0}^{d-1} |i\rangle_A |i\rangle_B, \tag{B.72}$$

where d is the dimension of systems A and B. We will use the notation $\Gamma_{AB} \coloneqq |\Gamma\rangle\langle\Gamma|_{AB}$ for conciseness. Now consider the following identity:

$$\mathrm{T}_B\left(\Pi_{AB}^{\mathrm{sym}} - \Pi_{AB}^{\mathrm{asym}}\right) = \mathrm{T}_B(F_{AB}) = \Gamma_{AB}. \tag{B.73}$$

The PPT condition for M_{AB} implies the following inequality:

$$0 \leq \mathrm{T}_B(M_{AB}) \leq I_{AB} \tag{B.74}$$

$$\implies 0 \leq \langle\Gamma|\mathrm{T}_B(M_{AB})|\Gamma\rangle \leq \langle\Gamma|\Gamma\rangle \tag{B.75}$$

$$\implies 0 \leq \mathrm{Tr}[T_B(M_{AB})\,\Gamma_{AB}] \leq d \tag{B.76}$$

$$\implies 0 \leq \mathrm{Tr}\left[T_B(M_{AB})\,\mathrm{T}_B\left(\Pi_{AB}^{\mathrm{sym}} - \Pi_{AB}^{\mathrm{asym}}\right)\right] \leq d \tag{B.77}$$

$$\implies 0 \leq \mathrm{Tr}\left[M_{AB}\left(\Pi_{AB}^{\mathrm{sym}} - \Pi_{AB}^{\mathrm{asym}}\right)\right] \leq d \tag{B.78}$$

$$\implies \mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right] \leq \mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right] \leq \mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right] + d, \tag{B.79}$$

where the fourth line follows from (B.73) and the fifth line follows from the fact that the partial transpose map is the Hilbert–Schmidt adjoint of itself as well as the inverse of itself.

Let us first assume that $q - \gamma p \geq 0$. We can rewrite (B.69) as follows:

$$
\begin{aligned}
&E_\gamma^{\mathrm{PPT}}\left(\omega_{AB}^q \| \omega_{AB}^p\right) \\
&= \sup_{\substack{0\leq M\leq I \\ 0\leq \mathrm{T}_B(M)\leq I}} \left(\frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]} - \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]} \right)(q-\gamma p) + \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}(1-\gamma) \quad \text{(B.80)} \\
&\leq \sup_{\substack{0\leq M\leq I \\ 0\leq \mathrm{T}_B(M)\leq I}} \left(\frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]} - \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]} \right)(q-\gamma p) + \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}(1-\gamma) \quad \text{(B.81)} \\
&\leq \sup_{\substack{0\leq M\leq I \\ 0\leq \mathrm{T}_B(M)\leq I}} \frac{d}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]}(q-\gamma p) + \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}(1-\gamma) \quad \text{(B.82)} \\
&\leq \frac{2(q-\gamma p)}{d+1}, \quad \text{(B.83)}
\end{aligned}
$$

where the first inequality follows by modifying the denominator of the second term because $\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right] \geq \mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]$, the second inequality follows from (B.78), and the final inequality follows by substituting the value of $\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]$ and ignoring the second term in the penultimate inequality because it is negative.

Now let us assume that $1 - q - \gamma(1-p) \geq 0$, which also implies that $q - \gamma p \leq 0$

since $\gamma \geq 1$. We can rewrite (B.80) as follows:

$$
\begin{aligned}
&E_\gamma^{\mathrm{PPT}}\left(\omega_{AB}^q \| \omega_{AB}^p\right) \\
&= \sup_{\substack{0\leq M\leq I \\ 0\leq \mathrm{T}_B(M)\leq I}} \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}\left(\left(\frac{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}\frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]} - 1\right)(q-\gamma p) + 1 - \gamma\right) && \text{(B.84)} \\
&= \sup_{\substack{0\leq M\leq I \\ 0\leq \mathrm{T}_B(M)\leq I}} \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}\left(\frac{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]}\frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{sym}}\right]}{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}(q-\gamma p) + 1 - q - \gamma(1-p)\right) && \text{(B.85)} \\
&\leq \sup_{\substack{0\leq M\leq I \\ 0\leq \mathrm{T}_B(M)\leq I}} \frac{\mathrm{Tr}\left[M_{AB}\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}\left(\frac{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]}(q-\gamma p) + 1 - q - \gamma(1-p)\right) && \text{(B.86)} \\
&\leq \max\left\{\frac{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]}{\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]}(q-\gamma p) + 1 - q - \gamma(1-p), 0\right\} && \text{(B.87)} \\
&= \max\left\{1 - \gamma - \frac{2(q-\gamma p)}{d+1}, 0\right\}, && \text{(B.88)}
\end{aligned}
$$

where the first inequality follows from (B.79) and the fact that $q-\gamma p \leq 0$, the second inequality follows from (B.71) and the fact that $M = 0$ is a valid choice, and the final equality follows by substituting the values of $\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{sym}}\right]$ and $\mathrm{Tr}\left[\Pi_{AB}^{\mathrm{asym}}\right]$.

For the last case, if both $q - \gamma p \leq 0$ and $1 - q - \gamma(1-p) \leq 0$, then the optimal choice is $M_{AB} = 0$ leading to $E^{\mathrm{PPT}_\gamma}\left(\omega_{AB}^q \| \omega_{AB}^p\right) = 0$. Therefore, the inequalities in (B.83) and (B.88) imply that

$$
E_\gamma^{\mathrm{PPT}}\left(\omega_{AB}^q \| \omega_{AB}^p\right) \leq \max\left\{0, \frac{2(q-\gamma p)}{d+1}, 1 - \gamma - \frac{2(q-\gamma p)}{d+1}\right\}. \tag{B.89}
$$

Recall that

$$
E_\gamma^{\mathrm{LO}^\star}\left(\omega_{AB}^q \| \omega_{AB}^p\right) \leq E_\gamma^{\mathrm{1W\text{-}LOCC}}\left(\omega_{AB}^q \| \omega_{AB}^p\right) \leq E_\gamma^{\mathrm{LOCC}}\left(\omega_{AB}^q \| \omega_{AB}^p\right) \leq E_\gamma^{\mathrm{PPT}}\left(\omega_{AB}^q \| \omega_{AB}^p\right). \tag{B.90}
$$

Therefore, the inequalities in (B.68) and (B.89) imply that all the aforementioned quanitities are equal for $\gamma \geq 1$.

B.6 Isotropic States

In this appendix, we analyze the measured hockey-stick divergence between two isotropic states. We first obtain a simpler expression for the measured hockey-stick divergence between two isotropic states using the symmetries of isotropic states, which we state in Lemma 11. We then use the statement of Lemma 11 to prove Propositions 37 and 38. We define a d-dimensional isotropic state as follows: for $p \in [0,1]$

$$\zeta^p := p\,\Phi + (1-p)\Phi^{\perp}, \tag{B.91}$$

where

$$\Phi := \frac{1}{d}\sum_{i,j=1}^{d} |i\rangle\langle j| \otimes |i\rangle\langle j| \quad \text{and} \quad \Phi^{\perp} := \frac{I_{AB} - \Phi}{d^2-1}. \tag{B.92}$$

Note that Φ_{AB} and $I_{AB} - \Phi_{AB}$ are orthogonal projections.

Lemma 11. *Fix $p, q \in [0,1]$. The following equality holds for all $\gamma \geq 1$:*

$$E_\gamma^{\mathcal{M}}\left(\omega_{AB}^q \| \omega_{AB}^p\right) = \sup_{M \in \mathcal{M}} \mathrm{Tr}[M_{AB}\Phi_{AB}]\,(q - \gamma p) + \frac{\mathrm{Tr}[M_{AB}(I_{AB} - \Phi_{AB})]}{d^2-1}(1 - q - \gamma(1-p)). \tag{B.93}$$

Proof. Consider the following twirling channel:

$$\widetilde{\mathcal{T}}_{AB}(\cdot) := \int dU (U_A \otimes \overline{U}_B)(\cdot)(U_A \otimes \overline{U}_B)^\dagger, \tag{B.94}$$

where $\overline{U}$ denotes the complex conjugate of U. The action of $\widetilde{\mathcal{T}}_{AB}$ on an arbitrary $X_{AB} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ results in an operator of the following form [66]:

$$\widetilde{\mathcal{T}}_{AB}(X_{AB}) = \mathrm{Tr}[X_{AB}\Phi_{AB}]\,\Phi_{AB} + \frac{\mathrm{Tr}[X_{AB}(I_{AB} - \Phi_{AB})]}{d^2-1}\,(I_{AB} - \Phi_{AB})\,. \tag{B.95}$$

It can be easily verified that isotropic states remain invariant under the action of $\widetilde{\mathcal{T}}_{AB}$. Also note that $\widetilde{\mathcal{T}}_{AB}$ is the Hilbert–Schmidt adjoint of itself. The measured

hockey-stick divergence between two isotropic states can then be calculated as follows:

$$E_\gamma^{\mathcal{M}}\left(\zeta_{AB}^q \| \zeta_{AB}^p\right) = \sup_{M\in\mathcal{M}} \operatorname{Tr}\left[M_{AB}\left(\zeta_{AB}^q - \gamma\zeta_{AB}^p\right)\right] \tag{B.96}$$

$$= \sup_{M\in\mathcal{M}} \operatorname{Tr}\left[M_{AB}\left(\widetilde{\mathcal{T}}_{AB}\left(\zeta_{AB}^q\right) - \gamma\widetilde{\mathcal{T}}_{AB}\left(\zeta_{AB}^p\right)\right)\right] \tag{B.97}$$

$$= \sup_{M\in\mathcal{M}} \operatorname{Tr}\left[\widetilde{\mathcal{T}}_{AB}(M_{AB})\left(\zeta_{AB}^q - \gamma\zeta_{AB}^p\right)\right] \tag{B.98}$$

$$= \sup_{M\in\mathcal{M}} \operatorname{Tr}[M_{AB}\Phi_{AB}]\,(q-\gamma p) + \frac{\operatorname{Tr}[M_{AB}(I_{AB}-\Phi_{AB})]}{d^2-1}(1-q-\gamma(1-p)), \tag{B.99}$$

where the second equality follows from the invariance of isotropic states under the twirling channel $\widetilde{\mathcal{T}}_{AB}$, the third equality follows from the fact that $\widetilde{\mathcal{T}}_{AB}$ is self-adjoint, and the final equality follows from (B.95) and the definition of isotropic states in (B.91). □

Proposition 37 (Hockey-Stick Divergence for Isotropic States)**.** *Let* $p, q \in [0,1]$. *We have that for* $\gamma \geq 1$

$$E_\gamma(\zeta^q \| \zeta^p) = \max\{0, q-\gamma p, (1-q)-\gamma(1-p)\}. \tag{B.100}$$

Proof. For $\gamma \geq 1$, at most one of the two quantities can be positive: $q-\gamma p$ or $1-q-\gamma(1-p)$. The equality in (B.93) leads to the following:

$$E_\gamma\left(\zeta_{AB}^q \| \zeta_{AB}^p\right)$$

$$= \sup_{0\leq M\leq I} \operatorname{Tr}[M_{AB}\Phi_{AB}]\,(q-\gamma p) + \frac{\operatorname{Tr}[M_{AB}(I_{AB}-\Phi_{AB})]}{d^2-1}(1-q-\gamma(1-p)) \tag{B.101}$$

$$\leq \sup_{0\leq M\leq I} \max\left\{\operatorname{Tr}[M_{AB}\Phi_{AB}]\,(q-\gamma p), \frac{\operatorname{Tr}[M_{AB}(I_{AB}-\Phi_{AB})]}{d^2-1}(1-q-\gamma(1-p))\right\} \tag{B.102}$$

$$\leq \max\{0, q-\gamma p, 1-q-\gamma(1-p)\}, \tag{B.103}$$

where the first inequality follows by ignoring the negative term in (B.101) and the last inequality follows from the fact that $0 \leq M_{AB} \leq I_{AB}$.

To show that the inequality in (B.103) is saturated, consider Φ_{AB} to be a specific choice for the measurement operator in (B.101), which implies $E_\gamma\left(\zeta^q_{AB}\|\zeta^p_{AB}\right) \geq q-\gamma p$. Furthermore, choosing $M_{AB} = I_{AB}-\Phi_{AB}$ leads to $E_\gamma\left(\zeta^q_{AB}\|\zeta^p_{AB}\right) \geq 1-q-\gamma(1-p)$, and choosing $M_{AB} = 0$ leads to $E_\gamma\left(\zeta^q_{AB}\|\zeta^p_{AB}\right) \geq 0$. Combining the three inequalities, we arrive at the following inequality:

$$E_\gamma\left(\zeta^q_{AB}\|\zeta^p_{AB}\right) \geq \max\left\{0, q-\gamma p, 1-q-\gamma(1-p)\right\}, \tag{B.104}$$

which completes the proof. □

Proposition 38 (Measured Hockey-Stick Divergence for Isotropic States)**.** *Let $p, q \in [0,1]$. We have the following equality for $\gamma \geq 1$ and $\mathcal{M} \in \{\mathcal{M}_{\text{LO}^\star}, \mathcal{M}_{\text{1W-LOCC}}, \mathcal{M}_{\text{LOCC}}, \mathcal{M}_{\text{PPT}}\}$:*

$$E_\gamma^{\mathcal{M}}(\zeta^q\|\zeta^p) = \max\left\{0, q-\gamma p + \frac{(1-q)-\gamma(1-p)}{d+1}, \frac{d}{d+1}\left((1-q)-\gamma(1-p)\right)\right\}. \tag{B.105}$$

Proof. Lower bound: Recall that for $\gamma \geq 1$

$$E_\gamma^{\text{LO}^\star}(\rho\|\sigma) \coloneqq \sup_{M\in\text{LO}^\star} \text{Tr}[M(\rho-\gamma\sigma)]. \tag{B.106}$$

Since the measurement operators $M = 0$ and $I - M = I$, trivially belong to the set of local operators, we have

$$E_\gamma^{\text{LO}^\star}(\zeta^q\|\zeta^p) \geq 0. \tag{B.107}$$

Observe that $M = \sum_{i=1}^d |i\rangle\langle i|\otimes|i\rangle\langle i|$ and $I-M = \sum_{i\neq j}|i\rangle\langle i|\otimes|j\rangle\langle j| = I-\sum_{i=1}^d|i\rangle\langle i|\otimes|i\rangle\langle i|$

belong to the set LO* operators. With the former, we arrive at the inequality

$$
\begin{aligned}
&E_\gamma^{\mathrm{LO}^\star}(\zeta^q\|\zeta^p) \\
&\geq \mathrm{Tr}\left[\sum_{i=1}^{d} |i\rangle\langle i| \otimes |i\rangle\langle i| \left((q-\gamma p)\Phi + ((1-q)-\gamma(1-p))\,\Phi^\perp\right)\right] && \text{(B.108)}\\
&= (q-\gamma p)\,\mathrm{Tr}\left[\left(\sum_{i=1}^{d} |i\rangle\langle i|_A \otimes |i\rangle\langle i|_B\right)\Phi_{AB}\right] + ((1-q)-\gamma(1-p))\,\mathrm{Tr}\left[\left(\sum_{i=1}^{d} |i\rangle\langle i|_A \otimes |i\rangle\langle i|_B\right)\Phi_{AB}^\perp\right] && \text{(B.109)}\\
&= (q-\gamma p) + ((1-q)-\gamma(1-p))\,\frac{1}{d+1}, && \text{(B.110)}
\end{aligned}
$$

where the first equality follows from the linearity of the trace operator and by substituting (B.92).

Furthermore, with the choice $I - M = \sum_{i\neq j} |i\rangle\langle i| \otimes |j\rangle\langle j|$, we arrive at another lower bound as follows:

$$
\begin{aligned}
E_\gamma^{\mathrm{LO}^\star}(\zeta^q\|\zeta^p) &\geq \mathrm{Tr}\left[\left(I - \sum_{i=1}^{d} |i\rangle\langle i| \otimes |i\rangle\langle i|\right)\left((q-\gamma p)\Phi + ((1-q)-\gamma(1-p))\,\Phi^\perp\right)\right] && \text{(B.111)}\\
&= 1-\gamma-\left((q-\gamma p) + ((1-q)-\gamma(1-p))\,\frac{1}{d+1}\right) && \text{(B.112)}\\
&= \frac{d}{d+1}\left((1-q)-\gamma(1-p)\right), && \text{(B.113)}
\end{aligned}
$$

where the first equality follows from (B.110).

Combining (B.107),(B.110) and (B.113), we obtain that

$$
E_\gamma^{\mathrm{LO}^\star}(\zeta^q\|\zeta^p) \geq \max\left\{0, q-\gamma p + \frac{1}{d+1}(1-q-\gamma(1-p)), \frac{d}{d+1}\left((1-q)-\gamma(1-p)\right)\right\}. \tag{B.114}
$$

Upper bound: Let us first consider the following inequalities that hold for

every PPT measurement operator:

$$0 \leq \mathrm{T}_B(M_{AB}) \leq I_{AB} \tag{B.115}$$

$$\implies 0 \leq \frac{1}{2}\operatorname{Tr}[\mathrm{T}_B(M_{AB})(I_{AB} - F_{AB})] \leq \frac{1}{2}\operatorname{Tr}[I_{AB} - F_{AB}] \tag{B.116}$$

$$\implies 0 \leq \frac{1}{2}\operatorname{Tr}[M_{AB}\mathrm{T}_B(I_{AB} - F_{AB})] \leq \frac{1}{2}\operatorname{Tr}[I_{AB} - F_{AB}] \tag{B.117}$$

$$\implies 0 \leq \frac{1}{2}\operatorname{Tr}[M_{AB}(I_{AB} - d\Phi_{AB})] \leq \frac{1}{2}(d^2 - d) \tag{B.118}$$

$$\implies 0 \leq \operatorname{Tr}[M_{AB}(I_{AB} - \Phi_{AB} + (1-d)\Phi_{AB})] \leq d^2 - d \tag{B.119}$$

$$\implies 0 \leq \operatorname{Tr}[M_{AB}(I_{AB} - \Phi_{AB})] - (d-1)\operatorname{Tr}[M_{AB}\Phi_{AB}] \leq d(d-1) \tag{B.120}$$

$$\implies \frac{1}{d+1} \leq \frac{1}{d^2-1} \times \frac{\operatorname{Tr}[M_{AB}(I_{AB} - \Phi_{AB})]}{\operatorname{Tr}[M_{AB}\Phi_{AB}]} \leq \frac{1}{d+1} + \frac{d}{(d+1)\operatorname{Tr}[M_{AB}\Phi_{AB}]}, \tag{B.121}$$

where F_{AB} is the swap operator defined just after (4.20). In the set of inequalities mentioned above, the second line follows from the fact that $\frac{1}{2}(I_{AB} - F_{AB})$ is a projector, the third line follows from the fact that the partial transpose is a self-adjoint map, the fourth line follows from the fact that the partial transpose of the swap operator is the unnormalized maximally entangled vector, and the last line follows by rearranging the terms.

Recall the equality in (B.93), which leads to the following:

$$E_\gamma^{\mathrm{PPT}}\left(\zeta_{AB}^q \| \zeta_{AB}^p\right)$$

$$= \sup_{\substack{0 \leq M \leq I, \\ 0 \leq \mathrm{T}_B(M) \leq I}} \operatorname{Tr}[M_{AB}\Phi_{AB}]\,(q - \gamma p) + \frac{\operatorname{Tr}[M_{AB}(I_{AB} - \Phi_{AB})]}{d^2-1}(1 - q - \gamma(1-p)) \tag{B.122}$$

$$= \sup_{\substack{0 \leq M \leq I, \\ 0 \leq \mathrm{T}_B(M) \leq I}} \operatorname{Tr}[M_{AB}\Phi_{AB}]\left(q - \gamma p + \frac{\operatorname{Tr}[M_{AB}(I_{AB} - \Phi_{AB})]}{(d^2-1)\operatorname{Tr}[M_{AB}\Phi_{AB}]}(1 - q - \gamma(1-p))\right). \tag{B.123}$$

Consider the case when $q - \gamma p \geq 0$, which implies that $1 - q - \gamma(1-p) \leq 0$. The inequality in (B.121) yields the following inequality:

$$E_\gamma^{\mathrm{PPT}}\left(\zeta_{AB}^q \| \zeta_{AB}^p\right) \leq \sup_{\substack{0 \leq M \leq I, \\ 0 \leq \mathrm{T}_B(M) \leq I}} \operatorname{Tr}[M_{AB}\Phi_{AB}]\left(q - \gamma p + \frac{1}{d+1}(1 - q - \gamma(1-p))\right) \tag{B.124}$$

$$\leq \max\left\{q - \gamma p + \frac{1}{d+1}(1 - q - \gamma(1-p)), 0\right\}, \tag{B.125}$$

where the last inequality follows from the fact that $M_{AB} \leq I_{AB}$ and $M = 0$ is a valid choice.

Now consider the case when $1 - q - \gamma(1-p) \geq 0$, which implies that $q - \gamma p \leq 0$. The equality in (B.123) and the inequality in (B.121) yield the following inequality:

$$\begin{aligned}
&E_\gamma^{\mathrm{PPT}}\left(\zeta_{AB}^q \| \zeta_{AB}^p\right) \\
&\leq \sup_{\substack{0\leq M\leq I,\\ 0\leq \mathrm{T}_B(M)\leq I}} \mathrm{Tr}[M_{AB}\Phi_{AB}]\left(q - \gamma p + \left(\frac{1}{d+1} + \frac{d}{(d+1)\,\mathrm{Tr}[M_{AB}\Phi_{AB}]}\right)(1 - q - \gamma(1-p))\right) \qquad \text{(B.126)}\\
&= \sup_{\substack{0\leq M\leq I,\\ 0\leq \mathrm{T}_B(M)\leq I}} \mathrm{Tr}[M_{AB}\Phi_{AB}] \times \\
&\quad \left(\left(1 - \frac{1}{d+1}\right)(q - \gamma p) - \frac{1}{d+1}(\gamma - 1) + \frac{d}{(d+1)\,\mathrm{Tr}[M_{AB}\Phi_{AB}]}(1 - q - \gamma(1-p))\right) \qquad \text{(B.127)}\\
&\leq \sup_{\substack{0\leq M\leq I,\\ 0\leq \mathrm{T}_B(M)\leq I}} \mathrm{Tr}[M_{AB}\Phi_{AB}]\left(\frac{d}{(d+1)\,\mathrm{Tr}[M_{AB}\Phi_{AB}]}(1 - q - \gamma(1-p))\right) \qquad \text{(B.128)}\\
&= \frac{d}{d+1}(1 - q - \gamma(1-p)), \qquad \text{(B.129)}
\end{aligned}$$

where we arrived at the last inequality by ignoring the non-positive terms.

Finally, if both $q - \gamma p < 0$ and $1 - q - \gamma(1-p) < 0$, then the optimal choice is $M_{AB} = 0$, which leads to $E_\gamma^{\mathrm{PPT}}\left(\zeta_{AB}^q \| \zeta_{AB}^p\right) = 0$. Combining this fact with (B.125) and (B.129) leads to the following inequality:

$$E_\gamma^{\mathrm{PPT}}\left(\zeta_{AB}^q \| \zeta_{AB}^p\right) \leq \max\left\{0, q - \gamma p + \frac{1}{d+1}(1 - q - \gamma(1-p)), \frac{d}{d+1}(1 - q - \gamma(1-p))\right\}. \qquad \text{(B.130)}$$

Recall that the right-hand side of the above inequality is a lower bound on $E_\gamma^{\mathrm{LO}^\star}\left(\zeta_{AB}^q \| \zeta_{AB}^p\right)$ in (B.114). Since

$$E_\gamma^{\mathrm{LO}^\star}\left(\zeta_{AB}^q \| \zeta_{AB}^p\right) \leq E_\gamma^{\mathrm{1W\text{-}LOCC}}\left(\zeta_{AB}^q \| \zeta_{AB}^p\right) \leq E_\gamma^{\mathrm{LOCC}}\left(\zeta_{AB}^q \| \zeta_{AB}^p\right) \leq E_\gamma^{\mathrm{PPT}}\left(\zeta_{AB}^q \| \zeta_{AB}^p\right), \qquad \text{(B.131)}$$

we conclude that all of the above quantities are equal to each other, and their value is given by $\max\left\{0, q - \gamma p + \frac{1}{d+1}(1 - q - \gamma(1-p)), \frac{d}{d+1}(1 - q - \gamma(1-p))\right\}$. □

Remark 43 (High Dimensions)**.** *When $d \to \infty$, we see that the quantity in Proposition 38 is equal to the quantity in Proposition 37.*

B.7 Channel Divergence with All Possible Measurements

Let $\bar{\mathcal{M}}$ denote the set of all measurements. Then,

$$E_\gamma(\mathcal{P}\|\mathcal{Q}) \equiv E_\gamma^{\bar{\mathcal{M}}}(\mathcal{P}\|\mathcal{Q}). \tag{B.132}$$

Proposition 39 (Properties of Hockey-Stick Divergence for Channels)**.** *The hockey-stick divergence for channels satisfies the following properties:*

1. *Quasi Sub-Additivity: Let $\mathcal{P}_i$ and $\mathcal{Q}_i$ for $i \in \{0,1\}$ be channels such that $\mathcal{P}_0$ and $\mathcal{Q}_0$ are linear mappings from A to A' and $\mathcal{P}_1$ and $\mathcal{Q}_1$ are linear mappings from A' to B. Also, let $\gamma_1, \gamma_2 \geq 1$. Then,*

$$\begin{aligned} &E_{\gamma_1\gamma_2}(\mathcal{P}_1 \circ \mathcal{P}_0 \| \mathcal{Q}_1 \circ \mathcal{Q}_0) \\ &\leq \min\left\{E_{\gamma_1}(\mathcal{P}_0\|\mathcal{Q}_0) + \gamma_1 E_{\gamma_2}(\mathcal{P}_1\|\mathcal{Q}_1), E_{\gamma_1}(\mathcal{P}_1\|\mathcal{Q}_1) + \gamma_1 E_{\gamma_2}(\mathcal{P}_0\|\mathcal{Q}_0)\right\}. \end{aligned} \tag{B.133}$$

Proof. Quasi sub-additivity: Let $\mathcal{P}_i$ and $\mathcal{Q}_i$ for $i \in \{0,1\}$ be channels such that $\mathcal{P}_0$ and $\mathcal{Q}_0$ are linear mappings from A to A' and $\mathcal{P}_1$ and $\mathcal{Q}_1$ are linear mappings from A' to B. Also, R is a reference system isomorphic to A.

Consider that

$$
\begin{aligned}
&E_{\gamma_1\gamma_2}(\mathcal{P}_1 \circ \mathcal{P}_0 \| \mathcal{Q}_1 \circ \mathcal{Q}_0) \\
&= \sup_{\rho_{RA}} E_{\gamma_1\gamma_2}(\mathcal{P}_1 \circ \mathcal{P}_0(\rho_{RA}) \| \mathcal{Q}_1 \circ \mathcal{Q}_0(\rho_{RA})) && \text{(B.134)} \\
&\leq \sup_{\rho_{RA}} E_{\gamma_1}(\mathcal{P}_1 \circ \mathcal{P}_0(\rho_{RA}) \| \mathcal{P}_1 \circ \mathcal{Q}_0(\rho_{RA})) + \gamma_1 E_{\gamma_2}(\mathcal{P}_1 \circ \mathcal{Q}_0(\rho_{RA}) \| \mathcal{Q}_1 \circ \mathcal{Q}_0(\rho_{RA})) && \text{(B.135)} \\
&\leq \sup_{\rho_{RA}} E_{\gamma_1}(\mathcal{P}_0(\rho_{RA}) \| \mathcal{Q}_0(\rho_{RA})) + \gamma_1 E_{\gamma_2}(\mathcal{P}_1 \circ \mathcal{Q}_0(\rho_{RA}) \| \mathcal{Q}_1 \circ \mathcal{Q}_0(\rho_{RA})) && \text{(B.136)} \\
&\leq \sup_{\rho_{RA}} E_{\gamma_1}(\mathcal{P}_0(\rho_{RA}) \| \mathcal{Q}_0(\rho_{RA})) + \gamma_1 \sup_{\rho_{RA}} E_{\gamma_2}(\mathcal{P}_1 \circ \mathcal{Q}_0(\rho_{RA}) \| \mathcal{Q}_1 \circ \mathcal{Q}_0(\rho_{RA})) && \text{(B.137)} \\
&= E_{\gamma_1}(\mathcal{P}_0 \| \mathcal{Q}_0) + \gamma_1 \sup_{\rho_{RA}} E_{\gamma_2}(\mathcal{P}_1 \circ \mathcal{Q}_0(\rho_{RA}) \| \mathcal{Q}_1 \circ \mathcal{Q}_0(\rho_{RA})) && \text{(B.138)} \\
&\leq E_{\gamma_1}(\mathcal{P}_0 \| \mathcal{Q}_0) + \gamma_1 \sup_{\sigma_{RA'}} E_{\gamma_2}(\mathcal{P}_1(\sigma_{RA'}) \| \mathcal{Q}_1(\sigma_{RA'})) && \text{(B.139)} \\
&= E_{\gamma_1}(\mathcal{P}_0 \| \mathcal{Q}_0) + \gamma_1 E_{\gamma_2}(\mathcal{P}_1 \| \mathcal{Q}_1), && \text{(B.140)}
\end{aligned}
$$

where the first equality follows from the triangular property of the hockey-stick divergence [62, Eq II.16]; the second inequality from the data processing of the hockey-stick divergence [119, Lemma 4].

The second expression can be obtained by choosing $\mathcal{Q}_1 \circ \mathcal{P}_0(\rho_{RA})$ instead of $\mathcal{P}_1 \circ \mathcal{Q}_0(\rho_{RA})$ in the second equality above and proceeding with the similar decompositions and arguments. □

We show that the hockey-stick channel divergence and the PPT-measured hockey-stick channel divergence are SDP computable.

Proposition 40 (SDP for E_γ Channel Divergence)**.** *Let $\mathcal{P}$ and $\mathcal{Q}$ be two quantum channels, and let $\gamma \geq 1$. Then, the channel divergence $E_\gamma(\mathcal{P}\|\mathcal{Q})$ is equivalent to the following expression:*

$$
E_\gamma(\mathcal{P}\|\mathcal{Q}) = \sup_{\Omega_{RB}\geq 0, \rho_R \geq 0} \left\{ \operatorname{Tr}\left[\Omega_{RB}(\Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB})\right] : \operatorname{Tr}[\rho_R] = 1,\ \Omega_{RB} \leq \rho_R \otimes I_B \right\}, \quad \text{(B.141)}
$$

where $\Gamma^{\mathcal{N}}_{RB}$ is the Choi operator of the channel $\mathcal{N}_{A\to B}$.

Furthermore, its dual expression evaluates to the following:

$$E_\gamma(\mathcal{P}\|\mathcal{Q}) = \inf_{\mu\geq 0, Z_{RB}\geq 0} \left\{ \mu : \Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB} \leq Z_{RB},\ \mu I_{RB} \geq \mathrm{Tr}_B[Z_{RB}] \right\}. \tag{B.142}$$

Proof. By the joint-convexity of hockey-stick divergence [62, Proposition II.5] together with the Schmidt decomposition, we have that the supremum in the channel divergence is achieved by pure states. Then, together with the variational form of hockey-stick divergence, we have that

$$E_\gamma(\mathcal{P}\|\mathcal{Q}) = \sup_{\phi_{RA}\in\mathcal{D}} \sup_{0\leq M_{RB}\leq I} \mathrm{Tr}[M_{RB}\left(\mathcal{P}_{A\to B}(\phi_{RA}) - \gamma\mathcal{Q}_{A\to B}(\phi_{RA})\right)], \tag{B.143}$$

where ϕ_{RA} is a pure state with the system R isomorphic to system A.

We can rewrite the above as follows:

$$E_\gamma(\mathcal{P}\|\mathcal{Q}) = \sup_{\substack{\phi_{RA}\geq 0,\\ M_{RB}\geq 0}} \left\{ \mathrm{Tr}[M_{RB}\left(\mathcal{P}_{A\to B}(\phi_{RA}) - \gamma\mathcal{Q}_{A\to B}(\phi_{RA})\right)] : \begin{array}{l} \mathrm{Tr}[\phi_{RA}] = 1,\ \mathrm{Tr}[\phi^2_{RA}] = 1, \\ M_{RB} \leq I \end{array} \right\}, \tag{B.144}$$

where ϕ_{RA} is a pure bipartitie state that satisfies $\mathrm{Tr}[\phi_{RA}] = 1,\ \mathrm{Tr}[\phi^2_{RA}] = 1, \phi_{RA} \geq 0$. Note also that it is equivalent to

$$E_\gamma(\mathcal{P}\|\mathcal{Q}) = \sup_{\substack{\phi_{RA}\geq 0,\\ M_{RB}\geq 0}} \left\{ \mathrm{Tr}[M_{RB}\left(\mathcal{P}_{A\to B}(\phi_{RA}) - \gamma\mathcal{Q}_{A\to B}(\phi_{RA})\right)] : \begin{array}{l} \mathrm{Tr}[\phi_{RA}] = 1,\ \mathrm{Tr}[\phi^2_{RA}] = 1, \\ M_{RB} \leq I,\ \phi_R > 0 \end{array} \right\}, \tag{B.145}$$

since the set of pure states with reduced state ϕ_R positive definite is dense in the set of all pure states. Then, any such pure state can be written as

$$\phi_{RA} = X_R \Gamma_{RA} X_R^\dagger \tag{B.146}$$

for some linear operator X such that $\mathrm{Tr}[X_R^\dagger X_R] = 1$ and $|X_R| > 0$, where Γ_{RA} is the unnormalized maximally entangled operator defined just after (B.72).

Using this, we find that the objective function can be rewritten as

$$\begin{aligned}
&\mathrm{Tr}[M_{RB}\,(\mathcal{P}_{A\to B}(\phi_{RA}) - \gamma\mathcal{Q}_{A\to B}(\phi_{RA}))] \\
&= \mathrm{Tr}\Big[M_{RB}\,(\mathcal{P}_{A\to B} - \gamma\mathcal{Q}_{A\to B})\,(X_R\Gamma_{RA}X_R^\dagger)\Big] &\text{(B.147)}\\
&= \mathrm{Tr}\Big[X_R^\dagger M_{RB}X_R\,(\mathcal{P}_{A\to B} - \gamma\mathcal{Q}_{A\to B})\,(\Gamma_{RA})\Big] &\text{(B.148)}\\
&= \mathrm{Tr}\Big[X_R^\dagger M_{RB}X_R\left(\Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB}\right)\Big]. &\text{(B.149)}
\end{aligned}$$

Also, the following equivalence holds:

$$0 \le M_{RB} \le I_{RB} \iff 0 \le X_R^\dagger M_{RB}X_R \le X_R^\dagger X_R \otimes I_B. \tag{B.150}$$

With that we choose $\Omega_{RB} \coloneqq X_R^\dagger M_{RB}X_R$, $\rho_R \coloneqq X_R^\dagger X_R$ since we have $\mathrm{Tr}[X_R^\dagger X_R] = 1$.

Using the substitutions selected along with (B.145) and (B.149), we arrive at

$$E_\gamma(\mathcal{P}\|\mathcal{Q}) = \sup_{\Omega_{RB}\ge 0,\rho_R\ge 0}\left\{\mathrm{Tr}\Big[\Omega_{RB}(\Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB})\Big] : \mathrm{Tr}[\rho_R] = 1,\ \Omega_{RB} \le \rho_R \otimes I_B\right\}. \tag{B.151}$$

This completes the proof of the primal formulation.

To obtain the dual representation, consider

$$E_\gamma(\mathcal{P}\|\mathcal{Q}) \tag{B.152}$$

$$= \sup_{\Omega_{RB}\geq 0,\rho_R\geq 0} \left\{ \mathrm{Tr}\Big[\Omega_{RB}(\Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB})\Big] + \inf_{\mu\in\mathbb{R},Z_{RB}\geq 0} \mu(1 - \mathrm{Tr}[\rho_R]) + \mathrm{Tr}[Z_{RB}(\rho_R \otimes I_B - \Omega_{RB})] \right\}$$

$$= \sup_{\Omega_{RB}\geq 0,\rho_R\geq 0} \inf_{\mu\in\mathbb{R},Z_{RB}\geq 0} \left\{ \mathrm{Tr}\Big[\Omega_{RB}(\Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB})\Big] + \mu(1 - \mathrm{Tr}[\rho_R]) + \mathrm{Tr}[Z_{RB}(\rho_R \otimes I_B - \Omega_{RB})] \right\} \tag{B.153}$$

$$\leq \inf_{\mu\in\mathbb{R},Z_{RB}\geq 0} \sup_{\Omega_{RB}\geq 0,\rho_R\geq 0} \left\{ \mathrm{Tr}\Big[\Omega_{RB}(\Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB})\Big] + \mu(1 - \mathrm{Tr}[\rho_R]) + \mathrm{Tr}[Z_{RB}(\rho_R \otimes I_B - \Omega_{RB})] \right\} \tag{B.154}$$

$$= \inf_{\mu\in\mathbb{R},Z_{RB}\geq 0} \sup_{\Omega_{RB}\geq 0,\rho_R\geq 0} \left\{ \mu + \mathrm{Tr}\Big[\Omega_{RB}(\Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB} - Z_{RB})\Big] + \mathrm{Tr}[\rho_R\,(-\mu I + \mathrm{Tr}_B[Z_{RB}])] \right\} \tag{B.155}$$

$$= \inf_{\mu\in\mathbb{R},Z_{RB}\geq 0} \left\{ \mu + \sup_{\Omega_{RB}\geq 0,\rho_R\geq 0} \mathrm{Tr}\Big[\Omega_{RB}(\Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB} - Z_{RB})\Big] + \mathrm{Tr}[\rho_R\,(-\mu I + \mathrm{Tr}_B[Z_{RB}])] \right\} \tag{B.156}$$

$$= \inf_{\mu\in\mathbb{R},Z_{RB}\geq 0} \left\{ \mu : \Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB} \leq Z_{RB},\ \mu I \geq \mathrm{Tr}_B[Z_{RB}] \right\} \tag{B.157}$$

$$= \inf_{\mu\geq 0,Z_{RB}\geq 0} \left\{ \mu : \Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB} \leq Z_{RB},\ \mu I \geq \mathrm{Tr}_B[Z_{RB}] \right\}, \tag{B.158}$$

where the last inequality follows since $Z_{RB} \geq 0$ and then $\mu I \geq \mathrm{Tr}_B[Z_{RB}]$ holds only when $\mu \geq 0$.

Now, what is remaining is to show that the strong duality holds. To see this, choose $Z_{RB} = \Gamma^{\mathcal{P}}_{RB} - \gamma\Gamma^{\mathcal{Q}}_{RB} + \delta I_{RB}$ and μ such that $\mu I_R = \mathrm{Tr}_B[Z_{RB}] + \delta I_R$ together form a strictly feasible point for all $\delta > 0$ and a feasible point for the primal (i.e., supremum formulation) is $\rho_R = \pi_R$, which is the maximally mixed state (in the Hilbert space $\mathcal{H}_R$) and $\Omega_{RB} = \pi_R \otimes I_B$. This concludes the proof. □

Proposition 41 (Channel Divergence for Depolarizing Channels)**.** *Let* $p, q \in [0, 1]$

and $\gamma \geq 1$. We have that

$$E_\gamma\left(\mathcal{A}^q_{\text{Dep}} \| \mathcal{A}^p_{\text{Dep}}\right) = \max\left\{0, (1-q) - \gamma(1-p) + \frac{q-\gamma p}{d^2}, q - \gamma p - \frac{q-\gamma p}{d^2}\right\}, \quad \text{(B.159)}$$

where $\mathcal{A}^p_{\text{Dep}}$ is the depolarizing channel with parameter p and d is the dimension of the input space of the channel $\mathcal{A}^p_{\text{Dep}}$.

Proof. Since E_γ satisfies data-processing, it is a generalized divergence. Also, with the fact that depolarizing channels are jointly covariant, using [75, Proposition 7.84], we have that

$$E_\gamma\left(\mathcal{A}^q_{\text{Dep}} \| \mathcal{A}^p_{\text{Dep}}\right) = E_\gamma\left(\frac{\Gamma^{\mathcal{A}^q_{\text{Dep}}}_{RB}}{d} \middle\| \frac{\Gamma^{\mathcal{A}^p_{\text{Dep}}}_{RB}}{d}\right), \quad \text{(B.160)}$$

where $\Gamma^{\mathcal{A}^p_{\text{Dep}}}_{RB}$ is the unnormalized Choi-state of the channel $\mathcal{A}^p_{\text{Dep}}$. By using the fact that

$$\Gamma^{\mathcal{A}^p_{\text{Dep}}}_{RB} = (1-p)\Gamma_{RA} + p\frac{I_{d^2}}{d}, \quad \text{(B.161)}$$

consider

$$\begin{aligned}
\frac{\Gamma^{\mathcal{A}^p_{\text{Dep}}}_{RB}}{d} &= (1-p)\frac{\Gamma_{RA}}{d} + \frac{p}{d^2} I_{d^2} && \text{(B.162)}\\
&= (1-p)\Phi + \frac{p}{d^2}\left(I_{d^2} - \Phi + \Phi\right) && \text{(B.163)}\\
&= \left(1 - p + \frac{p}{d^2}\right)\Phi + \frac{p}{d^2}(d^2-1)\frac{(I_{d^2} - \Phi)}{d^2-1} && \text{(B.164)}\\
&= \zeta^\eta && \text{(B.165)}
\end{aligned}$$

where $\eta := 1 - p + \frac{p}{d^2}$ in (B.91).

With that, by choosing $\eta' := 1 - q + \frac{q}{d^2}$, we arrive at

$$E_\gamma\left(\mathcal{A}^q_{\text{Dep}} \| \mathcal{A}^p_{\text{Dep}}\right) = E_\gamma\left(\zeta^{\eta'} \middle\| \zeta^\eta\right). \quad \text{(B.166)}$$

Then, we conclude the proof by applying Proposition 37. □

B.8 Proof of Proposition 20

The proof of the supremum formulation (primal) follows similarly to the proof of Proposition 40 with the added constraints as given below. For PPT measurements, we require

$$0 \leq \mathrm{T}_B(M_{RB}) \leq I. \tag{B.167}$$

This is equivalent to enforcing

$$0 \leq X_R^\dagger \mathrm{T}_B(M_{RB}) X_R \leq X_R^\dagger X_R \otimes I_B \quad \Longleftrightarrow \quad 0 \leq \mathrm{T}_B(\Omega_{RB}) \leq \rho_R \otimes I_B, \tag{B.168}$$

with the choices $\Omega_{RB} \coloneqq X_R^\dagger M_{RB} X_R$ and $\rho_R \coloneqq X_R^\dagger X_R$.

For the derivation of the dual, consider that

$$E_\gamma^{\mathrm{PPT}}(\mathcal{P}\|\mathcal{Q}) \tag{B.169}$$

$$= \sup_{\Omega_{RB}\geq 0,\rho_R\geq 0} \left\{ \begin{array}{l} \mathrm{Tr}\left[\Omega_{RB}(\Gamma_{RB}^{\mathcal{P}} - \gamma\Gamma_{RB}^{\mathcal{Q}})\right] + \inf\limits_{\substack{\mu\in\mathbb{R},Z_{RB}\geq 0\\ L_{RB},Y_{RB}\geq 0}} \left\{ \begin{array}{l} \mu(1-\mathrm{Tr}[\rho_R]) + \mathrm{Tr}[Z_{RB}(\rho_R\otimes I_B - \Omega_{RB})] \\ + \mathrm{Tr}[L_{RB}\mathrm{T}_B(\Omega_{RB})] \\ + \mathrm{Tr}[Y_{RB}\,(\rho_R\otimes I_B - \mathrm{T}_B(\Omega_{RB}))] \end{array} \right\} \end{array} \right\}$$

$$= \sup_{\Omega_{RB}\geq 0,\rho_R\geq 0} \inf_{\substack{\mu\in\mathbb{R},Z_{RB}\geq 0\\ L_{RB},Y_{RB}\geq 0}} \left\{ \begin{array}{l} \mathrm{Tr}\left[\Omega_{RB}(\Gamma_{RB}^{\mathcal{P}} - \gamma\Gamma_{RB}^{\mathcal{Q}})\right] + \mu(1-\mathrm{Tr}[\rho_R]) + \mathrm{Tr}[Z_{RB}(\rho_R\otimes I_B - \Omega_{RB})] \\ + \mathrm{Tr}[L_{RB}\mathrm{T}_B(\Omega_{RB})] + \mathrm{Tr}[Y_{RB}\,(\rho_R\otimes I_B - \mathrm{T}_B(\Omega_{RB}))] \end{array} \right\} \tag{B.170}$$

$$\leq \inf_{\substack{\mu\in\mathbb{R},Z_{RB}\geq 0\\ L_{RB},Y_{RB}\geq 0}} \sup_{\Omega_{RB}\geq 0,\rho_R\geq 0} \left\{ \begin{array}{l} \mathrm{Tr}\left[\Omega_{RB}(\Gamma_{RB}^{\mathcal{P}} - \gamma\Gamma_{RB}^{\mathcal{Q}})\right] + \mu(1-\mathrm{Tr}[\rho_R]) + \mathrm{Tr}[Z_{RB}(\rho_R\otimes I_B - \Omega_{RB})] \\ + \mathrm{Tr}[L_{RB}\mathrm{T}_B(\Omega_{RB})] + \mathrm{Tr}[Y_{RB}\,(\rho_R\otimes I_B - \mathrm{T}_B(\Omega_{RB}))] \end{array} \right\} \tag{B.171}$$

$$= \inf_{\substack{\mu\in\mathbb{R},Z_{RB}\geq 0\\ L_{RB},Y_{RB}\geq 0}} \left\{ \mu + \sup_{\Omega_{RB}\geq 0,\rho_R\geq 0} \begin{array}{r} \mathrm{Tr}\left[\Omega_{RB}(\Gamma_{RB}^{\mathcal{P}} - \gamma\Gamma_{RB}^{\mathcal{Q}} - Z_{RB} + \mathrm{T}_B(L_{RB}) - \mathrm{T}_B(Y_{RB}))\right] \\ + \mathrm{Tr}[\rho_R\,(-\mu I + \mathrm{Tr}_B[Z_{RB}] + \mathrm{Tr}_B[Y_{RB}])] \end{array} \right\} \tag{B.172}$$

$$= \inf_{\substack{\mu\in\mathbb{R},Z_{RB}\geq 0\\ L_{RB},Y_{RB}\geq 0}} \left\{\mu : \Gamma_{RB}^{\mathcal{P}} - \gamma\Gamma_{RB}^{\mathcal{Q}} \leq Z_{RB} + \mathrm{T}_B(Y_{RB}) - \mathrm{T}_B(L_{RB}),\ \mu I \geq \mathrm{Tr}_B[Z_{RB}] + \mathrm{Tr}_B[Y_{RB}]\right\} \tag{B.173}$$

$$\inf_{\substack{\mu\geq 0,Z_{RB}\geq 0\\ L_{RB},Y_{RB}\geq 0}} \left\{\mu : \Gamma_{RB}^{\mathcal{P}} - \gamma\Gamma_{RB}^{\mathcal{Q}} \leq Z_{RB} + \mathrm{T}_B(Y_{RB}) - \mathrm{T}_B(L_{RB}),\ \mu I \geq \mathrm{Tr}_B[Z_{RB}] + \mathrm{Tr}_B[Y_{RB}]\right\}, \tag{B.174}$$

where the last inequality follows since $Z_{RB}, Y_{RB} \geq 0$ and then $\mu I \geq \mathrm{Tr}_B[Z_{RB}] + \mathrm{Tr}_B[Y_{RB}]$ holds only when $\mu \geq 0$.

Now, what is remaining is to show that the strong duality holds. To see this, choose $Z_{RB} = \Gamma_{RB}^{\mathcal{P}} - \gamma\Gamma_{RB}^{\mathcal{Q}}$, $Y_{RB} = \delta_1 I_{RB}$, $L_{RB} = \delta_2 I_{RB}$ and μ such that $\mu I_R = \mathrm{Tr}_B[Z_{RB}] + \mathrm{Tr}_B[Y_{RB}] + \delta_3 I_R$ together form a strictly feasible point for all $\delta_i > 0$ with $i \in \{1,2,3\}$ and a feasible point for the primal (i.e., supremum formulation)

is $\rho_R = \pi_R$, which is the maximally mixed state (in the Hilbert space $\mathcal{H}_R$) and $\Omega_{RB} = \pi_R \otimes I_B$. This concludes the proof.

B.9 Data Processing under PPT Measurements

Note that $E_\gamma^{\mathcal{M}}$ for a measurement class $\mathcal{M}$ may not satisfy data processing under every quantum channel in general. Next, we show that E_γ^{PPT} satisfies data-processing under special classes that are PPT preserving. PPT preserving channels are defined as follows: A bipartite channel $\mathcal{N}_{AB\to A'B'}$ is called as PPT preserving channel if $\mathrm{T}_{A'} \circ \mathcal{N}_{AB\to A'B'} \circ \mathrm{T}_A$ is a completely positive map.

Note that we highlight what bipartition is considered when PPT-measured hockey-stick divergences are written, and we omit those systems when it is clear from the context.

Proposition 42 (Data processing of E_γ^{PPT})**.** *Let ρ_{AR} and σ_{AR} be quantum states, and let $\mathcal{P}_{AR\to BR'}$ be a PPT preserving channel with A, B systems belonging to one party and R, R' systems belonging to the other party. Then*

$$E_\gamma^{\mathrm{PPT}(B:R')}(\mathcal{P}_{AR\to BR'}(\rho_{AR})\|\mathcal{P}_{AR\to BR'}(\sigma_{AR})) \leq E_\gamma^{\mathrm{PPT}(A:R)}(\rho_{AR}\|\sigma_{AR})\,. \tag{B.175}$$

Furthermore, for a local isometric channel $\mathcal{W}_{R\to R'A'}$ acting on the input system R, and a channel $\mathcal{N}_{A\to B}$ on input system A, we also have that

$$\begin{aligned} &E_\gamma^{\mathrm{PPT}(B:R)}(\mathcal{N}_{A\to B}(\rho_{AR})\|\mathcal{N}_{A\to B}(\sigma_{AR})) \\ &= E_\gamma^{\mathrm{PPT}(B:R'A')}(\mathcal{W}_{R\to R'A'}\otimes\mathcal{N}_{A\to B}(\rho_{AR})\|\mathcal{W}_{R\to R'A'}\otimes\mathcal{N}_{A\to B}(\sigma_{AR}))\,. \end{aligned} \tag{B.176}$$

Proof. Let $M_{BR'}$ be such that $0 \leq M_{BR'} \leq I_{BR'}, 0 \leq \mathrm{T}_B(M_{BR'}) \leq I_{BR'}$ and consider that

$$\mathrm{Tr}[M_{BR'}\left(\mathcal{P}_{AR\to BR'}(\rho_{AR}) - \gamma\mathcal{P}_{AR\to BR'}(\sigma_{AR})\right)] = \mathrm{Tr}\left[\mathcal{P}^{\dagger}_{BR'\to AR}(M_{BR'})(\rho_{AR} - \gamma\sigma_{AR})\right]. \tag{B.177}$$

Since $0 \leq M_{BR'}, \mathrm{T}_B(M_{BR'}) \leq I_{BR'}$ and $\mathcal{P}$ is a CPTP map ($\mathcal{P}^{\dagger}$ is also CPTP and unital), we also have that

$$0 \leq \mathcal{P}^{\dagger}_{BR'\to AR}(M_{BR'}) \leq \mathcal{P}^{\dagger}_{BR'\to AR}(I_{BR'}) = I_{AR}. \tag{B.178}$$

Also consider

$$\mathrm{T}_A\left(\mathcal{P}^{\dagger}_{BR'\to AR}(M_{BR'})\right) = \mathrm{T}_A\left(\mathcal{P}^{\dagger}_{BR'\to AR}\mathrm{T}_B\left(\mathrm{T}_B(M_{BR'})\right)\right) \tag{B.179}$$

$$= \mathrm{T}_A \circ \mathcal{P}^{\dagger}_{BR'\to AR} \circ \mathrm{T}_B\left(\mathrm{T}_B(M_{BR'})\right), \tag{B.180}$$

where the first equality holds since the inverse of T_B is T_B itself.

Let $X_{BR'}$ and Y_{AR} be arbitrary positive semi-definite operators. Since $\mathrm{T}_B \circ \mathcal{P}_{AR\to BR'} \circ T_A$ is a positive map,

$$0 \leq \mathrm{Tr}[X_{BR'}\mathrm{T}_B\left(\mathcal{P}_{AR\to BR'}\left(\mathrm{T}_A(Y_{AR})\right)\right)] \tag{B.181}$$

$$= \mathrm{Tr}[\mathrm{T}_B(X_{BR'})\mathcal{P}_{AR\to BR'}\left(\mathrm{T}_A(Y_{AR})\right)] \tag{B.182}$$

$$= \mathrm{Tr}\left[\mathrm{T}_A \circ \mathcal{P}^{\dagger}_{BR'\to AR} \circ \mathrm{T}_B(X_{BR'})Y_{AR}\right], \tag{B.183}$$

which leads to $\mathrm{T}_A \circ \mathcal{P}^{\dagger}_{BR'\to AR} \circ \mathrm{T}_B$ being a positive map. With that, we arrive at

$$0 \leq \mathrm{T}_A \circ \mathcal{P}^{\dagger}_{BR'\to AR} \circ \mathrm{T}_B\left(\mathrm{T}_B(M_{BR'})\right) \leq \mathrm{T}_A \circ \mathcal{P}^{\dagger}_{BR'\to AR} \circ \mathrm{T}_B(I_{BR'}) = I_{AR} \tag{B.184}$$

with the use of $0 \leq \mathrm{T}_B(M_{BR'}) \leq I_{BR'}$ and (B.178). Then, this shows that

$$\mathcal{P}^{\dagger}_{BR'\to AR}(M_{BR'}) \in \{M_{AR} : 0 \leq M_{AR} \leq I_{AR}, 0 \leq \mathrm{T}_A(M_{AR}) \leq I_{AR}\}. \tag{B.185}$$

With that, we have

$$\mathrm{Tr}[M_{BR'}\,(\mathcal{P}_{AR\to BR'}(\rho_{AR})-\gamma\mathcal{P}_{AR\to BR'}(\sigma_{AR}))] = \mathrm{Tr}\left[\mathcal{P}^{\dagger}_{BR'\to AR}(M_{BR'})(\rho_{AR}-\gamma\sigma_{AR})\right] \tag{B.186}$$

$$\leq \sup_{0\leq M_{AR},\mathrm{T}_A(M_{AR})\leq I_{AR}} \mathrm{Tr}[M_{AR}(\rho_{AR}-\gamma\sigma_{AR})] \tag{B.187}$$

$$= E_{\gamma}^{\mathrm{PPT}(A:R)}(\rho_{AR}\|\sigma_{AR}). \tag{B.188}$$

Finally, by supremizing over $M_{BR'}$ such that $0 \leq M_{BR'}, \mathrm{T}_B(M_{BR'}) \leq I_{BR}$, we conclude the proof of the inequality presented in (B.175).

The equality in (B.176) follows by applying the data-processing inequality under isometric channel $\mathcal{W}_{R\to R'A'}\otimes\mathcal{I}_B$ and its inverse channel $\mathcal{W}^{-1}_{R\to R'A'}\otimes\mathcal{I}_B$, both of which belong to the set of PPT preserving channels with respect to the partial transpose on system B. □

B.10 Proof of Proposition 21

Let G be a finite group with $\{U_A(g)\}_{g\in G}$ and $\{V_B(g)\}_{g\in G}$ unitary representations. Since the set $\{\psi_{RA} : \psi_A = \mathcal{T}_G(\psi_A)\}$ is a subset of all pure states, we immediately obtain the following inequality:

$$E_{\gamma}^{\mathrm{PPT}}(\mathcal{P}\|\mathcal{Q}) \geq \sup_{\psi_{RA}} \left\{E_{\gamma}^{\mathrm{PPT}}(\mathcal{P}_{A\to B}(\psi_{RA})\|\mathcal{Q}_{A\to B}(\psi_{RA})) : \psi_A = \mathcal{T}_G(\psi_A)\right\}. \tag{B.189}$$

What remains to show is the reverse inequality.

Let $\rho_A \in \mathcal{L}(\mathcal{H}_A)$ and ψ_{RA} be a purification of state ρ_A. Let $\bar{\rho}_A$ be the group average of ρ_A, i.e.;

$$\bar{\rho}_A := \frac{1}{|G|}\sum_{g\in G} U_A(g)\rho_A U_A^{\dagger}(g), \tag{B.190}$$

and let $\phi^{\bar{\rho}}_{RA}$ be a purification of $\bar{\rho}_A$.

Let us also consider the following state $\psi^{\bar{\rho}}_{PR'A}$, which is also a purification of the state $\bar{\rho}_A$ with the purifying systems P and R' with P system being a classical system and R' being isomorphic to system R:

$$|\psi^{\bar{\rho}}\rangle_{PR'A} \coloneqq \frac{1}{\sqrt{|G|}} \sum_{g\in G} |g\rangle_P \otimes (I_{R'} \otimes U_A(g))|\psi\rangle_{RA}, \tag{B.191}$$

where $\{|g\rangle\}_{g\in G}$ is an orthonormal basis for $\mathcal{H}_P$ indexed by the elements of G.

Since all purifications of a state can be mapped to each other by isometries acting on the purifying systems, there exists an isometry $W_{R\to PR'}$ that forms the isometric channel $\mathcal{W}_{R\to PR'}$ such that

$$\psi^{\bar{\rho}}_{PRA} = \mathcal{W}_{R\to PR'}(\phi^{\bar{\rho}}_{RA}). \tag{B.192}$$

Consider that,

$$\begin{aligned}
&E_\gamma^{\mathrm{PPT}(R:B)}\left(\mathcal{P}_{A\to B}(\phi^{\bar{\rho}}_{RA})\middle\|\mathcal{Q}_{A\to B}(\phi^{\bar{\rho}}_{RA})\right)\\
&= E_\gamma^{\mathrm{PPT}(PR':B)}\left(\mathcal{W}_{R\to PR'}\circ\mathcal{P}_{A\to B}\circ(\phi^{\bar{\rho}}_{RA})\middle\|\mathcal{W}_{R\to PR'}\circ\mathcal{Q}_{A\to B}(\phi^{\bar{\rho}}_{RA})\right) &\text{(B.193)}\\
&= E_\gamma^{\mathrm{PPT}(PR':B)}\left(\mathcal{P}_{A\to B}(\psi^{\bar{\rho}}_{PR'A})\middle\|\mathcal{Q}_{A\to B}(\psi^{\bar{\rho}}_{PR'A})\right) &\text{(B.194)}\\
&\geq E_\gamma^{\mathrm{PPT}(PR':B)}\left(\frac{1}{|G|}\sum_{g\in G}|g\rangle\langle g|_P\otimes(\mathcal{P}_{A\to B}\circ\mathcal{U}_A(g))(\psi_{RA})\middle\|\frac{1}{|G|}\sum_{g\in G}|g\rangle\langle g|_P\otimes(\mathcal{Q}_{A\to B}\circ\mathcal{U}_A(g))(\psi_{RA})\right) &\text{(B.195)}\\
&\geq E_\gamma^{\mathrm{PPT}(PR':B)}\left(\frac{1}{|G|}\sum_{g\in G}|g\rangle\langle g|_P\otimes\left((\mathcal{V}_B(g))^\dagger\circ\mathcal{P}_{A\to B}\circ\mathcal{U}_A(g)\right)(\psi_{RA})\right.\\
&\qquad\left.\middle\|\frac{1}{|G|}\sum_{g\in G}|g\rangle\langle g|_P\otimes\left((\mathcal{V}_B(g))^\dagger\circ\mathcal{Q}_{A\to B}\circ\mathcal{U}_A(g)\right)(\psi_{RA})\right) &\text{(B.196)}\\
&= E_\gamma^{\mathrm{PPT}(PR':B)}\left(\frac{1}{|G|}\sum_{g\in G}|g\rangle\langle g|_P\otimes\mathcal{P}_{A\to B}(\psi_{RA})\middle\|\frac{1}{|G|}\sum_{g\in G}|g\rangle\langle g|_P\otimes\mathcal{Q}_{A\to B}(\psi_{RA})\right) &\text{(B.197)}\\
&\geq E_\gamma^{\mathrm{PPT}(R':B)}(\mathcal{P}_{A\to B}(\psi_{RA})\|\mathcal{Q}_{A\to B}(\psi_{RA}))\,, &\text{(B.198)}
\end{aligned}$$

where the first two equalities follow from (B.176), the first inequality by applying data processing of E_γ^{PPT} in Proposition 42 for the dephasing channel $X \to \sum_{g\in G} |g\rangle\langle g|X|g\rangle\langle g|$ on system P (a PPT preserving channel); the second inequality by applying the data processing under the unitary channel given by the unitary $\sum_{g\in G} |g\rangle\langle g|_P \otimes V_B^\dagger(g)$ [1]; third equality by the joint-covariance of the channels $\mathcal{P}$ and $\mathcal{Q}$ such that

$$(\mathcal{V}_B(g))^\dagger \circ \mathcal{P}_{A\to B} \circ \mathcal{U}_A(g) = \mathcal{P}_{A\to B}, \quad (\mathcal{V}_B(g))^\dagger \circ \mathcal{Q}_{A\to B} \circ \mathcal{U}_A(g) = \mathcal{Q}_{A\to B}; \tag{B.199}$$

and finally the last inequality by the data processing under the partial trace channel on the system P by using Proposition 42.

By definition, the pure state $\phi_{RA}^{\bar{D}}$ is such that its reduced state on A is invariant under the channel $\mathcal{T}_G$. Then, by optimizing over all such pure states, and noting that R' is isomorphic to system R, we obtain that

$$E_\gamma^{\mathrm{PPT}(R:B)}(\mathcal{P}_{A\to B}(\psi_{RA})\|\mathcal{Q}_{A\to B}(\psi_{RA})) \leq \sup_{\phi_{RA}} \left\{ E_\gamma^{\mathrm{PPT}}(\mathcal{P}_{A\to B}(\phi_{RA})\|\mathcal{Q}_{A\to B}(\phi_{RA})) : \phi_A = \mathcal{T}_G(\phi_A) \right\}. \tag{B.200}$$

Since the above inequality holds for all pure states ψ_{RA}, we obtain the reverse inequality,

$$E_\gamma^{\mathrm{PPT}}(\mathcal{P}\|\mathcal{Q}) \leq \sup_{\psi_{RA}} \left\{ E_\gamma^{\mathrm{PPT}}(\mathcal{P}_{A\to B}(\psi_{RA})\|\mathcal{Q}_{A\to B}(\psi_{RA})) : \psi_A = \mathcal{T}_G(\psi_A) \right\}. \tag{B.201}$$

Then, we conclude the first equality by combining (B.189) and (B.201).

To prove the next claim, if the representation $\{U_A(g)\}_{g\in G}$ acting on the input space of the channel $\mathcal{P}$ and $\mathcal{Q}$ is irreducible, then for every state ψ_{RA} such that

[1] This is because that one could implement this operation on these states as a classically controlled LOCC operation where the von-Neumann measurement $|g\rangle\langle g|_{g\in G}$ is applied on the register P followed by a rotation given by the unitary channel $(\mathcal{V}_B(g))^\dagger$. Recall that we chose the system P to be classical, so the mentioned procedure can be followed.

$\rho_A = \psi_A$, it holds that $\bar{\rho}_A = I_A/D_A$. Then, since the maximally entangled state is a purification of the maximally mixed state, we let $\phi_{RA}^{\bar{\rho}} = \Phi_{RA}$, which implies that

$$E_\gamma^{\mathrm{PPT}(R:B)}\left(\mathcal{P}_{A\to B}(\phi_{RA}^{\bar{\rho}})\middle\|\mathcal{Q}_{A\to B}(\phi_{RA}^{\bar{\rho}})\right) = E_\gamma^{\mathrm{PPT}(R:B)}(\mathcal{P}_{A\to B}(\Phi_{RA})\|\mathcal{Q}_{A\to B}(\Phi_{RA}))\,. \tag{B.202}$$

Using (B.198), we obtain that

$$E_\gamma^{\mathrm{PPT}(R:B)}(\mathcal{P}_{A\to B}(\Phi_{RA})\|\mathcal{Q}_{A\to B}(\Phi_{RA})) \geq E_\gamma^{\mathrm{PPT}(R:B)}(\mathcal{P}_{A\to B}(\psi_{RA})\|\mathcal{Q}_{A\to B}(\psi_{RA})) \tag{B.203}$$

for all states ψ_{RA}. So, by optimizing over all states ψ_{RA} we have that

$$E_\gamma^{\mathrm{PPT}}(\mathcal{P}\|\mathcal{Q}) \leq E_\gamma^{\mathrm{PPT}(R:B)}(\mathcal{P}_{A\to B}(\Phi_{RA})\|\mathcal{Q}_{A\to B}(\Phi_{RA}))\,. \tag{B.204}$$

By choosing the pure state to be the maximally entangled state in (4.25), we obtain the reverse inequality, concluding the proof of Proposition 21.

B.11 Proof of Proposition 22

Proof follows similar to the proof of Proposition 41. With the use of Proposition 21, we have that

$$E_\gamma^{\mathrm{PPT}}\left(\mathcal{A}_{\mathrm{Dep}}^q\|\mathcal{A}_{\mathrm{Dep}}^p\right) = E_\gamma^{\mathrm{PPT}}\left(\frac{\Gamma_{RB}^{\mathcal{A}_{\mathrm{Dep}}^q}}{d}\middle\|\frac{\Gamma_{RB}^{\mathcal{A}_{\mathrm{Dep}}^p}}{d}\right) \tag{B.205}$$

$$= E_\gamma^{\mathrm{PPT}}\left(\zeta^{\eta'}\middle\|\zeta^{\eta}\right), \tag{B.206}$$

where the last equality follows by the substitution in (B.165) together with $\eta := 1 - p + p/d^2$ and $\eta' := 1 - q + q/d^2$.

Finally, we conclude the proof by adapting Proposition 38 to (B.206).

APPENDIX C

CHAPTER 5 AND 6 OF APPENDIX

C.1 Properties of Hockey-Stick Divergence

Proposition 43. *Given a channel $\mathcal{A}$ and $\gamma \geq 1$, the following equality holds:*

$$\sup_{\rho,\sigma} E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = \sup_{\varphi_1 \perp \varphi_2} E_\gamma(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)), \tag{C.1}$$

where the optimization on the left-hand side is over all states ρ and σ and the optimization on the right-hand side is over orthogonal pure states φ_1 and φ_2.

Proof. First, consider that the inequality

$$\sup_{\rho,\sigma} E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \geq \sup_{\varphi_1 \perp \varphi_2} E_\gamma(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)) \tag{C.2}$$

trivially holds due to the containment $\{(\varphi_1, \varphi_2) : \varphi_1 \perp \varphi_2\} \subset \{(\rho, \sigma)\}$. So it remains to prove the opposite inequality:

$$\sup_{\rho,\sigma} E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \sup_{\varphi_1 \perp \varphi_2} E_\gamma(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)). \tag{C.3}$$

First, from the joint convexity of E_γ (see [62, Eq. (II.17)]), it readily follows that

$$\sup_{\rho,\sigma} E_\gamma(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \sup_{\psi,\phi} E_\gamma(\mathcal{A}(\psi)\|\mathcal{A}(\phi)), \tag{C.4}$$

so that it suffices to perform the optimization over pure states. As such, we now prove that

$$\sup_{\psi,\phi} E_\gamma(\mathcal{A}(\psi)\|\mathcal{A}(\phi)) \leq \sup_{\varphi_1 \perp \varphi_2} E_\gamma(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)). \tag{C.5}$$

From Lemma 12, the operator inequality in (C.11) holds. Applying that then gives

$$E_\gamma(\mathcal{A}(\psi)\|\mathcal{A}(\phi)) = \sup_{M\geq 0}\{\mathrm{Tr}[M(\mathcal{A}(\psi)-\gamma\mathcal{A}(\phi))] : M \leq I\} \tag{C.6}$$

$$\leq \sup_{M\geq 0}\{\mathrm{Tr}[M\lambda_1(\mathcal{A}(\varphi_1)-\gamma\mathcal{A}(\varphi_2)))] : M \leq I\} \tag{C.7}$$

$$\leq \sup_{M\geq 0}\{\mathrm{Tr}[M(\mathcal{A}(\varphi_1)-\gamma\mathcal{A}(\varphi_2)))] : M \leq I\} \tag{C.8}$$

$$= E_\gamma(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)) \tag{C.9}$$

$$\leq \sup_{\varphi_1\perp\varphi_2} E_\gamma(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)). \tag{C.10}$$

The first inequality follows from (C.11). Since this inequality holds for all pure states ψ and ϕ, we conclude (C.5). □

Lemma 12. *For pure states ψ and ϕ, a positive map $\mathcal{A}$, and $\gamma \geq 1$, the following operator inequality holds:*

$$\mathcal{A}(\psi)-\gamma\mathcal{A}(\phi) \leq \lambda_1\left[\mathcal{A}(\varphi_1)-\gamma\mathcal{A}(\varphi_2)\right], \tag{C.11}$$

where a spectral decomposition of $\psi-\gamma\phi$ is given by

$$\psi-\gamma\phi = \lambda_1\varphi_1 - \lambda_2\varphi_2, \tag{C.12}$$

with φ_1 and φ_2 orthogonal pure states and

$$\lambda_1 := \frac{1}{2}\left[\sqrt{(\gamma+1)^2-4\gamma F(\psi,\phi)}-(\gamma-1)\right] \in [0,1], \tag{C.13}$$

$$\lambda_2 := \frac{1}{2}\left[\sqrt{(\gamma+1)^2-4\gamma F(\psi,\phi)}+(\gamma-1)\right] \in [\gamma-1,\gamma], \tag{C.14}$$

and the fidelity $F(\psi,\phi) := |\langle\psi|\phi\rangle|^2$.

Proof. Consider $\gamma \geq 1$ and the operator $\psi-\gamma\phi$. Now consider that $|\psi\rangle$ and $|\phi\rangle$ span a two-dimensional subspace, and in this subspace we can write

$$|\psi\rangle = \cos(\theta)|\phi\rangle + \sin(\theta)|\phi^\perp\rangle, \tag{C.15}$$

where $|\phi^\perp\rangle$ is a pure state vector orthogonal to $|\phi\rangle$. Observe that $F(\psi,\phi) = \cos^2(\theta)$. Then

$$|\psi\rangle\langle\psi| = \cos^2(\theta)|\phi\rangle\langle\phi| + \sin(\theta)\cos(\theta)\left(|\phi\rangle\langle\phi^\perp| + |\phi^\perp\rangle\langle\phi|\right) + \sin^2(\theta)|\phi^\perp\rangle\langle\phi^\perp| \tag{C.16}$$

and thus, in the basis $\{|\phi\rangle, |\phi^\perp\rangle\}$,

$$\psi - \gamma\phi = |\psi\rangle\langle\psi| - \gamma|\phi\rangle\langle\phi| \tag{C.17}$$

$$= \begin{bmatrix} \cos^2(\theta) - \gamma & \sin(\theta)\cos(\theta) \\ \sin(\theta)\cos(\theta) & \sin^2(\theta) \end{bmatrix}. \tag{C.18}$$

For a 2×2 matrix A, the following well known expression exists for its eigenvalues:

$$\lambda_\pm = \frac{1}{2}\left(\mathrm{Tr}[A] \pm \sqrt{(\mathrm{Tr}[A])^2 - 4\det(A)}\right). \tag{C.19}$$

In this case, we find that $\mathrm{Tr}[\psi - \gamma\phi] = 1 - \gamma$ and $\det(\psi - \gamma\phi) = -\gamma\sin^2(\theta)$, so that the eigenvalues of $\psi - \gamma\phi$ are given by

$$\lambda_\pm = \frac{1}{2}\left(1 - \gamma \pm \sqrt{(1-\gamma)^2 + 4\gamma\sin^2(\theta)}\right) \tag{C.20}$$

$$= \frac{1}{2}\left(1 - \gamma \pm \sqrt{(\gamma-1)^2 + 4\gamma\left(1 - F(\psi,\phi)\right)}\right) \tag{C.21}$$

$$= \frac{1}{2}\left(1 - \gamma \pm \sqrt{(\gamma+1)^2 - 4\gamma F(\psi,\phi)}\right). \tag{C.22}$$

Thus, there exist orthogonal state vectors $|\varphi_1\rangle$ and $|\varphi_2\rangle$ such that

$$\psi - \gamma\phi = \lambda_+\varphi_1 + \lambda_-\varphi_2. \tag{C.23}$$

Observe that $\lambda_+ \in [0,1]$ and $\lambda_- \in [-\gamma, -(\gamma-1)]$ because $F(\psi,\phi) \in [0,1]$. Now identifying $\lambda_1 = \lambda_+$ and $\lambda_2 = -\lambda_-$, we can write

$$\psi - \gamma\phi = \lambda_1\varphi_1 - \lambda_2\varphi_2, \tag{C.24}$$

and conclude that $\lambda_1 \in [0,1]$ and $\lambda_2 \in [\gamma - 1, \gamma]$, as claimed.

To prove the operator inequality in (C.11), consider that

$$\mathcal{A}(\psi) - \gamma\mathcal{A}(\phi) = \mathcal{A}(\psi - \gamma\phi) \tag{C.25}$$

$$= \mathcal{A}(\lambda_1\varphi_1 - \lambda_2\varphi_2) \tag{C.26}$$

$$= \lambda_1\mathcal{A}(\varphi_1) - \lambda_2\mathcal{A}(\varphi_2) \tag{C.27}$$

$$= \lambda_1\mathcal{A}(\varphi_1) - \gamma\lambda_1\mathcal{A}(\varphi_2) + \gamma\lambda_1\mathcal{A}(\varphi_2) - \lambda_2\mathcal{A}(\varphi_2) \tag{C.28}$$

$$= \lambda_1\left[\mathcal{A}(\varphi_1) - \gamma\mathcal{A}(\varphi_2)\right] + \left[\gamma\lambda_1 - \lambda_2\right]\mathcal{A}(\varphi_2) \tag{C.29}$$

$$\leq \lambda_1\left[\mathcal{A}(\varphi_1) - \gamma\mathcal{A}(\varphi_2)\right]. \tag{C.30}$$

The last inequality follows because $\gamma\lambda_1 - \lambda_2 \leq 0$ and $\mathcal{A}$ is a positive map, so that $[\gamma\lambda_1 - \lambda_2]\,\mathcal{A}(\varphi_2) \leq 0$. Indeed, consider that

$$1 - \gamma = \mathrm{Tr}[\psi - \gamma\phi] = \mathrm{Tr}[\lambda_1\varphi_1 - \lambda_2\varphi_2] = \lambda_1 - \lambda_2, \tag{C.31}$$

which implies that

$$\gamma\lambda_1 - \lambda_2 = \gamma\lambda_1 - \lambda_1 + \lambda_1 - \lambda_2 \tag{C.32}$$

$$= (\gamma - 1)\,\lambda_1 + (1 - \gamma) \tag{C.33}$$

$$= -(\gamma - 1)\,(1 - \lambda_1) \tag{C.34}$$

$$\leq 0. \tag{C.35}$$

Here, we used the fact that $\gamma \geq 1$ and $\lambda_1 \in [0, 1]$. □

Remark 44 (Hockey-Stick Divergence for Pure States)**.** *For two pure states ψ and ϕ and $\gamma \geq 1$, their hockey-stick divergence is equal to*

$$E_\gamma(\psi\|\phi) = \frac{1}{2}\left(\sqrt{(\gamma+1)^2 - 4\gamma F(\psi,\phi)} + 1 - \gamma\right). \tag{C.36}$$

This follows by considering the positive eigenvalue in (C.12) together with (C.13). Note that with the above equality, it follows that the upper bound in Eq. (II.11) in [62] is saturated for pure states.

C.2 Datta–Leditzky Divergence

Proposition 44. *Given a channel $\mathcal{A}$, $\gamma \geq 1$, and $\delta \in [0,1)$, the following equality holds:*

$$\sup_{\rho,\sigma} \overline{D}^{\delta}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) = \sup_{\varphi_1 \perp \varphi_2} \overline{D}^{\delta}(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)), \tag{C.37}$$

where the optimization on the left-hand side is over all states ρ and σ and the optimization on the right-hand side is over orthogonal pure states φ_1 and φ_2.

Proof. First, consider that the inequality

$$\sup_{\rho,\sigma} \overline{D}^{\delta}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \geq \sup_{\varphi_1 \perp \varphi_2} \overline{D}^{\delta}(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)) \tag{C.38}$$

trivially holds due to the containment $\{(\varphi_1,\varphi_2) : \varphi_1 \perp \varphi_2\} \subset \{(\rho,\sigma)\}$. So it remains to prove the opposite inequality:

$$\sup_{\rho,\sigma} \overline{D}^{\delta}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \sup_{\varphi_1 \perp \varphi_2} E_{\gamma}(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)). \tag{C.39}$$

First, from the joint quasi-convexity of $\overline{D}^{\delta}$ (see [95, Eq. (67)]), it readily follows that

$$\sup_{\rho,\sigma} \overline{D}^{\delta}(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \sup_{\psi,\phi} \overline{D}^{\delta}(\mathcal{A}(\psi)\|\mathcal{A}(\phi)), \tag{C.40}$$

so that it suffices to perform the optimization over pure states. As such, we now prove that

$$\sup_{\psi,\phi} \overline{D}^{\delta}(\mathcal{A}(\psi)\|\mathcal{A}(\phi)) \leq \sup_{\varphi_1 \perp \varphi_2} \overline{D}^{\delta}(\mathcal{A}(\varphi_1)\|\mathcal{A}(\varphi_2)). \tag{C.41}$$

By [95, Corollary 1], we have that

$$\sup_{\psi,\phi} \overline{D}^{\delta}(\mathcal{A}(\psi)\|\mathcal{A}(\phi)) = \sup_{\psi,\phi} \ln \sup_{0 \leq W \leq I, \operatorname{Tr}[W\mathcal{A}(\psi)] \geq \delta} \frac{\operatorname{Tr}[W\mathcal{A}(\psi)] - \delta}{\operatorname{Tr}[W\mathcal{A}(\phi)]} \tag{C.42}$$

$$= \ln \sup_{\psi,\phi} \sup_{0 \leq W \leq I, \operatorname{Tr}\left[\mathcal{A}^{\dagger}(W)\psi\right] \geq \delta} \frac{\operatorname{Tr}\left[\mathcal{A}^{\dagger}(W)\psi\right] - \delta}{\operatorname{Tr}[\mathcal{A}^{\dagger}(W)\phi]}. \tag{C.43}$$

Then consider that

$$\sup_{\psi,\phi} \overline{D}^{\delta}(\mathcal{A}(\psi)\|\mathcal{A}(\phi)) = \ln \sup_{\psi,\phi} \sup_{0\leq W\leq I} \left\{ \frac{\mathrm{Tr}\left[\mathcal{A}^{\dagger}(W)\psi\right] - \delta}{\mathrm{Tr}[\mathcal{A}^{\dagger}(W)\phi]} : \mathrm{Tr}\left[\mathcal{A}^{\dagger}(W)\psi\right] \geq \delta \right\} \tag{C.44}$$

$$\leq \ln \sup_{\psi,\phi} \sup_{0\leq W\leq I} \left\{ \frac{\mathrm{Tr}\left[\mathcal{A}^{\dagger}(W)\psi\right] - \delta}{\mathrm{Tr}[\mathcal{A}^{\dagger}(W)\phi]} : \lambda_{\max}\left(\mathcal{A}^{\dagger}(W)\right) \geq \delta \right\} \tag{C.45}$$

$$\leq \ln \sup_{0\leq W\leq I} \left\{ \frac{\lambda_{\max}\left(\mathcal{A}^{\dagger}(W)\right) - \delta}{\lambda_{\min}(\mathcal{A}^{\dagger}(W))} : \lambda_{\max}\left(\mathcal{A}^{\dagger}(W)\right) \geq \delta \right\} \tag{C.46}$$

where the first inequality follows by the implication $\mathrm{Tr}\left[\mathcal{A}^{\dagger}(W)\psi\right] \geq \delta \implies \lambda_{\max}\left(\mathcal{A}^{\dagger}(W)\right) \geq \delta$ with $\lambda_{\max}(B)$ denoting the maximum eigenvalue of B; the last inequality follows by $\lambda_{\min}(B) \leq \mathrm{Tr}[B\psi] \leq \lambda_{\max}(B)$ for a state ψ with $\lambda_{\min}(B)$ denoting the minimum eigenvalue of B.

Next, observe that the upper bound in (C.46) is achieved by picking ψ and ϕ to be the pure states formed by the eigenvectors corresponding to the maximum and minimum eigenvalues of $\mathcal{A}^{\dagger}(W)$ for all W such that $0 \leq W \leq I$. Also note that these eigenvectors are orthogonal. With this, we conclude that (C.41) holds, and then together with (C.38), we conclude the proof of the proposition. □

C.3 Proof of Proposition 28

Converse:

For all $\mathcal{A} \in \mathcal{B}^{\varepsilon}$ and ρ and σ states, by (5.98), we have

$$\sup_{\substack{\mathcal{A}\in\mathcal{B}^{\varepsilon},\\ \rho,\sigma\in\mathcal{D}}} D_f(\mathcal{A}(\rho)\|\mathcal{A}(\sigma)) \leq \sup_{\substack{\mathcal{A}\in\mathcal{B}^{\varepsilon},\\ \rho,\sigma\in\mathcal{D}}} \frac{f(e^{\varepsilon}) + e^{\varepsilon} f(e^{-\varepsilon})}{e^{\varepsilon}+1} T(\rho,\sigma) \tag{C.47}$$

$$\leq \frac{f(e^{\varepsilon}) + e^{\varepsilon} f(e^{-\varepsilon})}{e^{\varepsilon}+1}, \tag{C.48}$$

where the last inequality follows because $T(\rho,\sigma) \leq 1$.

Achievability: First we define the measurement channel $\mathcal{M}$ as

$$\mathcal{M}(\omega) \coloneqq \mathrm{Tr}[M\omega]|0\rangle\langle 0| + \mathrm{Tr}[(I - M)\omega]|1\rangle\langle 1|, \tag{C.49}$$

where M is a measurement operator (satisfying $0 \leq M \leq I$) and ω is a quantum state. We also define the classical binary symmetric channel $\mathcal{A}^p_{\mathrm{BSC}}$ with the flip parameter $p \in [0, 1/2]$ as follows:

$$\mathcal{A}^p_{\mathrm{BSC}}(|0\rangle\langle 0|) = (1 - p)|0\rangle\langle 0| + p|1\rangle\langle 1|, \tag{C.50}$$

$$\mathcal{A}^p_{\mathrm{BSC}}(|1\rangle\langle 1|) = (1 - p)|1\rangle\langle 1| + p|0\rangle\langle 0|. \tag{C.51}$$

With that we consider the composite channel $\mathcal{A}^p_{\mathrm{BSC}} \circ \mathcal{M}$, and for the input state ω we have

$$\mathcal{A}^p_{\mathrm{BSC}} \circ \mathcal{M}(\omega) = q^\omega_0|0\rangle\langle 0| + q^\omega_1|1\rangle\langle 1|, \tag{C.52}$$

where

$$q^\omega_0 \coloneqq (1 - p)\mathrm{Tr}[M\omega] + p(1 - \mathrm{Tr}[M\omega]) \tag{C.53}$$

$$= p + (1 - 2p)\mathrm{Tr}[M\omega], \tag{C.54}$$

$$q^\omega_1 \coloneqq p\mathrm{Tr}[M\omega] + (1 - p)(1 - \mathrm{Tr}[M\omega]) \tag{C.55}$$

$$= p + (1 - 2p)(1 - \mathrm{Tr}[M\omega]). \tag{C.56}$$

For states ρ and σ, we consider

$$\frac{q^\rho_0}{q^\sigma_0} - 1 = \frac{q^\rho_0 - q^\sigma_0}{q^\sigma_0} \tag{C.57}$$

$$= \frac{p + (1 - 2p)\mathrm{Tr}[M\rho] - (p + (1 - 2p)\mathrm{Tr}[M\sigma])}{p + (1 - 2p)\mathrm{Tr}[M\sigma]} \tag{C.58}$$

$$\leq \frac{(1 - 2p)\mathrm{Tr}[M(\rho - \sigma)]}{p} \tag{C.59}$$

$$\leq \frac{(1 - 2p)T(\rho, \sigma)}{p} \tag{C.60}$$

$$\leq \frac{1 - 2p}{p}, \tag{C.61}$$

where the first equality follows since $(1-2p)\mathrm{Tr}[M\sigma] \geq 0$ when $p \in [0, 1/2]$; second inequality follows due to $T(\rho,\sigma) = \sup_{0\leq M\leq I} \mathrm{Tr}[M(\rho-\sigma)]$; and the last inequality by $T(\rho,\sigma) \leq 1$.

With that, we arrive at

$$\frac{q_0^\rho}{q_0^\sigma} \leq \frac{1}{p} - 1. \tag{C.62}$$

Now, we choose $p = 1/(e^\varepsilon + 1)$; then we see that

$$\frac{q_0^\rho}{q_0^\sigma} \leq \frac{1}{p} - 1 = e^\varepsilon. \tag{C.63}$$

Similarly, we can show that

$$\frac{q_1^\rho}{q_1^\sigma} \leq \frac{1}{p} - 1 = e^\varepsilon \tag{C.64}$$

by using the fact that

$$q_1^\omega = p + (1-2p)\mathrm{Tr}[(I-M)\omega] \tag{C.65}$$

and following analogous proof arguments as in (C.57).

With that, we show the following: for all ρ and σ with $p = 1/(e^\varepsilon + 1)$,

$$\begin{aligned}
&E_{e^\varepsilon}\left(\mathcal{A}^p_{\mathrm{BSC}} \circ \mathcal{M}(\rho) \| \mathcal{A}^p_{\mathrm{BSC}} \circ \mathcal{M}(\sigma)\right) \\
&= \mathrm{Tr}\left[(\mathcal{A}^p_{\mathrm{BSC}} \circ \mathcal{M}(\rho) - e^\varepsilon \mathcal{A}^p_{\mathrm{BSC}} \circ \mathcal{M}(\sigma))_+\right] && \text{(C.66)} \\
&= \mathrm{Tr}\left[\left((q_0^\rho - e^\varepsilon q_0^\sigma)|0\rangle\langle 0| + (q_1^\rho - e^\varepsilon q_1^\sigma)|1\rangle\langle 1|\right)_+\right] && \text{(C.67)} \\
&= \max\{0, q_0^\rho - e^\varepsilon q_0^\sigma\}\mathrm{Tr}[|0\rangle\langle 0|] + \max\{0, q_1^\rho - e^\varepsilon q_1^\sigma\}\mathrm{Tr}[|1\rangle\langle 1|] && \text{(C.68)} \\
&= \max\{0, q_0^\rho - e^\varepsilon q_0^\sigma\} + \max\{0, q_1^\rho - e^\varepsilon q_1^\sigma\} && \text{(C.69)} \\
&= 0, && \text{(C.70)}
\end{aligned}$$

where the last equality follows by using (C.64) and (C.65).

Then we conclude that

$$\sup_{\rho,\sigma\in\mathcal{D}} E_{e^\varepsilon}\Big(\mathcal{A}^p_{\mathrm{BSC}}\circ\mathcal{M}(\rho)\|\mathcal{A}^p_{\mathrm{BSC}}\circ\mathcal{M}(\sigma)\Big)=0, \tag{C.71}$$

so that $\mathcal{A}^p_{\mathrm{BSC}}\circ\mathcal{M}\in\mathcal{B}^\varepsilon$ satisfies ε-QLDP with $p=1/(e^\varepsilon+1)$ by (5.8).

Let us consider the special case in which ρ' and σ' are orthogonal states. Then choose $M=\Pi_{\rho'}$, which is the projection onto the support of the state ρ'. With that choice

$$\mathrm{Tr}[M\rho']=1 \quad \text{and} \quad \mathrm{Tr}[M\sigma']=0. \tag{C.72}$$

For states ρ' and σ', then we have

$$q_0^{\rho'}=1-p, \tag{C.73}$$

$$q_0^{\sigma'}=p. \tag{C.74}$$

With that, we define two binary classical distributions as follows: $\omega\in\{\rho',\sigma'\}$

$$q^\omega := q_0^\omega|0\rangle\langle 0|+(1-q_0^\omega)|1\rangle\langle 1|. \tag{C.75}$$

Recall that for classical discrete distributions p and q, f-divergences can be written as

$$D_f(p\|q) := \sum_x q(x) f\left(\frac{p(x)}{q(x)}\right) \tag{C.76}$$

$$= \int_1^\infty f''(\gamma)E_\gamma(p\|q)+\gamma^{-3}f''(\gamma^{-1})E_\gamma(q\|p)\,\mathrm{d}\gamma. \tag{C.77}$$

For this setting, we have

$$D_f\Big(\mathcal{A}^p_{\mathrm{BSC}}\circ\mathcal{M}(\rho')\|\mathcal{A}^p_{\mathrm{BSC}}\circ\mathcal{M}(\sigma')\Big)=D_f(q^{\rho'}\|q^{\sigma'}) \tag{C.78}$$

$$= \sum_{x\in\{0,1\}} q_x^{\sigma'} f\left(\frac{q_x^{\rho'}}{q_x^{\sigma'}}\right) \tag{C.79}$$

$$= pf\left(\frac{1-p}{p}\right)+(1-p)f\left(\frac{p}{1-p}\right) \tag{C.80}$$

$$= \frac{1}{e^\varepsilon+1}f(e^\varepsilon)+\frac{e^\varepsilon}{e^\varepsilon+1}f(e^{-\varepsilon}), \tag{C.81}$$

where the last equality follows by substituting $p = 1/(e^{\varepsilon}+1)$.

This then leads to

$$\frac{1}{e^{\varepsilon}+1} f(e^{\varepsilon}) + \frac{e^{\varepsilon}}{e^{\varepsilon}+1} f(e^{-\varepsilon}) = D_f\Big(\mathcal{A}^p_{\mathrm{BSC}} \circ \mathcal{M}(\rho') \| \mathcal{A}^p_{\mathrm{BSC}} \circ \mathcal{M}(\sigma')\Big) \leq \sup_{\substack{A \in \mathcal{B}^{\varepsilon}, \\ \rho,\sigma \in \mathcal{D}}} D_f(\mathcal{A}(\rho) \| \mathcal{A}(\sigma))\,, \tag{C.82}$$

by recalling that $\mathcal{A}^p_{\mathrm{BSC}} \circ \mathcal{M} \in \mathcal{B}^{\varepsilon}$.

Finally, we conclude the proof by combining (C.48) and (C.82).

C.4 Sample Complexity in the Low-Privacy Regime

In this appendix, we prove Proposition 30. We use the following shorthand:

$$d_B^2(\rho,\sigma) \coloneqq [d_B(\rho,\sigma)]^2\,. \tag{C.83}$$

Let ρ and σ be states with dimension d. Let $\mathcal{M}$ be the measurement channel comprised of the POVM formed by the eigenbasis of $\rho\#\sigma^{-1}$ with $k \leq d$ POVM elements, where $A\#B$ denotes the matrix geometric mean of positive semi-definite operators A and B. It is defined for A and B positive definite operators as

$$A\#B \coloneqq A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{1/2} A^{1/2}. \tag{C.84}$$

and as $\lim_{\varepsilon\to 0^+}(A+\varepsilon I)\#(B+\varepsilon I)$ in the more general case when A and B are positive semi-definite. With that measurement (POVM) achieving fidelity [49] and recalling (2.5), we have

$$d_B^2(\rho,\sigma) = d_B^2(\mathcal{M}(\rho),\mathcal{M}(\sigma))\,. \tag{C.85}$$

Also recall that

$$[T(\rho,\sigma)]^2 \leq d_B^2(\rho,\sigma) \leq 2T(\rho,\sigma). \tag{C.86}$$

Fix $p = 2/(e^{\varepsilon} + 1)$ and $d = 2$. Also choose $\mathcal{A} \coloneqq \mathcal{A}^{p}_{\mathrm{Dep}} \circ \mathcal{M} \in \mathcal{B}^{\varepsilon}$. With these in hand, consider that

$$
\begin{aligned}
d_B^2(\mathcal{A}(\rho), \mathcal{A}(\sigma)) &\geq \left[T(\mathcal{A}(\rho), \mathcal{A}(\sigma))\right]^2 && \text{(C.87)}\\
&= (1-p)^2 \left[T(\mathcal{M}(\rho), \mathcal{M}(\sigma))\right]^2 && \text{(C.88)}\\
&\geq \frac{(1-p)^2}{4} \left[d_B^2(\mathcal{M}(\rho), \mathcal{M}(\sigma))\right]^2 && \text{(C.89)}\\
&\geq \frac{1}{4} d_B^2(\mathcal{M}(\rho), \mathcal{M}(\sigma)) \min\left\{1, (1-p)^2 d_B^2(\mathcal{M}(\rho), \mathcal{M}(\sigma))\right\} && \text{(C.90)}\\
&= \frac{1}{4} d_B^2(\rho, \sigma) \min\left\{1, (1-p)^2 d_B^2(\rho, \sigma)\right\}, && \text{(C.91)}
\end{aligned}
$$

where the last equality follows from (C.85).

This leads to the inequality

$$
d_B^2(\mathcal{A}(\rho), \mathcal{A}(\sigma)) \geq \frac{1}{4} d_B^2(\rho, \sigma) \tag{C.92}
$$

if

$$
(1-p)^2 \geq \frac{1}{d_B^2(\rho, \sigma)}. \tag{C.93}
$$

By data processing, for all $\mathcal{A} \in \mathcal{B}^{\varepsilon}$, we have

$$
d_B^2(\mathcal{A}(\rho), \mathcal{A}(\sigma)) \leq d_B^2(\rho, \sigma). \tag{C.94}
$$

Combining both these bounds, we have

$$
\frac{1}{d_B^2(\rho, \sigma)} \leq \inf_{\mathcal{A} \in \mathcal{B}^{\varepsilon}} \frac{1}{d_B^2(\mathcal{A}(\rho), \mathcal{A}(\sigma))} \leq \frac{4}{d_B^2(\rho, \sigma)} \tag{C.95}
$$

if

$$
\left(\frac{e^{\varepsilon} - 1}{e^{\varepsilon} + 1}\right)^2 \geq \frac{1}{d_B^2(\rho, \sigma)}. \tag{C.96}
$$

Then under this constraint on ε, together with (6.9) and (6.11) we have that

$$
\mathrm{SC}^{\varepsilon}_{(\rho,\sigma)} = \Theta\left(\frac{1}{d_B^2(\rho, \sigma)}\right). \tag{C.97}
$$

Next, we consider the case where ρ and σ are general quantum states. Recall that k is the number of POVM elements formed by the eigenbasis of $\rho\#\sigma^{-1}$. Let $\mathcal{N}$ be a classical channel with input dimension k and output dimension two. By applying [106, Corollary 3.4] and considering that $\mathcal{M}' = \mathcal{N} \circ \mathcal{M}$ is a two-outcome measurement channel, we have that

$$1 \leq \frac{d_B^2(\mathcal{M}(\rho), \mathcal{M}(\sigma))}{d_B^2(\mathcal{M}'(\rho), \mathcal{M}'(\sigma))} \leq 1800 \max\left\{1, \frac{\min\{k, k'\}}{2}\right\}, \tag{C.98}$$

where $k' := \ln\left(4/d_B^2(\rho, \sigma)\right)$.

With the choice $\mathcal{A}' := \mathcal{A}^p_{\text{Dep}} \circ \mathcal{M}'$ and $p = 2/(e^\varepsilon + 1)$ following a similar approach as done from (C.87)-(C.90), we find that

$$\begin{aligned}
d_B^2(\mathcal{A}'(\rho), \mathcal{A}'(\sigma)) &\geq [T(\mathcal{A}'(\rho), \mathcal{A}'(\sigma))]^2 && \text{(C.99)}\\
&= (1-p)^2 [T(\mathcal{M}'(\rho), \mathcal{M}'(\sigma))]^2 && \text{(C.100)}\\
&\geq \frac{(1-p)^2}{4} \left[d_B^2(\mathcal{M}'(\rho), \mathcal{M}'(\sigma))\right]^2 && \text{(C.101)}\\
&\geq \frac{(1-p)^2}{4(1800)^2 L_{k,k'}} \left[d_B^2(\rho, \sigma)\right]^2 && \text{(C.102)}\\
&\geq \frac{1}{4(1800)^2 L_{k,k'}} d_B^2(\rho, \sigma) \min\left\{1, (1-p)^2 d_B^2(\rho, \sigma)\right\}, && \text{(C.103)}
\end{aligned}$$

where $L_{k,k'}$ is defined in (6.51).

If $(1-p)^2 \geq 1/d_B^2(\rho, \sigma)$, then we have that

$$d_B^2(\mathcal{A}'(\rho), \mathcal{A}'(\sigma)) \geq \frac{1}{4(1800)^2 L_{k,k'}} d_B^2(\rho, \sigma)\,. \tag{C.104}$$

This leads to

$$\frac{1}{d_B^2(\rho, \sigma)} \leq \inf_{\mathcal{A} \in \mathcal{B}^\varepsilon} \frac{1}{d_B^2(\mathcal{A}(\rho), \mathcal{A}(\sigma))} \leq \frac{4(1800)^2 L_{k,k'}}{d_B^2(\rho, \sigma)}. \tag{C.105}$$

Again together with (6.9) and (6.11), we conclude the proof.

C.5 Proof of the Lower Bound in Remark 36

In this appendix, we prove the lower bound given in (6.54).

First, we prove the following:

$$p_e(\rho, \sigma, p, q) \geq 2 \ \min\{p, q\} \ p_e(\rho, \sigma, 1/2, 1/2) \tag{C.106}$$

To obtain that, consider the following chain of relations:

$$\begin{aligned}
&p_e(\rho, \sigma, p, q) \\
&= \min_{\substack{M_1, M_2 \geq 0: \\ M_1 + M_2 = I}} p \operatorname{Tr}[M_2 \rho] + q \operatorname{Tr}[M_1 \sigma] && \text{(C.107)} \\
&\geq 2 \min\{p, q\} \min_{\substack{M_1, M_2 \geq 0: \\ M_1 + M_2 = I}} \frac{1}{2} \operatorname{Tr}[M_2 \rho] + \frac{1}{2} \operatorname{Tr}[M_1 \sigma] && \text{(C.108)} \\
&= 2 \min\{p, q\} \ p_e(\rho, \sigma, 1/2, 1/2). && \text{(C.109)}
\end{aligned}$$

Let the fixed error probability be α. Then, in the private setting choose $\mathcal{A}_i \in \mathcal{B}^{\varepsilon}$ for all $i \in \{1, \ldots, n\}$. With that we have

$$\begin{aligned}
\alpha &\geq p_e(\mathcal{A}_1(\rho) \otimes \cdots \otimes \mathcal{A}_n(\rho), \mathcal{A}_1(\sigma) \otimes \cdots \otimes \mathcal{A}_n(\sigma), p, q) && \text{(C.110)} \\
&\geq 2 \min\{p, q\} \ p_e(\mathcal{A}_1(\rho) \otimes \cdots \otimes \mathcal{A}_n(\rho), \mathcal{A}_1(\sigma) \otimes \cdots \otimes \mathcal{A}_n(\sigma), 1/2, 1/2) && \text{(C.111)} \\
&= \min\{p, q\} \left(1 - \frac{1}{2} \left\| \mathcal{A}_1(\rho) \otimes \cdots \otimes \mathcal{A}_n(\rho) - \mathcal{A}_1(\sigma) \otimes \cdots \otimes \mathcal{A}_n(\sigma) \right\|_1 \right). && \text{(C.112)}
\end{aligned}$$

Rearranging the last inequality yields the following:

$$T(\mathcal{A}_1(\rho) \otimes \cdots \otimes \mathcal{A}_n(\rho), \mathcal{A}_1(\sigma) \otimes \cdots \otimes \mathcal{A}_n(\sigma)) \geq 1 - \frac{\alpha}{\min\{p, q\}} =: \beta. \tag{C.113}$$

Applying Pinsker's inequality, we get

$$\begin{aligned}
&D(\mathcal{A}_1(\rho) \otimes \cdots \otimes \mathcal{A}_n(\rho) \| \mathcal{A}_1(\sigma) \otimes \cdots \otimes \mathcal{A}_n(\sigma)) \\
&\qquad \geq 2 \left[T(\mathcal{A}_1(\rho) \otimes \cdots \otimes \mathcal{A}_n(\rho), \mathcal{A}_1(\sigma) \otimes \cdots \otimes \mathcal{A}_n(\sigma)) \right]^2 . && \text{(C.114)}
\end{aligned}$$

Together with the former, consider that

$$2\beta^2 \leq D(\mathcal{A}_1(\rho) \otimes \cdots \otimes \mathcal{A}_n(\rho) \| \mathcal{A}_1(\sigma) \otimes \cdots \otimes \mathcal{A}_n(\sigma)) \tag{C.115}$$

$$= \sum_{i=1}^{n} D(\mathcal{A}_i(\rho) \| \mathcal{A}_i(\sigma)) \tag{C.116}$$

$$\leq \sum_{i=1}^{n} \varepsilon \frac{e^{\varepsilon} - 1}{e^{\varepsilon} + 1} T(\rho, \sigma) \tag{C.117}$$

$$\leq n\varepsilon \frac{e^{\varepsilon} - 1}{e^{\varepsilon} + 1} T(\rho, \sigma), \tag{C.118}$$

where the first equality follows from the additivity of quantum relative entropy and the first inequality follows from Proposition 25 with the assumption $\mathcal{A} \in \mathcal{B}^{\varepsilon}$.

Thus, by rearranging the terms, we conclude the proof of the lower bound in (6.54).

C.6 Other Variants of QLDP and Relation to ε-QLDP

In Chapter 6, we discussed the impact of privacy imposed by ε-QLDP with $\delta = 0$ on the sample complexity of private quantum hypothesis testing. Next, we show several relations between ε'-QLDP and (ε, δ)-QLDP framework when $\delta > 0$, so that we may use the derived results from Section 6.2.2 to infer about the impact of privacy imposed by (ε, δ)-QLDP mechanisms.

From [95, Lemma 5], we have that $(\varepsilon + \delta, 0)$-QLDP implies that (ε, δ')-QLDP, where

$$\delta' := 1 - \frac{e^{\varepsilon} + 1}{e^{\varepsilon + \delta} + 1}. \tag{C.119}$$

However, in this case δ' is dependent on ε. Next, we show another relation between the two frameworks without this dependence.

Proposition 45. *Every $\mathcal{A}$ that satisfies $(\varepsilon+\delta, 0)$-QLDP also satisfies (ε, δ)-QLDP. i.e.,*

$$(\varepsilon + \delta, 0)\text{−QLDP} \quad \Longrightarrow \quad (\varepsilon, \delta)\text{−QLDP}. \tag{C.120}$$

Proof. Proof follows analogous to the proof of one implication of [2, Lemma 5] established for classical differential privacy. Fix $0 \leq M \leq I$ and assume that $\mathcal{A}$ satisfies (ε, δ)-QLDP. Then, consider

$$\begin{aligned}
\mathrm{Tr}[M\mathcal{A}(\rho)] &\leq e^{\varepsilon+\delta}\mathrm{Tr}[M\mathcal{A}(\sigma)] && \text{(C.121)}\\
&= e^{\varepsilon}e^{\delta}\mathrm{Tr}[M\mathcal{A}(\sigma)] + e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)] - e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)] && \text{(C.122)}\\
&= e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)] + (e^{\delta}-1)e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)]\,. && \text{(C.123)}
\end{aligned}$$

If $e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)] + \delta \geq 1$, then the following holds:

$$\mathrm{Tr}[M\mathcal{A}(\rho)] \leq 1 \leq e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)] + \delta. \tag{C.124}$$

Now assuming $e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)] + \delta < 1$, we have $e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)] < (1-\delta)$. With this, consider

$$\begin{aligned}
(e^{\delta}-1)e^{\varepsilon}\mathrm{Tr}[M\mathcal{A}(\sigma)] &\leq (e^{\delta}-1)(1-\delta) && \text{(C.125)}\\
&\leq (e^{\delta}-1)e^{-\delta} && \text{(C.126)}\\
&= 1 - e^{-\delta} && \text{(C.127)}\\
&\leq \delta, && \text{(C.128)}
\end{aligned}$$

where the second and the last inequality follows from $1-x \leq e^{-x}$ for all $x \in \mathbb{R}$. □

With the above implication together with Theorem 8, we have

$$\mathrm{SC}^{\varepsilon,\delta}_{(\rho,\sigma)} \coloneqq \inf_{\mathcal{A}\in\mathcal{B}^{\varepsilon,\delta}} \mathrm{SC}^{\mathcal{A}}_{(\rho,\sigma)} \leq \inf_{\mathcal{A}\in\mathcal{B}^{\varepsilon+\delta}} \mathrm{SC}^{\mathcal{A}}_{(\rho,\sigma)} \leq \left\lceil 2\ln\left(\frac{\sqrt{pq}}{\alpha}\right)\left(\frac{(e^{\varepsilon+\delta}+1)}{(e^{\varepsilon+\delta}-1)T(\rho,\sigma)}\right)^2\right\rceil, \tag{C.129}$$

where $\mathcal{B}^{\varepsilon,\delta}$ is the set of all (ε, δ)-QLDP mechanisms. First inequality there follows from $(\varepsilon+\delta, 0)$-QLDP mechanisms are (ε, δ)-QLDP from Proposition 45, which

makes the latter a larger set of mechanisms. The second inequality follows from the upper bound in Theorem 8.

Inspired by [70, Lemma 3.2] for classical differential privacy, next we show that (ε, δ)-QLDP mechanism also implies $(\varepsilon', 0)$-QLDP for another mechanism that depends on the original mechanism.

Proposition 46. *Let $\mathcal{A}$ satisfy (ε, δ)-QLDP. Then, there exists $\mathcal{A}'$ such that $T(\mathcal{A}(\omega), \mathcal{A}'(\omega)) \leq \eta$ for $\eta \in [0, 1]$ $\mathcal{A}'$ satisfies $(\varepsilon', 0)$-QLDP, where*

$$\varepsilon' := \varepsilon + \ln\left(1 + \frac{d\delta}{\eta} e^{-\varepsilon}\right) \tag{C.130}$$

with d being the dimension of the Hilbert space $\mathcal{H}$.

Proof. Choose $\mathcal{A}' = \mathcal{A}^{\eta}_{\text{Dep}} \circ \mathcal{A}$ such that

$$\mathcal{A}'(\omega) = (1 - \eta)\mathcal{A}(\omega) + \eta \frac{I}{d}, \tag{C.131}$$

for $\eta \in (0, 1)$. For that choice,

$$T(\mathcal{A}(\omega), \mathcal{A}'(\omega)) = \eta T\left(\mathcal{A}(\omega), \frac{I}{d}\right) \tag{C.132}$$

$$\leq \eta, \tag{C.133}$$

where the last inequality follows from the normalized trace distance being bounded from above by one.

Since $\mathcal{A}$ satisfies (ε, δ)-QLDP and due to the data-processing property of QLDP mechanisms, $\mathcal{A}'$ also satisfies (ε, δ)-QLDP. This leads to the following inequality for all $0 \leq M \leq I$ and states ρ, σ:

$$\text{Tr}[M\mathcal{A}'(\rho)] \leq e^{\varepsilon} \text{Tr}[M\mathcal{A}'(\sigma)] + \delta. \tag{C.134}$$

By choosing a rank one projection P, we obtain the above inequality by replacing M with P. For that choice we have

$$\begin{aligned}\mathrm{Tr}[P\mathcal{A}'(\sigma)] &= \mathrm{Tr}\left[P\left((1-\eta)\mathcal{A}(\sigma)+\eta\frac{I}{d}\right)\right] && \text{(C.135)}\\ &= (1-\eta)\mathrm{Tr}[P\mathcal{A}(\sigma)]+\frac{\eta\mathrm{Tr}[P]}{d} && \text{(C.136)}\\ &\geq \frac{\eta}{d}, && \text{(C.137)}\end{aligned}$$

where the last inequality follows from $(1-\eta)\mathrm{Tr}[P\mathcal{A}(\sigma)] \geq 0$ and $\mathrm{Tr}[P] = 1$.

Now, let us consider

$$\begin{aligned}\mathrm{Tr}[P\mathcal{A}'(\rho)] &\leq e^{\varepsilon}\mathrm{Tr}[P\mathcal{A}'(\sigma)]+\delta && \text{(C.138)}\\ &= e^{\varepsilon}\mathrm{Tr}[P\mathcal{A}'(\sigma)]+\frac{d\delta}{\eta}\times\frac{\eta}{d} && \text{(C.139)}\\ &\leq e^{\varepsilon}\mathrm{Tr}[P\mathcal{A}'(\sigma)]+\frac{d\delta}{\eta}\mathrm{Tr}[P\mathcal{A}'(\sigma)] && \text{(C.140)}\\ &= \left(e^{\varepsilon}+\frac{d\delta}{\eta}\right)\mathrm{Tr}[P\mathcal{A}'(\sigma)], && \text{(C.141)}\end{aligned}$$

where the penultimate inequality follows from (C.137).

The above holds for all rank-one projections. For $M = 0$, the above inequality holds trivially. We can write a general measurement $0 < M \leq I$ using spectral theorem as $M = \sum_i \lambda_i P_i$ with $\lambda_i > 0$ and P_i for all i are rank one projections. Having $\mathrm{Tr}[P_i\mathcal{A}'(\rho)] \leq \left(e^{\varepsilon}+\frac{d\delta}{\eta}\right)\mathrm{Tr}[P_i\mathcal{A}'(\sigma)]$ for all i, then multiplying both sides by λ_i, and summing over all i, we arrive at

$$\sum_i \lambda_i \mathrm{Tr}[P_i\mathcal{A}'(\rho)] \leq \left(e^{\varepsilon}+\frac{d\delta}{\eta}\right)\sum_i \lambda_i \mathrm{Tr}[P_i\mathcal{A}'(\sigma)]. \tag{C.142}$$

This is equivalent to

$$\mathrm{Tr}[M\mathcal{A}'(\rho)] \leq \left(e^{\varepsilon}+\frac{d\delta}{\eta}\right)\mathrm{Tr}[M\mathcal{A}'(\sigma)], \tag{C.143}$$

proving that $\mathcal{A}'$ satisfies $(\varepsilon', 0)$-QLDP with

$$\varepsilon' := \ln\left(e^{\varepsilon}+\frac{d\delta}{\eta}\right). \tag{C.144}$$

Simplifying the above with the identity $\ln(ab) = \ln(a) + \ln(b)$ for positive a and b, provides the desired result concluding the proof. □

However, it is not quite clear how to use Proposition 46 directly to obtain a lower bound on the sample complexity of quantum hypothesis testing under (ε, δ)-QLDP. We leave this as an open question.

BIBLIOGRAPHY

[1] Scott Aaronson and Guy N. Rothblum. Gentle measurement of quantum states and differential privacy. In *Proceedings of ACM SIGACT Symposium on Theory of Computing*, pages 322–333, 2019.

[2] Jayadev Acharya, Ziteng Sun, and Huanyu Zhang. Differentially private testing of identity and closeness of discrete distributions. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 6879–6891, 2018.

[3] Dorit Aharonov, Vaughan Jones, and Zeph Landau. A polynomial quantum algorithm for approximating the Jones polynomial. *Algorithmica*, 55(3):395–421, 2009.

[4] Armando Angrisani, Mina Doosti, and Elham Kashefi. Differential privacy amplification in quantum and quantum-inspired algorithms. arXiv:2203.03604, 2022.

[5] Armando Angrisani, Mina Doosti, and Elham Kashefi. A unifying framework for differentially private quantum algorithms, 2023. arXiv:2307.04733.

[6] Armando Angrisani and Elham Kashefi. Quantum local differential privacy and quantum statistical query model. arXiv:2203.03591, 2022.

[7] Shahab Asoodeh, Maryam Aliakbarpour, and Flavio P Calmon. Local differential privacy is equivalent to contraction of E_γ-divergence. *arXiv preprint arXiv:2102.01258*, 2021.

[8] Shahab Asoodeh and Huanyu Zhang. Contraction of locally differentially private mechanisms. *IEEE Journal on Selected Areas in Information Theory*, pages 1–1, 2024. arXiv:2210.13386.

[9] Koenraad M. R. Audenaert, John Calsamiglia, Ramón Munoz-Tapia, Emilio Bagan, Lluis Masanes, Antonio Acin, and Frank Verstraete. Discriminating states: The quantum Chernoff bound. *Physical Review Letters*, 98(16):160501, 2007.

[10] Koenraad M. R. Audenaert, Michael Nussbaum, Arleta Szkoła, and Frank Verstraete. Asymptotic error rates in quantum hypothesis testing. *Communications in Mathematical Physics*, 279:251–283, 2008.

[11] Joonwoo Bae and Leong-Chuan Kwek. Quantum state discrimination and its applications. *Journal of Physics A: Mathematical and Theoretical*, 48(8):083001, 2015.

[12] Raef Bassily, Adam Groce, Jonathan Katz, and Adam Smith. Coupled-worlds privacy: Exploiting adversarial uncertainty in statistical data privacy. In *Proceedings of IEEE Symposium on Foundations of Computer Science*, pages 439–448. IEEE, 2013.

[13] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S. Kottmann, Tim Menke, Wai-Keong Mok, Sukin Sim, Leong-Chuan Kwek, and Alan Aspuru-Guzik. Noisy intermediate-scale quantum (NISQ) algorithms. *Reviews of Modern Physics*, 94(1):015004, February 2022. arXiv:2101.08448.

[14] Olivier Bousquet and André Elisseeff. Stability and generalization. *The Journal of Machine Learning Research*, 2:499–526, 2002.

[15] Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Proceedings of Theory of Cryptography Conference*, pages 635–658. Springer, 2016.

[16] Clément L. Canonne, Gautam Kamath, Audra McMillan, Adam Smith, and Jonathan Ullman. The structure of optimal private tests for simple hypotheses. In *Proceedings of the Annual ACM SIGACT Symposium on Theory of Computing*, page 310–321, 2019.

[17] Rachel Mary Cardell-Oliver and Brandon Ke. Towards an activity-aware pufferfish framework for local privacy of household smart water meter data. In *Proceedings of the ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, pages 328–332, 2023.

[18] Matthias C. Caro, Tom Gur, Cambyse Rouzé, Daniel Stilck França, and Sathyawageeswar Subramanian. Information-theoretic generalization bounds for learning from quantum data. In *Proceedings of Thirty Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pages 775–839. PMLR, 30 Jun–03 Jul 2024. arXiv:2311.05529v1.

[19] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan,

Lukasz Cincio, and Patrick J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3:625–644, August 2021. arXiv:2012.09265.

[20] Jingxuan Chen, Hanna Westerheim, Zoë Holmes, Ivy Luo, Theshani Nuradha, Dhrumil Patel, Soorya Rethinasamy, Kathie Wang, and Mark M Wilde. Qslack: A slack-variable approach for variational quantum semidefinite programming. arXiv:2312.03830, 2023.

[21] Ranyiliu Chen, Zhixin Song, Xuanqiang Zhao, and Xin Wang. Variational quantum algorithms for trace distance and fidelity estimation. *Quantum Science and Technology*, 7(1):015–019, 2021.

[22] Hao-Chung Cheng, Nilanjana Datta, Nana Liu, Theshani Nuradha, Robert Salzmann, and Mark M. Wilde. An invitation to the sample complexity of quantum hypothesis testing. arXiv:2403.17868v3, 2024.

[23] Hao-Chung Cheng, Christoph Hirche, and Cambyse Rouzé. Sample complexity of locally differentially private quantum hypothesis testing, 2024. arXiv:2406.18658.

[24] Hao-Chung Cheng, Andreas Winter, and Nengkun Yu. Discrimination of quantum states under locality constraints in the many-copy setting. *Communications in Mathematical Physics*, 404(1):151–183, 2023.

[25] Herman Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *The Annals of Mathematical Statistics*, pages 493–507, 1952.

[26] Giulio Chiribella, G Mauro D'Ariano, and Paolo Perinotti. Transforming quantum operations: Quantum supermaps. *Europhysics Letters*, 83(3):30004, 2008.

[27] Paul Cuff and Lanqing Yu. Differential privacy as a mutual information constraint. In *Proceedings of ACM SIGSAC Conference on Computer and Communications Security*, pages 43–54, 2016.

[28] Nilanjana Datta. Min-and max-relative entropies and a new entanglement monotone. *IEEE Transactions on Information Theory*, 55(6):2816–2826, 2009.

[29] Nilanjana Datta and Felix Leditzky. Second-order asymptotics for source coding, dense coding, and pure-state entanglement conversions. *IEEE Transactions on Information Theory*, 61(1):582–608, 2014.

[30] E. Brian Davies and John T. Lewis. An operational approach to quantum probability. *Communications in Mathematical Physics*, 17(3):239–260, September 1970.

[31] Edward B. Davies. *Quantum Theory of Open Systems*. Academic Press, 1976.

[32] 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.

[33] Zeyu Ding, Yuxin Wang, Guanhong Wang, Danfeng Zhang, and Daniel Kifer. Detecting violations of differential privacy. In *Proceedings of ACM SIGSAC Conference on Computer and Communications Security*, pages 475–489, 2018.

[34] David P. DiVincenzo, Debbie W. Leung, and Barbara M. Terhal. Quantum data hiding. *IEEE Transactions on Information Theory*, 48(3):580–598, 2002.

[35] Carles Domingo-Enrich and Youssef Mroueh. Auditing differential privacy in high dimensions with the kernel quantum Rényi divergence. *arXiv preprint arXiv:2205.13941*, 2022.

[36] Matthew J. Donald. On the relative entropy. *Communications in Mathematical Physics*, 105:13–34, 1986.

[37] Yuxuan Du, Min-Hsiu Hsieh, Tongliang Liu, Dacheng Tao, and Nana Liu. Quantum noise protects quantum classifiers against adversaries. *Physical Review Research*, 3(2):023153, 2021.

[38] John C. Duchi, Michael I. Jordan, and Martin J. Wainwright. Local privacy and statistical minimax rates. In *IEEE Annual Symposium on Foundations of Computer Science*, pages 429–438, 2013.

[39] John C Duchi, Michael I Jordan, and Martin J Wainwright. Minimax optimal procedures for locally private estimation. *Journal of the American Statistical Association*, 113(521):182–201, 2018.

[40] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Proceedings of Conference on Theory of Cryptography, TCC*, pages 265–284, 2006.

[41] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science (FnT-TCS)*, 9(3-4):211–407, 2014.

[42] T. Eggeling and R. F. Werner. Hiding classical data in multipartite quantum states. *Physical Review Letters*, 89(9):097905, August 2002.

[43] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of ACM SIGSAC Conference on Computer and Communications Security*, pages 1054–1067, 2014.

[44] Alexandre Evfimievski, Johannes Gehrke, and Ramakrishnan Srikant. Limiting privacy breaches in privacy preserving data mining. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 211–222, 2003.

[45] Anthony W Flores, Kristin Bechtel, and Christopher T Lowenkamp. False positives, false negatives, and false analyses: A rejoinder to machine bias: There's software used across the country to predict future criminals. and it's biased against blacks. *Fed. Probation*, 80:38, 2016.

[46] Rupert L. Frank and Elliott H. Lieb. Monotonicity of a relative Rényi entropy. *Journal of Mathematical Physics*, 54:122201, 2013.

[47] C.A. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, may 1999.

[48] Christopher Fuchs. *Distinguishability and Accessible Information in Quantum Theory*. PhD thesis, University of New Mexico, December 1996. arXiv:quant-ph/9601020.

[49] Christopher A Fuchs and Carlton M Caves. Mathematical techniques for quantum communication theory. *Open Systems & Information Dynamics*, 3(3):345–356, 1995.

[50] Ian George, Christoph Hirche, Theshani Nuradha, and Mark M Wilde. Quantum doeblin coefficients: Interpretations and applications. arXiv:2503.22823, 2025.

[51] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum

singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 193–204, 2019.

[52] Ziv Goldfeld, Kristjan Greenewald, Theshani Nuradha, and Galen Reeves. k-sliced mutual information: a quantitative study of scalability with dimension. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22, 2024.

[53] Gilad Gour. Comparison of quantum channels by superchannels. *IEEE Transactions on Information Theory*, 65(9):5880–5904, 2019.

[54] Ji Guan. Optimal mechanisms for quantum local differential privacy. arXiv:2407.13516, 2024.

[55] Ji Guan, Wang Fang, and Mingsheng Ying. Verifying fairness in quantum machine learning. In *Proceedings of International Conference on Computer Aided Verification*, pages 408–429. Springer, 2022.

[56] Samuel Haney, Ashwin Machanavajjhala, John M Abowd, Matthew Graham, Mark Kutzbach, and Lars Vilhuber. Utility cost of formal privacy for releasing national employer-employee statistics. In *Proceedings of the ACM International Conference on Management of Data*, pages 1339–1354, 2017.

[57] Patrick Hayden, Debbie Leung, Peter W. Shor, and Andreas Winter. Randomizing quantum states: Constructions and applications. *Communications in Mathematical Physics*, 250(2):371–391, 2004.

[58] Patrick Hayden, Debbie Leung, and Graeme Smith. Multiparty data hiding of quantum information. *Physical Review A*, 71(6):062339, June 2005.

[59] Carl W. Helstrom. Detection theory and quantum mechanics. *Information and Control*, 10(3):254–291, 1967.

[60] Fumio Hiai and Milán Mosonyi. Different quantum f-divergences and the reversibility of quantum operations. *Reviews in Mathematical Physics*, 29(07):1750023, 2017.

[61] Fumio Hiai and Dénes Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physics*, 143:99–114, 1991.

[62] Christoph Hirche, Cambyse Rouzé, and Daniel Stilck França. Quantum differential privacy: An information theory perspective. *IEEE Transactions on Information Theory*, 69(9):5771–5787, 2023. arXiv:2202.10717.

[63] Christoph Hirche and Marco Tomamichel. Quantum Rényi and f-divergences from integral representations. *Communications in Mathematical Physics*, 405(9):208, 2024. arXiv:2306.12343.

[64] Wassily Hoeffding. Asymptotically optimal tests for multinomial distributions. *The Annals of Mathematical Statistics*, pages 369–401, 1965.

[65] Alexander S. Holevo. Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*, 3(4):337–394, 1973.

[66] Michał Horodecki and Paweł Horodecki. Reduction criterion of separability and limits for a class of distillation protocols. *Physical Review A*, 59(6):4206–4216, June 1999. arXiv:quant-ph/9708015.

[67] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Physical Review A*, 60:1888–1898, September 1999.

[68] Hsin-Yuan Huang. Learning quantum states from their classical shadows. *Nature Reviews Physics*, 4(2):81–81, 2022.

[69] 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 *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2023.

[70] Yiyang Huang and Clément L. Canonne. Lemmas of differential privacy. *arXiv preprint arXiv:2211.11189*, 2022.

[71] Matthew Jagielski, Jonathan Ullman, and Alina Oprea. Auditing differentially private machine learning: How private is private SGD? *Advances in Neural Information Processing Systems*, 33:22205–22216, 2020.

[72] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. Extremal mechanisms for local differential privacy. In *Advances in Neural Information Processing Systems*, volume 27, pages 1–9, 2014.

[73] 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.

[74] Stephan Kessler, Erik Buchmann, and Klemens Böhm. Deploying and evaluating pufferfish privacy for smart meter data. In *IEEE Intl Conf on Ubiquitous Intelligence and Computing and IEEE Intl Conf on Autonomic and Trusted Computing and IEEE Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom)*, pages 229–238, 2015.

[75] Sumeet Khatri and Mark M. Wilde. Principles of quantum communication theory: A modern approach. arXiv:2011.04672v2, 2020.

[76] D. Kifer and Bing-R. Lin. An axiomatic view of statistical privacy and utility. *Journal of Privacy and Confidentiality*, 4(1), 2012.

[77] D. Kifer and A. Machanavajjhala. Pufferfish: A framework for mathematical privacy definitions. *ACM Transactions on Database Systems*, 39(1):1–36, 2014.

[78] Robbie King, David Gosset, Robin Kothari, and Ryan Babbush. Triply efficient shadow tomography. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 914–946. SIAM, 2025.

[79] Alexei Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52:1191–1249, 1997.

[80] Ludovico Lami, Carlos Palazuelos, and Andreas Winter. Ultimate data hiding in quantum mechanics and beyond. *Communications in Mathematical Physics*, 361(2):661–708, 2018.

[81] Wenjuan Liang, Hong Chen, Ruixuan Liu, Yuncheng Wu, and Cuiping Li. A pufferfish privacy mechanism for monitoring web browsing behavior under temporal correlations. *Computers & Security*, 92:101754, 2020.

[82] Elliott H. Lieb and Mary Beth Ruskai. Proof of the strong subadditivity of quantum-mechanical entropy. *Journal of Mathematical Physics*, 14(12):1938–1941, December 1973.

[83] Göran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40:147–151, June 1975.

[84] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum principal component analysis. *Nature Physics*, 10(9):631–633, July 2014.

[85] Cosmo Lupo, Mark M. Wilde, and Seth Lloyd. Quantum data hiding in the presence of noise. *IEEE Transactions on Information Theory*, 62(6):3745–3756, 2016.

[86] Keiji Matsumoto. A new quantum version of f-divergence. In *Reality and Measurement in Algebraic Quantum Theory*, volume 261, pages 229–273, Singapore, 2018. Springer Singapore. Series Title: Springer Proceedings in Mathematics & Statistics.

[87] William Matthews, Stephanie Wehner, and Andreas Winter. Distinguishability of quantum states under restricted families of measurements with an application to quantum data hiding. *Communications in Mathematical Physics*, 291:813–843, 2009.

[88] Ilya Mironov. Rényi differential privacy. In *Proceedings of IEEE computer security foundations symposium (CSF)*, pages 263–275, 2017.

[89] 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, 2013.

[90] Takao Murakami and Yusuke Kawamoto. Utility-optimized local differential privacy mechanisms for distribution estimation. In *Proceedings of USENIX Security Symposium*, pages 1877–1894, 2019.

[91] Sloan Nietert, Ziv Goldfeld, and Rachel Cummings. Outlier-robust optimal transport: Duality, structure, and statistical analysis. In *Proceedings of The International Conference on Artificial Intelligence and Statistics*, volume 151, pages 11691–11719, 2022.

[92] Theshani Nuradha and Ziv Goldfeld. An information-theoretic characterization of pufferfish privacy. In *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, pages 2005–2010. IEEE, 2022.

[93] Theshani Nuradha and Ziv Goldfeld. An information-theoretic characterization of pufferfish privacy. In *2022 IEEE International Symposium on Information Theory (ISIT)*, pages 2005–2010, 2022.

[94] Theshani Nuradha and Ziv Goldfeld. Pufferfish privacy: An information-

theoretic study. *IEEE Transactions on Information Theory*, 69(11):7336–7356, 2023.

[95] Theshani Nuradha, Ziv Goldfeld, and Mark M. Wilde. Quantum pufferfish privacy: A flexible privacy framework for quantum systems. *IEEE Transactions on Information Theory*, 70(8):5731–5762, August 2024. arXiv:2306.13054.

[96] Theshani Nuradha, Hemant K Mishra, Felix Leditzky, and Mark M Wilde. Multivariate fidelities. *Journal of Physics A: Mathematical and Theoretical*, 58(16):165304, April 2025.

[97] Theshani Nuradha, Vishal Singh, and Mark M Wilde. Measured hockey-stick divergence and its applications to quantum pufferfish privacy, 2025. arXiv:2501.12359.

[98] Theshani Nuradha and Mark M. Wilde. Fidelity-based smooth min-relative entropy: Properties and applications. *IEEE Transactions on Information Theory*, 70(6):4170–4196, 2024.

[99] Theshani Nuradha and Mark M. Wilde. Contraction of private quantum channels and private quantum hypothesis testing. *Accpeted to IEEE Transactions on Information Theory*, 2025. arXiv:2406.18651v2.

[100] Theshani Nuradha and Mark M Wilde. Query complexity of classical and quantum channel discrimination, 2025. arXiv:2504.12989.

[101] Michael Nussbaum and Arleta Szkoła. The Chernoff lower bound for symmetric quantum hypothesis testing. *The Annals of Statistics*, 37(2):1040–1057, 2009.

[102] Masanori Ohya and Dénes Petz. *Quantum Entropy and its Use*. Springer Science & Business Media, 1993.

[103] Masanao Ozawa. Quantum measuring processes of continuous observables. *Journal of Mathematical Physics*, 25(1):79–87, 1984.

[104] Balázs Pejó and Damien Desfontaines. *Guide to Differential Privacy Modifications: A Taxonomy of Variants and Extensions*. Springer Briefs in Computer Science Serries. Springer International Publishing AG, Cham, 2022.

[105] Ankit Pensia, Amir Reza Asadi, Varun Jog, and Po-Ling Loh. Simple binary hypothesis testing under local differential privacy and communication constraints. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 3229–3230. PMLR, 2023.

[106] Ankit Pensia, Varun Jog, and Po-Ling Loh. Communication-constrained hypothesis testing: Optimality, robustness, and reverse data processing inequalities. *IEEE Transactions on Information Theory*, 70(1):389–414, 2024.

[107] Ankit Pensia, Varun Jog, and Po-Ling Loh. The sample complexity of simple binary hypothesis testing. *arXiv preprint arXiv:2403.16981*, 2024.

[108] Dénes Petz. Quasi-entropies for States of a von Neumann Algebra. *Publications of the Research Institute for Mathematical Sciences*, 21:787–800, 1985.

[109] Dénes Petz. Quasi-entropies for finite quantum systems. *Reports in Mathematical Physics*, 23:57–65, 1986.

[110] Marco Piani. Relative entropy of entanglement and restricted measurements. *Physical Review Letters*, 103(16):160504, October 2009. arXiv:0904.2705 [quant-ph].

[111] Stefano Pirandola. Quantum reading of a classical digital memory. *Physical Review Lett.*, 106:090504, March 2011.

[112] Yihui Quek, Srinivasan Arunachalam, and John A Smolin. Private learning implies quantum stability. In *Proceedings of International Conference on Advances in Neural Information Processing Systems*, volume 34, pages 20503–20515, 2021.

[113] Maxim Raginsky, Alexander Rakhlin, Matthew Tsao, Yihong Wu, and Aolin Xu. Information-theoretic analysis of stability and bias of learning algorithms. In *2016 IEEE Information Theory Workshop (ITW)*, pages 26–30. IEEE, 2016.

[114] Bartosz Regula, Ludovico Lami, and Mark M. Wilde. Postselected quantum hypothesis testing. *Accepted for publication in IEEE Transactions on Information Theory*, 2022. arXiv:2209.10550.

[115] Soorya Rethinasamy, Rochisha Agarwal, Kunal Sharma, and Mark M. Wilde. Estimating distinguishability measures on quantum computers. *Physical Review A*, 108(1):012409, July 2023.

[116] Tobias Rippchen, Sreejith Sreekumar, and Mario Berta. Locally-measured Rényi divergences, 2024. arXiv:2405.05037.

[117] Robert Salzmann, Nilanjana Datta, Gilad Gour, Xin Wang, and Mark M. Wilde. Symmetric distinguishability as a quantum resource. *New Journal of Physics*, 23(8):083016, 2021.

[118] Makhamisa Senekane, Mhlambululi Mafu, and Benedict Molibeli Taele. Privacy-preserving quantum machine learning using differential privacy. In *Proceedings of IEEE AFRICON*, pages 1432–1435, 2017.

[119] Naresh Sharma and Naqueeb Ahmad Warsi. On the strong converses for the quantum channel capacity theorems. arXiv:1205.1712, 2012.

[120] Vishal Singh, Theshani Nuradha, and Mark M Wilde. Extendible quantum measurements and limitations on classical communication, 2024. arXiv:2412.18556.

[121] Shuang Song, Yizhen Wang, and Kamalika Chaudhuri. Pufferfish privacy mechanisms for correlated data. In *Proceedings of ACM SIGMOD*, pages 1291–1306, 2017.

[122] David Sutter, Volkher B Scholz, Andreas Winter, and Renato Renner. Approximate degradable quantum channels. *IEEE Transactions on Information Theory*, 63(12):7832–7844, 2017.

[123] Ryuji Takagi, Hiroyasu Tajima, and Mile Gu. Universal sampling lower bounds for quantum error mitigation. *Physical Review Letters*, 131(21):210602, 2023.

[124] Barbara M. Terhal, David P. DiVincenzo, and Debbie W. Leung. Hiding bits in Bell states. *Physical Review Letters*, 86(25):5807–5810, June 2001.

[125] Anthony C Thompson. On certain contraction mappings in a partially ordered vector space. *Proceedings of the American Mathematical Society*, 14(3):438–443, 1963.

[126] Marco Tomamichel. *Quantum Information Processing with Finite Resources: Mathematical Foundations*, volume 5. Springer, 2015.

[127] Armin Uhlmann. The 'Transition Probability' in the State Space of a *-Algebra. *Reports on Mathematical Physics*, 9:273–279, 1976.

[128] Qisheng Wang, Ji Guan, Junyi Liu, Zhicheng Zhang, and Mingsheng Ying. New quantum algorithms for computing quantum entropies and distances. *arXiv preprint arXiv:2203.13522*, 2022.

[129] Qisheng Wang and Zhicheng Zhang. Fast quantum algorithms for trace distance estimation. *IEEE Transactions on Information Theory*, 70(4):2720–2733, 2024.

[130] Xin Wang and Mark M. Wilde. Resource theory of asymmetric distinguishability. *Physical Review Research*, 1(3):033170, 2019.

[131] William M Watkins, Samuel Yen-Chi Chen, and Shinjae Yoo. Quantum machine learning with differential privacy. *Scientific Reports*, 13(1):2453, 2023.

[132] John Watrous. Limits on the power of quantum statistical zero-knowledge. *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 459–468, November 2002. arXiv:quant-ph/0202111.

[133] John Watrous. Semidefinite programs for completely bounded norms. *Theory of Computing*, 5(11):217–238, 2009.

[134] John Watrous. Zero-knowledge against quantum attacks. *SIAM Journal on Computing*, 39(1):25–58, 2009. arXiv:quant-ph/0511020.

[135] Reinhard F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Physical Review A*, 40(8):4277–4281, October 1989.

[136] Hanna Westerheim, Jingxuan Chen, Zoë Holmes, Ivy Luo, Theshani Nuradha, Dhrumil Patel, Soorya Rethinasamy, Kathie Wang, and Mark M Wilde. Dual-vqe: A quantum algorithm to lower bound the ground-state energy, 2023. arXiv:2312.03083.

[137] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, second edition, 2017.

[138] Mark M. Wilde. Optimized quantum f-divergences and data processing. *Journal of Physics A*, 51:374002, 2018.

[139] Mark M. Wilde, Andreas Winter, and Dong Yang. Strong converse for

the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Communications in Mathematical Physics*, 331:593–622, 2014.

[140] Behnoosh Zamanlooy and Shahab Asoodeh. Strong data processing inequalities for locally differentially private mechanisms. In *2023 IEEE International Symposium on Information Theory (ISIT)*, pages 1794–1799. IEEE, 2023.

[141] Yuying Zeng, Yingpeng Sang, Shunchao Luo, and Mingyang Song. A pufferfish privacy mechanism for the trajectory clustering task. In *Parallel Architectures, Algorithms and Programming*, pages 307–317, Singapore, 2021. Springer Singapore.

[142] Lin Zhang, Kaifeng Bu, and Junde Wu. A lower bound on the fidelity between two states in terms of their trace-distance and max-relative entropy. *Linear and Multilinear Algebra*, 64(5):801–806, 2016.

[143] Wanrong Zhang, Olga Ohrimenko, and Rachel Cummings. Attribute privacy: Framework and mechanisms. In *Proceedings of ACM Conference on Fairness, Accountability, and Transparency*, pages 757–766, 2022.

[144] Li Zhou and Mingsheng Ying. Differential privacy in quantum computation. In *Proceedings of IEEE Computer Security Foundations Symposium (CSF)*, pages 249–262. IEEE, 2017.