UNIVERSITY OF TECHNOLOGY SYDNEY

Applications of Matrix Spaces in Quantum Information and Computational Complexity

by

Yinan Li

A thesis submitted in partial fulfilment of the
requirements for the degree of Doctor of Philosophy

Supervisor: Dr. Runyao Duan

Co-supervisor: Dr. Youming Qiao

at the
Centre for Quantum Software and Information
Faculty of Engineering and Information Technology
University of Technology Sydney

May 2018

Certificate of Original Authorship

I, Yinan Li, declare that this thesis, is submitted in fulfillment of the requirements for the award of Doctor of Philosophy, in the School of Software, Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualification at any other academic institution.

Signature:

Production Note:
Signature removed prior to publication.

Date: 31-05-2018

Applications of Matrix Spaces in Quantum Information and Computational Complexity

by

Yinan Li

Abstract

This thesis explores the applications of matrix spaces in quantum information and computational complexity, specifically in the areas of quantum state/channel discrimination, entanglement transformation and isomorphism testing. We achieve the following contributions:

- We derive a necessary condition which determines whether a set of orthogonal bipartite states can be distinguished by positive-partial-transpose (PPT) operations, in the many-copy scenario. We reduce the discrimination task to the following problem: Decide whether there exist a nonzero bipartite matrix with positive partial-transpose, such that all its eigenvectors with nonzero eigenvalues is orthogonal to a given bipartite vector space. As applications, we reprove and extend a variety of existing results, including the local indistinguishability of the bipartite maximally entangled state and its orthogonal complement [Yu, Duan and Ying, *IEEE transaction on Information Theory*, 2014] and the maximum dimension of non-positive-partial-transpose subspaces [Johnston, *Physical Review A*, 2013].

- We show that determining the parallel distinguishability of quantum channels is equivalent to determining whether the orthogonal complement of a given matrix space contains nonzero positive semidefinite matrices. Our characterization immediately implies a necessary condition to decide the parallel distinguishability. We further prove that our condition is also sufficient for two large families of quantum channels, which leads to an alternating proof for the parallel distinguishability of

unitary operations [Acín, *Physical Review Letters*, 2001]. We also present an illustrative example showing our necessary condition cannot determine the parallel distinguishability.

- We systematically study the tripartite-to-bipartite entanglement transformations by stochastic local operations and classical communication (SLOCC). Such a problem is equivalent to computing the maximal rank of a matrix space [Chitambar, Duan and Shi, *Physical Review A*, 2010]. We analyze the SLOCC convertibility in both the finite-copy and asymptotic regimes. In particular, we derive explicit formulas which compute asymptotic entanglement transformation rates for two families of tripartite states by resorting to certain results of the structure of matrix spaces, including the study of matrix semi-invariants. Significantly, we show that the problem of deciding the asymptotic SLOCC convertibility of tripartite pure states to the bipartite maximally entangled states and the non-commutative symbolic determinant identity testing problem is algorithmically equivalent, which builds a new connection between problems in algebraic complexity and problems in asymptotic SLOCC entanglement transformations.

- We devise algorithms which test isometry between alternating matrix spaces over finite field. Algorithms for such a problem in time polynomial in the underlying vector space size resolves the believed bottleneck case of the group isomorphism problem, i.e. testing isomorphism of p-groups of class 2 and exponent p in time polynomial in the group order. Our approach is to view it as a linear algebraic analog of the graph isomorphism problem. We devise an average-case efficient algorithm for the alternating matrix space isometry problem over a key range of parameters. in a random model of alternating matrix spaces in vein of the Erdős-Rényi model of random graphs. To devise our algorithm, we developed linear algebraic individualization and refinement techniques, which are crucial in the first average-case efficient algorithm for graph isomorphism, devised by Babai, Erdős, and Selkow in the 1970s [Babai, Erdős and Selkow, *SIAM Journal on Computing*, 1980]. We also adapt Luks' dynamic programming technique for graph isomorphism [Luks, *STOC*, 1999] to slightly improve the worst-case time complexity of the alternating matrix space isometry problem.

Acknowledgements

Very little of the work in this thesis would have been possible without the support of my supervisors, collaborators, colleagues, friends and family.

First and foremost, I would like to thank my supervisor, Runyao Duan. Without his guidance and insight, I would not have complete my PhD study. I am greatly indebted to him for leading me into the quantum world. Dating back to 2013 in Wuhan University, when I was still a third-year undergraduate student, I was so inspired by his talk about quantum computation. He encouraged me to investigate the parallel distinguishability of quantum channels, which becomes the departure point of my research career. Furthermore, I am grateful for Runyao's spiritual support for my life. He is not only my academic supervisor but also a spiritual mentor, sharing his experience and scope in both research and life.

Thanks to my co-supervisor, Youming Qiao, my research interests have been developed and extended. I would like to thank him for his inspirational discussions and conversations from which I really benefit a lot. I have learned about many interesting topics from him, especially for introducing me to the beauty of computational complexity theory. Meanwhile, I also want to extend my great indebtedness to my external supervisor, Andreas Winter, at Universitat Autónoma de Barcelona. Professor Winter has been my academic hero since I started my PhD, and I always believe that he is the kind of scientist I would like to become in the future.

I would also like to thank Ashely Montanaro and an anonymous examiner for their careful reading of my thesis and constructive comments and suggestions. I am very fortunate to have been part of the Centre for Quantum Software and Information (and the Centre for Quantum Computation and Intelligence Systems from 2014 to 2016). The centre has offered me an example of academic excellence for which to strive.

I want to thank my co-authors Runyao Duan, Cheng Guo, Chi-Kwong Li, Youming Qiao, Xin Wang for giving me the opportunity to work with them. I wish to acknowledge Mingsheng Ying, Yuan Feng, Zhengfeng Ji, Michael Bremner, Ching-Yi Lai, Nengkun Yu, Chris Ferrie, Kun Fang, Wei Xie, Hao-Chung Cheng, Ryan Mann, Mario Berta, Andreas

Winter, Li Gao, Marius Junge, David Gross, Andrew Childs, Ronald de Wolf, Xie Chen, Thomas Vidick, Fernando Brandao, Chris Umans, Micheal Walter, Xiaoming Sun and Man-Hong Yung for their help and discussions in the past four years. I also want to thank Andrew Childs, Xie Chen and Man-Hong Yung for their hospitality during my visit in University of Maryland, California Institute of Technology and South University of Science and Technology of China, and offering me the opportunities to introduce my work in their groups.

I have been blessed with great friends who have made the years spent in Sydney full of fun and adventures. My heartfelt thanks also goes to my family. My love and gratitude for my parents and my grandparents can hardly be expressed in words. Thanks for the endless love and support over all these years. It has been a great privilege to grow in such a loving family and they are and will always be my greatest strength and comfort. Last but by no means least, I would like to thank my beloved fiancè, Zhuoling Tian. Thanks for being part of my life and light it up. We missed in Wuhan University, and we also missed in Europe and USA. But we finally meet down under in Australia, both doing our PhD. I really feel lucky to have someone who is so supportive and understandable, sharing happiness and sorrow together, and at the same time being the greatest comrade in academia, encouraging and inspiring each other and exploring the mysteries of science and arts hand in hand. The most romantic thing that conjures up in my mind is waking up with her smile aside.

Contents

List of Figures

List of Publications

[1] Runyao Duan, Cheng Guo, Chi-Kwong Li, and Yinan Li. Parallel distinguishability of quantum operations. In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 2259 - 2263, July 2016.

[2] Yinan Li, Xin Wang, and Runyao Duan. Indistinguishability of bipartite states by positive-partial-transpose operations in the many-copy scenario. *Phys. Rev. A*, 95:052346, May 2017.

[3] Yinan Li and Youming Qiao. On rank-critical matrix spaces. *Differential Geometry and its Applications*, 55:68 - 77, 2017.

[4] Yinan Li and Youming Qiao. Linear algebraic analogues of the graph isomorphism problem and the Erdős-Rènyi model. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 463-474, Oct 2017.

[5] Yinan Li, Youming Qiao, Xin Wang, and Runyao Duan. Tripartite-to-bipartite entanglement transformation by stochastic local operations and classical communication and the structure of matrix spaces. *Communications in Mathematical Physics*, 358(2):791814, Mar 2018.

Chapter 1

Introduction

Quantum mechanics, in addition to serving as the physical law to govern the universe, has been gradually revolutionizing our understanding of computation and information processing. Quantum algorithms and protocols which utilize the puzzling features of quantum mechanics, such as *quantum superposition* and *quantum entanglement*, are devised to deal with classically hard problems, such as integer factorization and key distribution. On the other hand, the intrinsic *non-commutative* feature of quantum mechanics unveils a richer structure in the manipulation of quantum information. The impact of quantum mechanics has infiltrated many aspects of computer and information science, including the rapid development of *quantum cryptography*, *quantum complexity theory* and *quantum Shannon theory*.

However, our knowledge of quantum computation and information is still limited. Even studies of fundamental tasks, such as quantum state discrimination and entanglement transformation, have not advanced due to complicated mathematical structures. New mathematical theories need to be discovered or adopted in order to characterize complicated quantum processes. In this thesis, we exploit the theory of *matrix spaces*, which studies the algebraic and geometric properties of the linear spaces of matrices, to explore the feasibility and limitations in *quantum state/channel discrimination* and *entanglement*

transformation. Such a theory not only provides powerful mathematical tools to characterize the distinguishability of quantum states and channels and the convertibility of entanglement, it also builds up new connections between quantum information and invariant theory and complexity theory. Moreover, we utilize the emerging "linear algebraic analog" viewpoint to attack the *alternating matrix space isometry problem*, which is equivalent to the long-believed bottleneck case of the group isomorphism problem. In the following four chapters, we present our contributions in detail. Each of these chapters is self-contained and can be read separately.

We start with the study of the *PPT-distinguishability of bipartite quantum states in the many-copy scenario*, which is to distinguish the bipartite quantum states chosen from a set of known ones using operations which completely preserve the positivity of partial transpose. The many-copy scenario stands for the case that arbitrarily many but finite copies of the unknown states are provided. Such a problem plays a central role in the context of *quantum nonlocality*, as sets of PPT-indistinguishable orthogonal bipartite states exhibit the *strongest* form of nonlocality manifested in the setting of LOCC discrimination [Ban11].

A similar problem is the *distinguishability of quantum channels*, which is to identify the unknown quantum channel which is chosen from a set of known channels. Determining the perfect distinguishability of quantum channels has been shown more sophisticated, as we have more freedom in choosing the input states and discrimination schemes (e.g. parallel schemes or sequential schemes). The first characterization for the perfect distinguishability of quantum channels is by Duan, Feng and Ying [DFY09], who presented discrimination protocols using the unknown quantum channels as well as an additional quantum channel, determined by the set of known quantum states. However, as a natural generalization from the quantum state discrimination problem, the distinguishability of quantum channels by the parallel scheme, which is to use the unknown channel alone within finite blocklength, remains unknown.

The aforementioned two problems can be converted into the formalism of *extendibility problems*. Given a subspace S of some Hilbert space $\mathcal{H}$, the goal of an extendibility problem is to determine whether there exists a finite integer k, such that the orthogonal

complement of $S^{\otimes k}$ admits some particular elements or properties. Extendibility problems has been used to characterize many theoretical problems in quantum information theory. We show that the two discrimination problems can be formulated as extendibility problem with respect to matrix spaces, i.e. a subspace of the whole matrix space. In Chapter 2 and Chapter 3, with such formalisms, we derive necessary conditions for the PPT-distinguishability of orthogonal bipartite states and the parallel-distinguishability of quantum channels. As byproducts, we also reprove and extend several well-known results in the literature. We turn to the study of transforming a tripartite pure state to a bipartite pure state via stochastic local operations and classical communication (SLOCC) in Chapter 4. Chitambar, Duan and Shi [CDS10] proved that deciding the feasibility of tripartite-to-bipartite SLOCC entanglement transformation is equivalent to computing the *maximal rank* of the matrix space, determined by the given tripartite pure state. The maximal rank of a matrix space is defined as the largest rank of matrices within. We exhibit novel results in both finite-copy and asymptotic settings, which lead to an algorithmically effective characterization of those tripartite pure states which can be transformed to the bipartite maximally entangled state by SLOCC, in the asymptotic setting.

It is worth noting that the problem of computing the maximal rank has been studied for over 50 years in the context of computational complexity theory, under the name of *Edmonds' problem* [Edm67]. Since 2003, the study of Edmonds' problem has been to investigate the implications of *circuit lower bounds* [KI04]. On the other hand, recent progress on its "non-commutative" version reveals a close connection with *invariant theory*. In particular, an important idea was raised in [IKQS15], which postulates that *matrix spaces can be viewed and studied as a linear algebraic analog of bipartite graphs*. Such a "*linear algebraic analog*" viewpoint can be understood by viewing vectors as a linear algebraic analog of vertices, and matrices as a linear algebraic analog of edges. This analog opens up the possibilities of transforming combinatorial techniques for graph-theoretical problems to the study of algebraic properties of matrix spaces and has been shown to be extremely useful in quantum information theory. An exemplification is the concept of *non-commutative graphs* of quantum channels, which is generalized by Duan, Severini and Winter [DSW13] from the confusability graphs of classical channels, based also on the linear algebraic analog viewpoint.

In Chapter 5, we refine the linear algebraic idea to study the *alternating matrix space isometry problem* (AltMatSpIso) as a linear algebraic analog of the *graph isomorphism problem* (GraphIso). The AltMatSpIso problem asks to decide whether two given alternating matrix spaces are the same up to a basis-change transformations. [1] The motivation of studying AltMatSpIso is to tackle the classic hard isomorphism testing problem, testing isomorphism of p-groups of class 2 and exponent p in time polynomial in the group order, which is the believed bottleneck case of *group isomorphism problem* (GroupIso). To obtain a better understanding on the "linear algebraic analog" viewpoint, we present a comparison between GraphIso and AltMatSpIso in Figure 1.1. In particular, we devise

	GraphIso	AltMatSpIso
Objects	n-vertex graphs with m edges	dimension m $n \times n$ alternating space over $\mathbb{F}_q$
Symmetry	Symmetric group	General linear group
Worst-case Complexity	Quasipolynomial [Bab16a, Bab16b]	$q^{n^2} \cdot \mathrm{poly}(n, m, \log q)$ [FN70, Mil78]
Average-case Complexity	Linear time [BES80]	$q^{O(n+m)}$ [LQ17a]
Random Model	Erdős-Rényi model [ER59]	Linear algebraic Erdős-Rényi model [LQ17a]
Practical	Nauty & Traces [MP14]	Magma & Gap
Group-Theoretic Technique	Permutation group algorithm	Matrix group algorithm
Combinatorial Technique	Individualization and refinement	Linear algebraic analog of Individualization and refinement

Figure 1.1: A systematic comparison between GraphIso and AltMatSpIso.

the first average-case efficient algorithm for AltMatSpIso over a key range of parameters in a random model of alternating matrix spaces in vein of the Erdős-Rényi model of random graphs. Notably, our algorithm is inspired by the first *average-case efficient* algorithm for graph isomorphism, presented by Babai, Erdős and Selkow in the 1970s [BES80]. To devise such an algorithm, we develop linear algebraic analog of *individualization and refinement techniques*, which are crucial in the development of the worst-case complexity of GraphIso.

[1] A matrix $A \in M(n, \mathbb{F})$ is alternating, if for any $v \in \mathbb{F}^n$, $v^{\mathrm{t}} A v = 0$. An alternating matrix space is a linear space of alternating matrices. We point out that the *general* matrix space isometry problem is at least as hard as AltMatSpIso. We will not discuss it in this thesis, despite that it is an interesting algorithmic problem. One of the reason is that we have not observed any connections with other isomorphism problems. In a follow-up work [LQ18], we reduce other isomorphism problems, including *code equivalence* and *lattice isomorphism*, into problems with respect to matrix spaces.

In Chapter 6, we provide a brief summary of our results presented in Chapter 2, 3, 4 and 5. We highlight some interesting and important results, which indicates the power and importance of the theory of matrix spaces. Besides, we discuss more on the "linear algebraic analog" viewpoint, and provide several promising directions for further study.

1.1 Quantum Mechanics in a Nutshell

This section briefly introduces some basic quantum mechanical notions which shall be employed throughout the remainder of this thesis. It is written from a more mathematical, rather than physical, perspective. The contents are based on the postulates of quantum mechanics. We assume the readers are familiar with the Dirac notations and basic linear algebra. For a complete introduction to quantum information and computation, we refer the readers to the excellent textbook by Nielsen and Chuang [NC11], or more comprehensive ones by Wilde [Wil13] and Watrous [Wat18]. More advanced concepts will be introduced later, when necessary.

We use $\mathbb{R}$ ($\mathbb{R}^+$), $\mathbb{N}$ and $\mathbb{C}$ to denote the set of (positive) real numbers, positive integers and complex numbers. Quantum systems are completely characterized by the *quantum states* in the *state space*. The state spaces are *Hilbert spaces* with finite dimension[2], denoted as $\mathcal{H}_n$ where n indicates its dimension. A *pure quantum state* is a unit vector $|\psi\rangle$ in the underlaying *Hilbert space* $\mathcal{H}_n$. The inner product of two quantum states $|\psi\rangle, |\phi\rangle$ is denoted by $\langle\psi|\phi\rangle$. We use $\{|0\rangle, \ldots, |d-1\rangle\}$ to denote the *computational basis* of $\mathcal{H}_n$, where $|k\rangle$ also stands for the vector where the $(k+1)$th entry equals to 1 and the rest are 0. The state space of M composition quantum systems $\mathcal{H}_{n_0}, \ \ldots, \ \mathcal{H}_{n_{M-1}}$ is given by the *tensor product* of each individual state space, denoted by $\mathcal{H}_{n_0} \otimes \cdots \otimes \mathcal{H}_{n_{M-1}}$. Quantum states in composition quantum systems are normally called *multipartite quantum states*. For instance, we use $|\Psi\rangle_{\mathrm{ABC}} \in \mathcal{H}_{n_A} \otimes \mathcal{H}_{n_B} \otimes \mathcal{H}_{n_C}$ to denote a tripartite pure state shared by Alice Bob and Charlie. A multipartite pure state $|\Phi\rangle \in \mathcal{H}_{n_0} \otimes \cdots \otimes \mathcal{H}_{n_{M-1}}$ is called a *product state*, if it is of the form $|\varphi_0\rangle \otimes \cdots \otimes |\varphi_{M-1}\rangle$, where $|\varphi_j\rangle \in \mathcal{H}_{n_j}$ for $j = 0, \ldots, M-1$; otherwise we call $|\Phi\rangle$ an *entangled state*. In the bipartite case, any pure state $|\psi\rangle_{\mathrm{AB}} \in \mathcal{H}_m \otimes \mathcal{H}_n$ admits the *Schmidt decomposition*, i.e. $|\psi\rangle_{\mathrm{AB}} = \sum_{i=0}^{r-1} \sqrt{\lambda_i} \, |\alpha_i\rangle_{\mathrm{A}} \, |\beta_i\rangle_{\mathrm{B}}$, where

[2]Throughout this thesis we focus on finite dimensional spaces only.

$r = \min\{m, n\}$, $0 \leq \lambda_0, \ldots, \lambda_{r-1} \leq 1$ satisfying $\sum_{i=0}^{r-1} \lambda_i = 1$, and $\{|\alpha_0\rangle_{\mathrm{A}}, \ldots, |\alpha_{r-1}\rangle_{\mathrm{A}}\}$ and $\{|\beta_0\rangle_{\mathrm{B}}, \ldots, |\beta_{r-1}\rangle_{\mathrm{B}}\}$ from complete basis of $\mathcal{H}_m$ and $\mathcal{H}_n$, respectively. $\lambda_0, \ldots, \lambda_{r-1}$ are called the Schmidt coefficients, and the number of nonzero Schmidt coefficients of $|\psi\rangle_{\mathrm{AB}}$ is called the *Schmidt rank* of $|\psi\rangle_{\mathrm{AB}}$, denoted by $\mathrm{Sch}(\psi_{\mathrm{AB}})$. It is easy to see that $|\psi\rangle_{\mathrm{AB}}$ is entangled if and only if $\mathrm{Sch}(\psi) \geq 2$.

Let $\mathcal{L}(\mathcal{H}_n, \mathcal{H}_m)$ denote the set of linear operators from $\mathcal{H}_n$ to $\mathcal{H}_m$ and use $\mathcal{L}(\mathcal{H}_n)$ for the abbreviation of $\mathcal{L}(\mathcal{H}_n, \mathcal{H}_n)$. An operator $\rho \in \mathcal{L}(\mathcal{H}_n)$ is positive (semi)definite if for any nonzero $|\psi\rangle \in \mathcal{H}_n$, $\langle\psi|\,\rho\,|\psi\rangle > 0$ ($\langle\psi|\,\rho\,|\psi\rangle \geq 0$). We use $\rho > 0$ ($\rho \geq 0$) to represent that ρ is positive definite (positive semidefinite). The trace of ρ is defined as $\mathrm{Tr}(\rho) := \sum_{i=0}^{n-1} \langle i|\,\rho\,|i\rangle$. We say ρ is a density operator if it is *positive semidefinite with trace* 1. A *mixed quantum state* of a quantum system is described by a *density operator* ρ on the state space $\mathcal{H}_n$. We use $\mathcal{D}(\mathcal{H}_n)$ to denote the set of density operators over $\mathcal{H}_n$. Specifically, the pure state $|\psi\rangle$ corresponding to the density operator $|\psi\rangle\langle\psi|$, which is the outer product of itself. The *spectral decomposition* of $\rho \in \mathcal{L}(\mathcal{H}_n)$ is given by $\rho := \sum_{i=0}^{n-1} \lambda_i\,|\varphi_i\rangle\langle\varphi_i|$, where λ_i is the eigenvalue of ρ and $|\varphi_i\rangle$ is the corresponding eigenvector. The *support* of ρ is then defined as the vector space spanned by those eigenvectors with positive eigenvalues, denoted by $\mathrm{supp}(\rho) = \mathrm{span}\{|\varphi_i\rangle : \lambda_i > 0\}$.

The *closed evolutions* of quantum systems are characterized by the *unitary operators*, which correspond to the solution of Schrödinger equation [Sch26], as it describes the changes over time of the quantum system. $U \in \mathcal{L}(\mathcal{H}_n)$ is unitary if and only if $U^\dagger U = \mathrm{I}_n$[3], where $U^\dagger$ denotes the conjugate transpose of U. The actions of a unitary operator on a pure state $|\psi\rangle$ and a mixed state ρ are given by $U\,|\psi\rangle$ and $U\rho U^\dagger$, respectively. The *dynamics of open quantum systems* are characterized by *quantum operations*, or equivalently, *quantum channels*. From a mathematical perspective, a quantum channel is a *completely positive and trace-preserving* (CPTP) linear map from the input space $\mathcal{L}(\mathcal{H}_n)$ to the output space $\mathcal{L}(\mathcal{H}_m)$. A linear map $\mathcal{E} : \mathcal{L}(\mathcal{H}_n) \to \mathcal{L}(\mathcal{H}_m)$ is *completely positive* if for every positive integer k, $\mathcal{I} \otimes \mathcal{E} : \mathcal{L}(\mathcal{H}_k \otimes \mathcal{H}_n) \to \mathcal{L}(\mathcal{H}_k \otimes \mathcal{H}_m)$ maps all positive semidefinite operators in $\mathcal{L}(\mathcal{H}_k \otimes \mathcal{H}_n)$ to semidefinite ones in $\mathcal{L}(\mathcal{H}_k \otimes \mathcal{H}_m)$, where $\mathcal{I}$ denotes the identity map. A linear map $\mathcal{E} : \mathcal{L}(\mathcal{H}_n) \to \mathcal{L}(\mathcal{H}_m)$ is *trace-preserving*, if $\mathrm{Tr}(\rho) = \mathrm{Tr}(\mathcal{E}(\rho))$ for all $\rho \in \mathcal{L}(\mathcal{H}_n)$.

[3] I_n denotes the identity operator on $\mathcal{H}_n$. The subscript n may be abbreviated when there is no confusion.

Each CPTP map can be represented by a set of *Choi-Kraus operators*, i.e. there exist $\{E_i \in \mathcal{L}(\mathcal{H}_n, \mathcal{H}_m) : i = 1, \ldots, k\}$, such that $\mathcal{E}(\rho) = \sum_{i=1}^{k} E_i \rho E_i^\dagger$ [Cho75, Kra83].

The way to observe the quantum system is through *quantum measurement*. Mathematically speaking, a k-outcome quantum measurement $\mathcal{M}$ is identified by a collection of *measurement operators* $\{M_i \in \mathcal{L}(\mathcal{H}_n, \mathcal{H}_m) : i = 0, \ldots, k-1\}$ satisfying $\sum_{i=0}^{k-1} M_i^\dagger M_i = \mathrm{I}_n$. The index i refers to the measurement outcome. When measuring a quantum system in state $|\psi\rangle$ with $\mathcal{M}$, the probability that the outcome is k is $\langle\psi| M_i^\dagger M_i |\psi\rangle$. The state after the measurement becomes $\frac{M_i|\psi\rangle}{\langle\psi|M_i^\dagger M_i|\psi\rangle}$. When we only focus on the classical outcomes, it is convenient to analyze the measurement via *Positive Operator-Valued Measure* (POVM) *formalism*. A k-outcome POVM is identified by $\{P_i \in \mathcal{L}(\mathcal{H}_n, \mathcal{H}_m) : i = 0, \ldots, k-1\}$, where P_i is semidefinite for $i = 0, \ldots, k-1$ and $\sum_{i=0}^{k-1} P_i = \mathrm{I}$. The probability that the outcome is i is $\langle\psi| P_i |\psi\rangle$.

1.2 Matrix Spaces in a Nutshell

A matrix space is a linear subspace of the linear space of all $n \times n$ matrices. In algorithms, a matrix space is given by a tuple of matrices. As the central concept in this thesis, matrix spaces has been studied in different area of mathematics for decades. In the following, we describe some useful ideas to study matrix spaces.

Let $\mathbb{F}$ be a field. $M(n, \mathbb{F})$ stands for the space of all $n \times n$ matrices over $\mathbb{F}$. A dimension m matrix space $\mathcal{B}$ of $M(n, \mathbb{F})$ is the linear space spanned by m linearly independent matrices. As a *linear space*, the matrix space $\mathcal{B}$ is isomorphic to a *vector space* $V_\mathcal{B}$ of $\mathbb{F}^{n^2}$, by mapping each matrix $B \in \mathcal{B}$ to a "longer vector" v_B, whose entries are matrix elements of B. We can also induce a *inner product* structure on $M(n, \mathbb{F})$ based on this connection. Namely, the inner product of two matrices B_1 and B_2 is defined as the inner product of v_{B_1} and v_{B_2}. This property will be utilized in Chapter 2 and 4, to study the discriminations of quantum states and quantum operations.

On the other hand, matrix spaces inherit linear algebraic properties. One of the natural generalization would be the concept of *rank*: the (maximal) rank of a matrix space is defined as the largest rank of matrices therein. To better understand this concept, it is

convenient to view matrix space as a *symbolic matrix*, whose entries are linear functions in variables $\{x_1, \ldots, x_m\}$ over the field $\mathbb{F}$. Take a linear basis of $\mathcal{B}$, say $B_1, \ldots, B_m$, the matrix $x_1B_1 + \cdots + x_mB_m$ is a symbolic matrix, and we shall use the same symbol $\mathcal{B}$ to represent it. When the variables are *commutative*, determine the rank of the symbolic matrix $\mathcal{B}$ over *rational function field* is equivalent to determine the maximal rank of the matrix space $\mathcal{B}$. More details of the algebraic properties of matrix space can be found in Chapter 5.

Finally, we point out that matrix spaces can be also studied as 3-*tensors*, by listing the tuple of (linear basis) matrices (e.g. Figure 1.2). Although two 3-tensors may represent the same matrix spaces, the representation is more intuitive and illustrative when we view matrix spaces as a linear algebraic analog of graphs, in particular, when transferring graph-theoretical techniques into the linear algebraic setting. In Chapter 6, deriving the linear algebraic analog of the *individualization and refinement* technique benefits a lot from such a viewpoint.

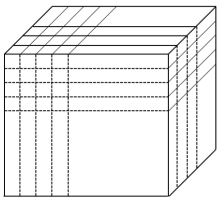


FIGURE 1.2: A 3-tensor.

Chapter 2

PPT-distinguishability of Orthogonal Bipartite States

In this chapter, we investigate the PPT-distinguishability of orthogonal bipartite states in the many-copy scenario, i.e. arbitrarily many but finite copies (uses) of the unknown state are provided. In particular, we reduce the discrimination problem into the formalism of extendibility problem, which helps us to devise an efficient computable and necessary condition to decide the PPT-distinguishability. This chapter is organized as follows: We first review the local distinguishability of orthogonal quantum states in Section 2.1. After providing essential notations and preliminaries (Section 2.2), we discuss our main results in Section 2.3, based on [LWD17]. We close in Section 2.4 with a brief summary.

2.1 The Local Distinguishability of Orthogonal Quantum States

The discrimination of quantum states is the task of identifying the unknown quantum state of a quantum system from a prior-known set of possibilities. Such a task is one of the pillars of quantum information theory. When we can access the whole quantum system, quantum mechanics ensures that only orthogonal quantum states can be distinguished perfectly, i.e. with probability 1. However, in most situations, we only have local access to a composite quantum systems, hence the distinguishability can be completely different. Generally speaking, we are given a multipartite quantum system which is distributed to spatially separated parties. Each party is restricted to act locally on its respective subsystem by performing arbitrary quantum operations. In addition, they can communicate classical data with each other. To achieve the perfect discrimination, they need to propose a joint protocol in such a manner. These protocols are normally referred to *local operations and classical communication* (LOCC). When there are only two possible multipartite pure states to be distinguished by LOCC, Walgate *et al.* [WSHV00] proved that LOCC discrimination protocols always exist. However, when the number of possibilities increases, local discriminations may be infeasible. For instance, the 4 $2 \otimes 2$ Bell states are not locally distinguishable [GKR$^+$01].

Essentially, the discrepancy of the global and local distinguishability of multipartite orthogonal states reveals *quantum nonlocality*, as there is a gap between the local information which we have access to and the global information stored in the composite system. In the setting of local discrimination, it can be understood as the identity of the unknown state can be only obtained by applying global operations to the composite system instead of having local access only. Thus, the investigation of the local discrimination of orthogonal multipartite states provides a natural strategy to study quantum nonlocality despite the fact that most of the celebrated manifestations of quantum nonlocality arise from entangled states. Bennett *et al.* [BDF$^+$99] observed the strange phenomenon referred to as *nonlocality without entanglement*, which revealed that entanglement is not always necessary for exhibiting quantum nonlocality. They presented a set of 9 orthogonal bipartite *product* pure states which cannot be distinguished perfectly by LOCC. Since

then, the local discrimination of orthogonal quantum states has been extendedly studied in [BDM$^+$99, WH02, DMS$^+$03, HSSH03, Wat05, HMM$^+$06, DFJY07, FS09] (and references therein).

However, the aforementioned local indistinguishability can be overcome if we increase the number of available copies of the unknown state. For instance, we have previously discussed that any $2\otimes 2$ Bell state cannot be distinguished from the rest by LOCC [Fan04, GKRS04]. Ghosh *et al.* [GKRS04] pointed out that two copies of the unknown state are sufficient for local distinguishability. Remarkably, Bandyopadhyay asserted that any N orthogonal multipartite pure states are local distinguishable with at most $N-1$ copies [Ban11]. Consequently, if we remove the restriction to have only single access to the unknown pure state, the nonlocality exhibited in distinguishing multipartite pure states will no longer exist. On the other hand, when at least one of the possible states is mixed, we may not distinguish these states perfectly, even if arbitrarily many but finite copies are provided. For instance, it is impossible to distinguish an arbitrary bipartite maximally entangled state with the normalized projection onto its orthogonal complement by LOCC, even given arbitrarily many but finite many copies [YDY14]. In some sense, th nonlocality presented in local discrimination is more robust in mixed states for it persists even in the domain of multiple copies, whereas in case of pure states it does not.

Moreover, we can further ask that what is the *strongest* form of nonlocality manifested in the setting of local discrimination? To answer such a question, note that we are comparing the amount of information we can obtain using LOCC and global operations, which can be interpreted as the probability of successful identifying the unknown states. We focus on the *unambiguous local discrimination* [DFJY07, BW09]. Protocols for unambiguous local discrimination determine the identity of the unknown states perfectly with some nonzero probability. Namely, we say a set of quantum states is *locally indistinguishable in the many-copy scenario* if they cannot be distinguished by LOCC unambiguously, when arbitrarily many but finite copies of the unknown state are given. Clearly, these sets of quantum states can be viewed as the strongest form of nonlocality manifested in the setting of local discrimination. It is of great interest to consider the construction and verification of such sets of states. In [Ban11], Bandyopadhyay provided a way to construct

locally indistinguishable orthogonal quantum states in the many-copy scenario using the *unextendible product basis* (UPB).

In general, the local distinguishability is hard to characterize even in the single-copy setting, as the structure of LOCC operations are rather sophisticated. A common method to remedy this obstacle is to study distinguishability by quantum operations which completely preserve the positivity of partial transpose, usually abbreviated by PPT-distinguishability or distinguishability by PPT operations. The family of PPT operations admits a neat mathematical characterization and contains all possible LOCC operations as a subset. Therefore, if a set of orthogonal states cannot be distinguished by PPT operations, it is also locally indistinguishable. Meanwhile, the notion of PPT is crucial in many areas of quantum information theory. It was originally developed by Peres [Per96] and Horodecki *et al.* [HHH96] for the separability tests. Later on, the notion of PPT was introduced to the study of entanglement transformation [Ish04], entanglement distillation [Rai01], and the classical and quantum capacity of quantum channels [LM15, WXD17].

In Section 2.3, we investigate the PPT-indistinguishability of orthogonal bipartite quantum states in the many-copy scenario. We first reduce the problem of determining PPT-distinguishability of orthogonal bipartite states with k copies to decide whether a bipartite subspace S is *k-PPT-extendible*, that is, if there exist nonzero PPT operators of which the support is orthogonal to $S^{\otimes k}$. Note that the k-PPT-extendible spaces are naturally generalized from the *k-extendible subspaces*, introduced by Cubitt, Chen and Harrow [CCH11]. With such a formalism, we derive a simple criterion to decide whether a given bipartite space is not k-PPT-unextendible for arbitrary $k \in \mathbb{N}$. This criterion then can be utilized to decide the PPT-indistinguishability in the many-copy scenario. Moreover, we utilize this criterion to show that an arbitrary *bipartite entangled state* and its orthogonal complement are PPT-indistinguishable in the many-copy scenario, which reproves and extends the results in [YDY14]. On the other hand, our criterion can be applied to study the minimum dimension of PPT-unextendible subspaces in $\mathcal{H}_m \otimes \mathcal{H}_n$. Johnston [Joh13] explicitly constructs a dimension-$(m-1)(n-1)$ subspace S in $\mathcal{H}_m \otimes \mathcal{H}_n$ which is 1-PPT-unextendible, matching the bounds for unextendible subspaces [CMW08]. We extend Johnston's result by proving the subspace he constructed is k-PPT-unextendible for any $k \in \mathbb{N}$.

2.2 Notations and Preliminaries

We focus on distinguishing orthogonal bipartite quantum states. We say two quantum states $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H}_n)$ are orthogonal, denoted by $\rho_1 \perp \rho_2$, if and only if $\mathrm{Tr}(\rho_1^\dagger \rho_2) = 0$. We use $\leq$ to denote the subspace relation, i.e. $S' \leq S$ if and only if S is a subspace of S'. We say a density operator ρ is support on S, if $\mathrm{supp}(\rho) \leq S$. For a subspace $S \leq \mathcal{H}_n$, the orthogonal complement of S is denoted and defined as $S^\perp := \{|\psi\rangle \in \mathcal{H}_n | \langle\psi|\varphi\rangle = 0,\ \forall |\varphi\rangle \in S\}$. The projection onto the subspace S is denoted and defined by $P_S := |\psi_0\rangle\langle\psi_0| + \cdots + |\psi_{N-1}\rangle\langle\psi_{N-1}|$, where $\{|\psi_0\rangle, \ldots, |\psi_{N-1}\rangle\}$ is a set of orthonormal basis of S. The normalized projection onto S is defined as $\rho_S = \frac{P_S}{\mathrm{Tr}(P_S)}$. We use $\{|i\rangle\langle j| : i = 0, \ldots, n-1,\ j = 0, \ldots, m-1\}$ to denote the computational basis of $\mathcal{L}(\mathcal{H}_m, \mathcal{H}_n)$. The partial transpose operation (with respect to the second system) $\Gamma : \mathcal{L}(\mathcal{H}_m \otimes \mathcal{H}_n) \to \mathcal{L}(\mathcal{H}_m \otimes \mathcal{H}_n)$ maps $|i\rangle\langle j| \otimes |k\rangle\langle l|$ to $|i\rangle\langle j| \otimes |l\rangle\langle k|$, and we simply use ρ^Γ to denote the operator after taking partial transpose. We say $\rho \in \mathcal{D}(\mathcal{H}_m \otimes \mathcal{H}_n)$ is a PPT state, if ρ^Γ is positive semidefinite. In addition, w say $\rho \in \mathcal{D}(\mathcal{H}_m \otimes \mathcal{H}_n)$ is a PPT-definite state, if ρ^Γ is positive definite.

For the definition of (unambiguous) PPT distinguishability of orthogonal bipartite quantum states, we simply adapt the definitions in [YDY14].

Definition 2.1. Promise that a bipartite quantum system is prepared in one of the possible states in $\{\rho_i \in \mathcal{D}(\mathcal{H}_m \otimes \mathcal{H}_n) : \rho_j \perp \rho_k,\ \forall\ ^{\prime\prime} j, k = 0, \ldots, N-1\}$. Given one copy of the unknown state, we say a set of orthogonal bipartite states $\{\rho_0, \ldots, \rho_{N-1}\}$ is

- **unambiguous local distinguishable**, if there exists a set of POVM operators $\{P_1, \ldots, P_N\}$ which can be implemented by a LOCC protocol (LOCC POVM), such that $\mathrm{Tr}(\mathrm{P}_i \rho_j) = p_j \delta_{ij}$ for all $i, j = 0, \ldots, N-1$, where $\delta_{ij} = 0$ if and only if $i \neq j$ and $\delta_{ii} = 1$ for all possible i.
- **unambiguous PPT-distinguishable**, if there exists a set of POVM operators $\{P_0, \ldots, P_{N-1}\}$ where P_i^Γ is positive semidefinite for $i = 0, \ldots, N-1$ (PPT POVM), such that $\mathrm{Tr}(\mathrm{P}_i \rho_j) = p_j \delta_{ij}$ for all $i, j = 0, \ldots, N-1$. ($\delta_{ij} = 0$ if and only if $i \neq j$ and $\delta_{ij} = 1$ if and only if $i = j$.)
- **local (PPT) indistinguishable**, if it is not unambiguous local (PPT) indistinguishable.

Since any LOCC protocols belongs to the set of all PPT operations, $\{\rho_0, \ldots, \rho_{N-1}\}$ is unambiguous local distinguishable implies $\{\rho_0, \ldots, \rho_{N-1}\}$ is unambiguous PPT-distinguishable. Thus If we can show a set of orthogonal quantum states is PPT-indistinguishable, they must be locally indistinguishable. Note that the definition can be generalized to the situation when multiple copies are provided. Thus the PPT-indistinguishability in the many-copy scenario implies local indistinguishability in the many-copy scenario.

We say a subspace $S \leq \mathcal{H}_m \otimes \mathcal{H}_n$ is PPT-extendible, if there exist nonzero PPT operators $\rho \in \mathcal{D}(\mathcal{H}_m \otimes \mathcal{H}_n)$, such that $\mathrm{supp}(\rho) \leq S^{\perp}$. In other words, there exist nonzero PPT operators of which the support contains in the orthogonal complement of S. We say S is k-PPT-extendible if $S^{\otimes k}$ is PPT-extendible. We say S is strongly PPT-unextendible if S is not k-PPT-extendible for any $k \in \mathbb{N}$. Note that, $S \leq \mathcal{H}_m \otimes \mathcal{H}_n$ is PPT-indistinguishable implies $S^{\perp}$ contains only entangled states, as any product states have positive partial-transpose. Wallach [Wal02] and Parthasarathy [Par04] proved that the maximum dimension of subspaces in $\mathcal{H}_m \otimes \mathcal{H}_n$ which contain only entangled state is $(m-1)(n-1)$. This implies the minimum dimension of PPT-unextendible subspaces is at least $m+n-1$. Johnston [Joh13] construct the following subspace $S_{m,n}$, of which the dimension is $m+n-1$ and is PPT-unextendible [1]:

$$\begin{aligned} S_{m,n} = \mathrm{span}\{|\psi_s\rangle = \sum_{j=0}^{m-1-s} |j\rangle\,|m-1-s-j\rangle : \ s = 0, \ldots, m-1; \\ |\varphi_t\rangle = \sum_{j=t-m+1}^{\min\{n-1,t\}} |t-j\rangle\,|j\rangle : \ t = m, \ldots, m+n-2\}. \end{aligned} \tag{2.1}$$

Eventually, we introduce several notions and results from matrix analysis. Given an $n \times n$ matrix H, $H_{I,J}$ denotes the $|I| \times |J|$ submatrix of H with respect to the row index subset $I = \{i_1, \ldots, i_{|I|}\}$ and column index subset $J = \{j_1, \ldots, j_{|J|}\}$. When $I = J$, $H_{I,I}$ is called a principal submatrix and the determinant of $H_{I,I}$ is called a principal minor. For $I = \{1, \ldots, i\}$, $H_{I,I}$ is the ith leading principal submatrix for $i = 1, \ldots, n$. For $J = \{j, \ldots, n\}$, $H_{J,J}$ is the jth trailing principal submatrix for $j = 1, \ldots, n$. The determinants of leading

[1] In [Joh13], the constructed subspace is $S(m,n) = \mathrm{span}\{|j\rangle\,|k+1\rangle - |j+1\rangle\,|k\rangle : 0 \leq j \leq m-2,\ 0 \leq k \leq n-2\}$, which satisfies that any $\rho \in \mathcal{D}(\mathcal{H}_m \otimes \mathcal{H}_n)$ supporting on S has non-positive partial transpose. The subspace $S_{m,n}$ is taken as the orthogonal complement of $S(m,n)$.

(trailing) principal submatrices are called leading (trailing) principal minors. With these notations, we exhibit the following useful results to decide the positive (semi)definiteness:

Theorem 2.2 ([HJ12])**.**

- *Let H be Hermitian, i.e. $H^\dagger = H$. If all the leading (trailing) principal minors of H are positive, H is positive definite. (*Sylvester's criterion*)*
- *For any positive definite matrix A, the determinant and all principal minors are positive.*

2.3 PPT-indistinguishability of Orthogonal Bipartite States

2.3.1 A Sufficient Condition for PPT-indistinguishability in the Many-copy Scenario

We start from deriving a sufficient condition for PPT-indistinguishability of orthogonal bipartite quantum states in the many-copy scenario. We first observe that, if there exists $k \in \{0, \ldots, N-1\}$ such that for any nonzero PPT operator $P \in \mathcal{L}(\mathcal{H}_m \otimes \mathcal{H}_n)$, $\mathrm{Tr}(P\rho_k) \neq 0$, then S cannot be distinguished by PPT POVM by definition 2.1. Based on this observation, we can easily derive the following:

Proposition 2.3. *For a set of orthogonal quantum states $\{\rho_0, \ldots, \rho_{N-1}\}$, if there exists $k \in \{0, \ldots, N-1\}$ such that* $\mathrm{supp}(\rho_k)$ *is strongly PPT-unextendible, then $\{\rho_0, \ldots, \rho_{N-1}\}$ is PPT-indistinguishable in the many-copy scenario.*

This proposition enables us to reduce determining PPT-indistinguishability in the many-copy scenario to determining the strongly PPT-unextendibility. Such problems are referred to the *extendibility problem*. A typical one is to determine whether the orthogonal complement of a given bipartite subspace contains nonzero product states, which plays an important role in the study of super-activation of zero-error classical capacity of quantum channel [CCH11, Dua09]. We can convert the k-PPT-extendibility of

$S = \mathrm{span}\{|\psi_i\rangle \in \mathcal{H}_m \otimes \mathcal{H}_n : i = 0, \ldots, M-1\}$ to determining the feasibility of the following semidefinite constraints:

$$\begin{aligned} &\langle\psi| P |\psi\rangle = 0, \ \forall \ |\psi\rangle \in \{|\psi_0\rangle, \ldots, |\psi_{M-1}\rangle\}^{\otimes k}, \\ &P \in \mathcal{L}((\mathcal{H}_m \otimes \mathcal{H}_n)^{\otimes k}), \ P \geq 0, \ P^{\Gamma} \geq 0, \end{aligned} \tag{2.2}$$

where $\{|\psi_0\rangle, \ldots, |\psi_{M-1}\rangle\}^{\otimes k} = \{|\psi_{j_0}\rangle \otimes \cdots \otimes |\psi_{j_{k-1}}\rangle : j_0, \ldots, j_{k-1} \in \{0, \ldots, M-1\}\}$. Although SDP can be efficiently solved by convex optimization packages such as CVX [GB14] in MATLAB, the dimensions of the above constrains grows exponentially when k increasing, which makes it infeasible when dealing with even $k = 20$. Thus, we require a witness for the strongly PPT-unextendibility, which is efficient computable.

Theorem 2.4. *If there is a PPT-definite operator $P \in \mathcal{L}(\mathcal{H}_m \otimes \mathcal{H}_n)$ supporting on $S \leq \mathcal{H}_m \otimes \mathcal{H}_n$, then S is strongly PPT-unextendible.*

Proof. It is easy to see that P is PPT-definite if and only if $\mathcal{T}(P) > 0$, where $\mathcal{T}(P)$ is given by the following semidefinite program:

$$\mathcal{T}(P) = \max\{t \in \mathbb{R} : \ 0 \leq R \leq P, \ R^{\Gamma} \geq t\mathrm{I}_{mn}\}. \tag{2.3}$$

Utilizing this SDP characterization, we then prove that $P^{\otimes k}$ supports on $S^{\otimes k}$ for any $k \in \mathbb{N}$. To see this, we only need to show that $T(P_1 \otimes P_2) \geq T(P_1) \times T(P_2)$ for any PPT-definite operators P_1 and P_2. Assume that the optimal solutions to SDP (2.3) of $T(P_1)$ and $T(P_2)$ are $\{R_1, t_1\}$ and $\{R_2, t_2\}$, respectively. It is easy to derive that $0 \leq R_1 \otimes R_2 \leq P_1 \otimes P_2$ and $R_1^{\Gamma} \otimes R_2^{\Gamma} \geq t_1 t_2 \mathrm{I}$. Then $\{R_1 \otimes R_2, t_1 t_2\}$ is a feasible solution to SDP (2.3) when replacing P by $P_1 \otimes P_2$, which means that $T(\sigma_1 \otimes \sigma_2) \geq t_1 t_2 > 0$. It follows immediately that $T(P^{\otimes k}) > 0$, or equivalently $P^{\otimes k}$ is PPT-definite, for any $k \in \mathbb{N}$.

Now assume that there exists $k_0 \in \mathbb{N}$ such that $Q \in \mathcal{L}((\mathcal{H}_m \otimes \mathcal{H}_n)^{\otimes k_0})$ is a nonzero PPT operator supporting on the orthogonal complement of $S^{\otimes k_0}$. Since $P^{\otimes k_0}$ supports on $S^{\otimes k_0}$, we know that

$$\mathrm{Tr}(P^{\otimes k} Q) = \mathrm{Tr}((P^{\otimes k})^{\Gamma} Q^{\Gamma}) > 0, \tag{2.4}$$

where the second inequality holds since $P^{\otimes k}$ is PPT-definite, which leads to the contradiction. □

Following Theorem 2.4, we can directly obtain a sufficient condition for PPT-indistinguishability in the many-copy scenario:

Corollary 2.5. *For a set of orthogonal quantum states* $\{\rho_0, \ldots, \rho_{N-1}\}$*, if there exists* $k \in \{0, \ldots, N-1\}$ *such that there is a PPT-definite operator supporting on* $\mathrm{supp}(\rho_k)$*, then* $\{\rho_0, \ldots, \rho_{N-1}\}$ *is PPT-indistinguishable in the many-copy scenario.*

2.3.2 Constructions of PPT-indistinguishable Orthogonal Bipartite States in the Many-copy Scenario

One major method for constructing orthogonal quantum states which are locally indistinguishable in the many-copy scenario is to invoke the *unextendible product basis* (UPB). A set of orthogonal product states $\{|\Psi_0\rangle, \ldots, |\Psi_{N-1}\rangle\} \subseteq \mathcal{H}_m \otimes \mathcal{H}_n$ is a UPB if the orthogonal complement of $S = \mathrm{span}\{|\Psi_0\rangle, \ldots, |\Psi_{N-1}\rangle\}$ contains no product state. Denote the orthogonal complement of S by $S^{\perp}$. Then ρ_S and $\rho_{S^{\perp}}$ cannot be distinguished by LOCC, even given arbitrarily large but finite copies [Ban11]. In the following, we use corollary 2.5 to construct orthogonal local (PPT) indistinguishable states in the many-copy scenario.

Theorem 2.6. *Given a bipartite entangled state* $|\varphi\rangle \in \mathcal{H}_m \otimes \mathcal{H}_n$*, let* $S_\varphi = \{|\psi\rangle \in \mathcal{H}_m \otimes \mathcal{H}_n : \langle\psi|\varphi\rangle = 0\}$*. Then* $|\varphi\rangle\langle\varphi|$ *and* ρ_{S_φ} *are PPT-indistinguishable in the many copy scenario.*

Proof. Piani and Mora [PM07] implicitly proved that ρ_{S_φ} is PPT-definite, and the theorem then follows by corollary 2.5. Here we provide an alternating proof to the PPT-definiteness of ρ_{S_φ}. Let $\lambda_{\min}(P)$ and $\lambda_{\max}(P)$ be the minimum and maximum eigenvalues of the operator P, respectively. Note that $\rho_{S_\varphi} = \frac{1}{mn-1}(\mathrm{I}_{mn} - |\varphi\rangle\langle\varphi|)$. ρ_{S_φ} is PPT-definite if and only if $\lambda_{\min}(\mathrm{I}_{mn} - |\varphi\rangle\langle\varphi|^{\Gamma}) > 0$, or equivalently, $\lambda_{\max}(|\varphi\rangle\langle\varphi|^{\Gamma}) < 1$. Let the Schmidt decomposition of $|\varphi\rangle$ be $|\varphi\rangle = \sum_{i=0}^{r-1} \lambda_i |\alpha_i\rangle |\beta_i\rangle$ with $1 > \lambda_0^2 \geq \cdots \geq \lambda_{r-1}^2 > 0$ and

$\lambda_0^2 + \cdots + \lambda_{r-1}^2 = 1$. Then $|\varphi\rangle\langle\varphi|^\Gamma$ can be represented as

$$\begin{aligned}
|\varphi\rangle\langle\varphi|^\Gamma =& \sum_{i=0}^{r-1} \lambda_i^2 |\alpha_i\rangle\langle\alpha_i| \otimes |\beta_i\rangle\langle\beta_i| + \sum_{i\neq j} \lambda_i\lambda_j |\alpha_i\rangle \langle\alpha_j| \otimes |\beta_j\rangle \langle\beta_i| \\
=& \sum_{i=0}^{r-1} \lambda_i^2 |\alpha_i\rangle |\beta_i\rangle\langle\alpha_i| \langle\beta_i| + \sum_{i>j} \frac{\lambda_i\lambda_j}{2}[(|\alpha_i\rangle |\beta_j\rangle + |\alpha_j\rangle |\beta_i\rangle)(\langle\alpha_i| \langle\beta_j| + \langle\alpha_j| \langle\beta_i|)] \\
& - \sum_{i>j} \frac{\lambda_i\lambda_j}{2}[(|\alpha_i\rangle |\beta_j\rangle - |\alpha_j\rangle |\beta_i\rangle)(\langle\alpha_i| \langle\beta_j| - \langle\alpha_j| \langle\beta_i|)].
\end{aligned} \tag{2.5}$$

Thus, the eigenvalues of $|\varphi\rangle\langle\varphi|^\Gamma$ are λ_i^2 and $\pm\frac{\lambda_i\lambda_j}{2}$. Since $1 > \lambda_0 \geq \cdots \geq \lambda_{r-1} > 0$, $\lambda_{max}(|\varphi\rangle\langle\varphi|^\Gamma) = \lambda_1^2 < 1$, which derives $\mathrm{I}_{n^2} - |\varphi\rangle\langle\varphi|^\Gamma$ is PPT-definite. □

Remark 2.7. Theorem 2.6 immediately implies that any maximally entangled state and the normalized projection onto its orthogonal complement are PPT-indistinguishable in the many-copy scenario [YDY14]. The proof of Theorem 2.6 is more intrinsic, as the proof in [YDY14] highly relies on the symmetry of maximally entangled state.

2.3.3 Minimum Dimension of Strongly PPT-unextendible Spaces in $\mathcal{H}_m \otimes \mathcal{H}_n$

In this subsection we show the subspace $S_{m,n} \leq \mathcal{H}_m \otimes \mathcal{H}_n$ (see Equation (2.1)) constructed by Johnston in [Joh13] is strongly PPT-unextendible, which derives the minimum dimension of strongly PPT-unextendible subspaces in $\mathcal{H}_m \otimes \mathcal{H}_n$ is also $m + n - 1$.

Theorem 2.8. *Let*

$$\begin{aligned}
S_{m,n} = \mathrm{span}\{|\psi_s\rangle = \sum_{j=0}^{m-1-s} |j\rangle |m-1-s-j\rangle :\ s = 0, \ldots, m-1; \\
|\varphi_t\rangle = \sum_{j=t-m+1}^{\min\{n-1,t\}} |t-j\rangle |j\rangle :\ t = m, \ldots, m+n-2\}.
\end{aligned}$$

Then there exist PPT-definite operators supporting on $S_{m,n}$. Moreover, $S_{m,n}$ is strongly PPT-unextendible.

Proof. We shall prove that there exist $x_0, x_1, \ldots, x_{m-1} \in \mathbb{R}^+$ and $y_m, \ldots, y_{m+n-2} \in \mathbb{R}^+$, such that

$$\rho_{m,n}^{\Gamma} = \sum_{s=0}^{m-1} x_{m-1-s} \left|\psi_s\right\rangle\!\left\langle\psi_s\right|^{\Gamma} + \sum_{t=m}^{m+n-2} y_t \left|\varphi_t\right\rangle\!\left\langle\varphi_t\right|^{\Gamma} \tag{2.6}$$

is PPT-definite. For $s = 0, \ldots, m-1$ and $t = m, \ldots, m+n-2$, we have

$$\begin{aligned} \left|\psi_s\right\rangle\!\left\langle\psi_s\right|^{\Gamma} &= \sum_{j_1, j_2=0}^{m-1-s} \left|j_1\right\rangle \left\langle j_2\right| \otimes \left|m-1-s-j_2\right\rangle \left\langle m-1-s-j_1\right|, \\ \left|\varphi_t\right\rangle\!\left\langle\varphi_t\right|^{\Gamma} &= \sum_{j_1, j_2=t-m+1}^{\min\{n-1, t\}} \left|t-j_1\right\rangle \left\langle t-j_2\right| \otimes \left|j_2\right\rangle \left\langle j_1\right|. \end{aligned} \tag{2.7}$$

Represent $\rho_{m,n}^{\Gamma}$ as a matrix with respect to the computational basis. We participate $\{\left|j\right\rangle \left|k\right\rangle :\ j = 0, \ldots, m-1,\ k = 0, \ldots, n-1\}$ into the following families:

$$\begin{aligned} &\mathcal{P}_a = \{\left|m-1-a+k\right\rangle \left|k\right\rangle :\ 0 \le k \le a\},\ a = 0, \ldots, m-1; \\ &\mathcal{Q}_b = \{\left|l\right\rangle \left|l+b\right\rangle :\ 0 \le l \le \min\{n-1-b, m-1\}\},\ b = 1, \ldots, n-1. \end{aligned} \tag{2.8}$$

Let P_a and Q_b denote the submatrices spanned by $\mathcal{P}_a$ and $\mathcal{Q}_b$, respectively, for $a = 0, \ldots, m-1$ and $b = 1, \ldots, n-1$. We have the decomposition $\rho_{m,n}^{\Gamma} = (\oplus_{a=0}^{m-1} P_a) \oplus (\oplus_{b=1}^{n-1} Q_b)$. Thus, we reduce ourselves to show the existence of $x_0, x_1, \ldots, x_{m-1} \in \mathbb{R}^+$, $y_m, \ldots, y_{m+n-2} \in \mathbb{R}^+$, such that

$$P_a = \begin{pmatrix} x_a & x_{a-1} & \cdots & x_0 \\ x_{a-1} & \ddots & \ddots & y_m \\ \ddots & \ddots & \ddots & \vdots \\ x_0 & y_m & \cdots & y_{m+a-1} \end{pmatrix}_{(a+1)\times(a+1)} \qquad 0 \le a \le m-1, \tag{2.9}$$

$$Q_b = \begin{pmatrix} x_{m-1-b} & \cdots & x_0 & y_m & \cdots & y_{m+b-1} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ x_0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ y_m & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ y_{m+b-1} & \cdots & \cdots & \cdots & \cdots & y_{2m+b-2} \end{pmatrix}_{m\times m} \qquad 1 \le b \le n-m, \tag{2.10}$$

$$Q_b = \begin{pmatrix} x_{m-1-b} & \cdots & x_0 & y_m & \cdots & y_{n-1} \\ \vdots & \iddots & \iddots & \iddots & \iddots & \vdots \\ x_0 & \iddots & \iddots & \iddots & \iddots & \vdots \\ y_m & \iddots & \iddots & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & \iddots & \iddots & \vdots \\ y_{n-1} & \cdots & \cdots & \cdots & \cdots & y_{2n-b-2} \end{pmatrix}_{(n-b)\times(n-b)} \qquad n-m < b \leq n-1, \quad (2.11)$$

are positive definite.

We first show that $P_0, \ldots, P_{m-1}$ can be positive definite. Let $x_a = y_{m-1+a}$ for $a = 1, \ldots, m-1$. Then P_a reduce to

$$P_a = \begin{pmatrix} x_a & x_{a-1} & \cdots & x_0 \\ x_{a-1} & \iddots & \iddots & x_1 \\ \vdots & \iddots & \iddots & \vdots \\ x_0 & x_1 & \cdots & x_a \end{pmatrix}_{(a+1)\times(a+1)} . \quad (2.12)$$

When $a = 0$, $P_a = x_0$ and we set x_0 as an arbitrary positive real number x'_0. Assume we have chosen $x'_0, x'_1, \ldots, x'_{m-2}$ such that P_a is positive definite for $a = 0, \ldots, m-2$. Note that

$$P_{m-1} = \begin{pmatrix} x_{m-1} & x'_{m-2} & \cdots & x'_0 \\ x'_{m-2} & \iddots & \iddots & x'_1 \\ \vdots & \iddots & \iddots & \vdots \\ x'_0 & x'_1 & \cdots & x_{m-1} \end{pmatrix}_{m\times m} . \quad (2.13)$$

We shall choose x_{m-1} such that P_{m-1} is positive definite. One possibility is to show the leading principal minors of P_{m-1} are positive definite, then P_{m-1} is positive definite by Sylvester's criterion (Theorem 2.2). As x_{m-1} is the only variable, the kth leading principal minor of P_{m-1} being positive definite will put a linear constraint on x_{m-1} for $k = 2, \ldots, m-1$. Moreover, the coefficient of x_{m-1} is positive, as it equals the $(k-1)$th leading principal minor of P_{m-2} (which is positive as P_{m-2} is positive definite by our assumption). In addition, to make the determinant of P_m positive, we put a quadratic constraint on x_m, where the coefficient of x_m^2 is positive as well. This is due to the fact that it equals the determinant of P'_{m-3}. Thus we have $m-1$ linear constraints and 1 quadratic

constraint on x_m, where the coefficients of x_m in the linear constraints are positive and the coefficients of x_m^2 in the quadratic constraint is positive. We can always choose a sufficient large x'_{m-1} to satisfy all these constraints, which makes P_{m-1} positive definite as well.

Next, we claim that there exist y'_{2m-1} such that Q_1 is positive definite, where

$$Q_1 = \begin{pmatrix} x_{m-2} & \cdots & x_0 & y_m \\ \vdots & \iddots & \iddots & y_{m+1} \\ x_0 & \iddots & \iddots & \vdots \\ y_m & y_{m+1} & \cdots & y_{2m-1} \end{pmatrix}_{m\times m} = \begin{pmatrix} x'_{m-2} & \cdots & x'_0 & x'_1 \\ \vdots & \iddots & \iddots & x'_2 \\ x'_0 & \iddots & \iddots & \vdots \\ x'_1 & x'_2 & \cdots & y_{2m-1} \end{pmatrix}_{m\times m} . \tag{2.14}$$

Note that, the leading principal minors of Q_1 is positive as they coincide with the leading principal minors and determinant of P'_{m-2}. This ensures the first $m-1$ leading principal minor of Q_1 is positive. We can choose $y_{2m-1} \in \mathbb{R}^+$ such that the determinant of Q_1 is positive, which implies that Q_1 is positive definite by Sylvester's criterion (Theorem 2.2). This put a linear constraint on y_{2m-1} with positive coefficient, which equals the determinant of P'_{m-2}. It is easy to see that we can always find $y_{2m-1} \in \mathbb{R}^+$ satisfy the constraint. For $Q_2, \ldots, Q_{n-m}$, we determine $y_{2m}, \ldots, y_{m+n-2}$ adaptively by repeating the above argument, such that Q_b is positive definite for $b = 2, \ldots n-m$.

Finally, we prove that with the above choices of all the variables, Q_b is also positive definite for $b = n-m+1, \ldots, n-1$. Note that Q_{n-m+1} is exactly the $(m-1)$th trailing principal submatrix of Q_{n-m-1}.

$$Q_{n-m+1} = \begin{pmatrix} x_{2m-n-2} & \cdots & x_0 & y_m & \cdots & y_{n-1} \\ \vdots & \iddots & \iddots & \iddots & \iddots & \vdots \\ x_0 & \iddots & \iddots & \iddots & \iddots & \vdots \\ y_m & \iddots & \iddots & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & \iddots & \iddots & \vdots \\ y_{n-1} & \cdots & \cdots & \cdots & \cdots & y_{n+m-3} \end{pmatrix}_{(m-1)\times(m-1)} . \tag{2.15}$$

$$Q_{n-m-1} = \begin{pmatrix} x_{2m-n} & \cdots & x_0 & y_m & \cdots & y_{n-2} \\ \vdots & \iddots & \iddots & \iddots & \iddots & y_{n-1} \\ x_0 & \iddots & \iddots & \iddots & \iddots & \vdots \\ y_m & \iddots & \iddots & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & \iddots & \iddots & \vdots \\ y_{n-2} & y_{n-1} & \cdots & \cdots & \cdots & y_{n+m-3} \end{pmatrix}_{m\times m} . \tag{2.16}$$

By Theorem 2.2, we know the trailing principal minors of Q_{n-m-1} are positive definite, and derive the trailing principal minors of Q_{n-m+1} are positive definite as well, which means Q_{n-m+1} is positive definite by Sylvester's criterion. In fact, it is easy to verify that Q_{n-m+k} corresponding to some principal submatrix of Q_{n-m-1}, for $k = 1, \ldots, m$. By Sylvester's criterion again, we can derive the positive definiteness. This concludes our proof. □

2.4 Summary and Discussion

In this chapter, we have studied the PPT-distinguishability of orthogonal bipartite states in the many-copy scenario. Constructing PPT-indistinguishable orthogonal bipartite states is helpful to deepen our understanding on quantum nonlocality. We reduce deciding the PPT-indistinguishability of orthogonal bipartite states to deciding whether a given bipartite subspace is (strongly) PPT-unexntedible. Such a reduction implies a necessary condition for the PPT-distinguishability. We exploit our condition to show that arbitrary entangled pure state and its orthogonal complement are PPT indistinguishable, even arbitrarily large but finite copies are provided. This extends one of the main results in [YDY14]. We also proved tight lower bound for the dimension of strongly PPT-unextendible subspaces.

It is worth noting that bipartite subspaces are one-to-one corresponding to matrix spaces by utilizing the Choi-Jamiołkowski isomorphism. Thus the PPT-extendibility of bipartite subspaces can be also defined with respect to matrix spaces. Although we have not been benefit from such a conversion, we believe that the language of matrix spaces may provide a general framework in the study of discrimination problems. Such a point will be further emphasized in the next chapter.

Chapter 3

Distinguishing Quantum Channels with Parallel Schemes

In this chapter, we study the distinguishability of quantum channels using parallel schemes. We completely characterize the distinguishability with respect to the matrix space generated by the Choi-Kraus operators of the two given quantum channels, which can be also formulated as a type of extendibility problem. This chapter is organized as follows: We first review the development of quantum channel discrimination in Section 3.1. We introduce some preliminaries in Section 3.2, followed by the study of the parallel distinguishability of quantum channels in Section 3.3. Our main results are based on [DGLL16]. We close in Section 3.4 with a brief summary.

3.1 Introduction: Quantum Channel Discrimination

The quantum channel discrimination problem is a natural generalization of quantum states discrimination. It is the task of identifying an unknown quantum channel, given the promise that it is secretly chosen from a set of known quantum channels. However, characterizing the perfect distinguishability of quantum channels is not as easy as characterizing the perfect distinguishability of quantum states. There are three main differences between these two discrimination problems. First, quantum channels are reusable; second, the input states can be chosen freely, and thus can be entangled with an auxiliary system or between different uses; and third, the unknown quantum channel can be applied in many essentially different ways, such as in parallel or in sequence. The first two viewpoints are quite accessible. While, we review several seminal results to get a better understanding on the third difference. Unitary channels are fundamental to both quantum computation and quantum mechanics, as important concepts such as quantum circuits and the time evolution of quantum systems are contained in the family of unitary channels. Acín [Ací01] and D'Ariano *et al.* [DLPP01] firstly proved that any set of unitary channels can be distinguished perfectly. Their strategy is to prepare an N-partite quantum state and apply the unknown unitary channels N times, each party one time. The resulting quantum states will be orthogonal and can be distinguished perfectly. Intuitively, this kind of protocol is called *parallel scheme*. On the other hand, Duan, Feng and Ying [DFY07] proved that unitary channels can also be distinguished perfectly in a *sequential scheme*, that is, we apply the unknown channel on the input state step by step. These two schemes stand for the use of spatial resources (entanglement or circuits) and the temporal resources (running steps or discriminating time), which enable us to optimize the discrimination protocols by combining these two schemes to fulfill some certain resource requirements. Interestingly, there exist a pair of (entanglement breaking) channels which cannot be perfect distinguished using parallel schemes, this can be done using sequential schemes with only two channel evaluations [HHLW10].

With respect to the fruitful structure of quantum channel discrimination, the perfect or optimal distinguishabilities of other specific families of quantum channels, including the projective measurements, Pauli channels and oracle operators, have been addressed

in [JFDY06, DSK05, CKT$^+$07, Wat08, CY10]. However, until 2009, a complete (mathematical) characterization of the perfect distinguishability of two quantum channels was finally settled in the seminal work of Duan, Feng and Ying [DFY09]. The protocol is a combination of parallel and sequential schemes. We first apply the unknown quantum channel (secretly chosen from channels $\mathcal{E}$ and $\mathcal{F}$) onto an N-partite input states, where N is a finite integer. Then we adapt an auxiliary quantum channel, determined by $\mathcal{E}$ and $\mathcal{F}$, to obtain orthogonal output states. In [DFY09], the condition on $\mathcal{E}$ and $\mathcal{F}$ ensures the two output states can be orthogonal, which can be distinguished perfectly.

However, such perfect discrimination protocols may not be feasible in practice, as implementing the auxiliary channels could be expensive, and they depends crucially on the set of (priori known) quantum channels. Thus, it would be important to obtain a better understanding of the distinguishability of quantum channels using parallel schemes only, as shown in figure 3.1. Moreover, we point out that the parallel scheme is the natural generalization of the discrimination of quantum states *with multiple copies*. Recent highlights are the indentification of the (multiple) *quantum Chernoff bound* [ACMnT$^+$07, Li16].

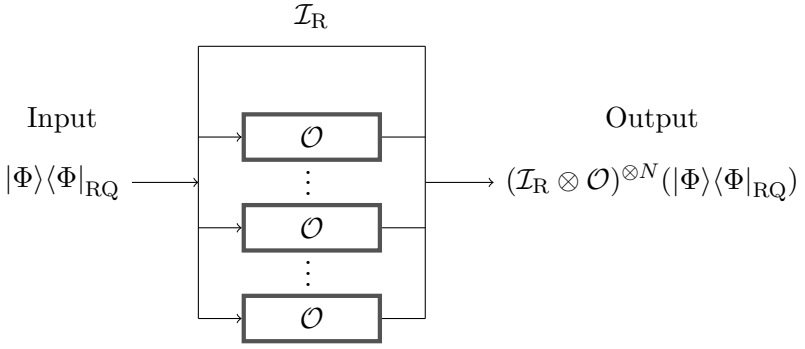


FIGURE 3.1: Parallel scheme to distinguish an unknown quantum channel $\mathcal{O} \in \{\mathcal{E}, \mathcal{F}\}$ with N uses on the input state $|\Phi\rangle\langle\Phi|_{\mathrm{RQ}}$, where $\mathcal{I}_R$ representing the identity channel is applied on the auxiliary system R.

In section 3.3, we investigate the parallel distinguishability for arbitrary two quantum channels. We first convert the problem of deciding parallel discrimination of two quantum channels $\mathcal{E}$ and $\mathcal{F}$ to the problem of deciding whether there exists $k \in \mathbb{N}$, such that the

orthogonal complement of $\mathcal{S}_{\mathcal{E},\mathcal{F}}^{\otimes k}$ contains nonzero positive semidefinite matrices. Here $\mathcal{S}_{\mathcal{E},\mathcal{F}}$ stands for a matrix space generated by the Choi-Kraus operators of $\mathcal{E}$ and $\mathcal{F}$. We then obtain a necessary condition for parallel discrimination, following a similar idea as that used for the PPT-distinguishability of quantum states. Such a condition is also sufficient if the given matrix space is of dimension-1 or self-adjoint. Interestingly, the dimension-1 case provides an alternate proof for the parallel distinguishability of unitary channels [Acín01]. However, we also demonstrate that this condition is not sufficient by providing several illustrative examples.

3.2 Notations and Preliminaries

In this chapter we focus on quantum channels discrimination, and some auxiliary systems are allowed. Quantum channels (CPTP maps) such as $\mathcal{E}, \mathcal{F} : \mathcal{L}(\mathcal{H}_n) \to \mathcal{L}(\mathcal{H}_m)$ are given by their Choi-Kraus operators $\{E_i \in \mathcal{L}(\mathcal{H}_n, \mathcal{H}_m) : i = 1, \ldots, N_{\mathcal{E}}\}$ and $\{F_j \in \mathcal{L}(\mathcal{H}_n, \mathcal{H}_m) : j = 1, \ldots, N_{\mathcal{F}}\}$, respectively. We say $\mathcal{E}$ is an isometry, if $\mathcal{E}$ has only one Choi-Kraus operator[1]. If $\mathcal{E}$ is an isometry and $m = n$, we say $\mathcal{E}$ is a unitary channel. The matrix space $\mathcal{S}_{\mathcal{E},\mathcal{F}} \leq \mathcal{L}(\mathcal{H}_n)$ is generated by the Choi-Kraus operators of $\mathcal{E}$ and $\mathcal{F}$ by $\mathcal{S}_{\mathcal{E},\mathcal{F}} := \mathrm{span}\{E_i^\dagger F_j : i = 1, \ldots, N_{\mathcal{E}},\ j = 1, \ldots, N_{\mathcal{F}}\}$, and the orthogonal complement of $\mathcal{S}_{\mathcal{E},\mathcal{F}}$ is defined as $\mathcal{S}_{\mathcal{E},\mathcal{F}}^{\perp} := \{A \in \mathcal{L}(\mathcal{H}_n) : \mathrm{Tr}(A^\dagger B) = 0,\ \forall\ B \in \mathcal{S}_{\mathcal{E},\mathcal{F}}\}$. Since $\mathcal{L}(\mathcal{H}_n) \cong M(n, \mathbb{C})$, the space of $n \times n$ matrices over complex field, operator spaces such as $\mathcal{S}_{\mathcal{E},\mathcal{F}}$ can be also represented as matrix spaces after claiming a set of linear bases of $\mathcal{H}_n$. We denote the principal system by Q and the auxiliary system by R in the subscript. Note that for any quantum state $\rho \in \mathcal{L}(\mathcal{H}_n)$ and quantum channel $\mathcal{E} : \mathcal{L}(\mathcal{H}_n) \to \mathcal{L}(\mathcal{H}_m)$, there exists a pure state $|\Psi\rangle_{\mathrm{RQ}} \in \mathcal{H}_{n_R} \otimes \mathcal{H}_{n_Q}$, such that $\mathcal{E}(\rho) = \mathrm{Tr}_{\mathrm{R}}(\mathcal{I}_{\mathrm{R}} \otimes \mathcal{E}_{\mathrm{Q}}(|\Psi\rangle\langle\Psi|_{\mathrm{RQ}}))$. Thus in most of the quantum channel discrimination protocols, we allow auxiliary systems (with underlaying Hilbert space $\mathcal{H}_n$) and use pure states as inputs. We say two quantum channel $\mathcal{E}$ and $\mathcal{F}$ are *(entanglement-assisted) disjoint*, if there exist $|\Psi\rangle_{\mathrm{RQ}} \in \mathcal{H}_{n_R} \otimes \mathcal{H}_{n_Q}$ such that $\mathrm{supp}(\mathcal{I}_{\mathrm{R}} \otimes \mathcal{E}_{\mathrm{Q}}(|\Psi\rangle\langle\Psi|_{\mathrm{RQ}})) \cap \mathrm{supp}(\mathcal{I}_{\mathrm{R}} \otimes \mathcal{F}_{\mathrm{Q}}(|\Psi\rangle\langle\Psi|_{\mathrm{RQ}})) = \{0\}$.

[1] Let the Choi-Kraus operator of an isometry $\mathcal{E}$ be $E \in \mathcal{L}(\mathcal{H}_n, \mathcal{H}_m)$, then $E^\dagger E = \mathrm{I}_n$. We also call such an linear operator E as an isometry.

For the perfect distinguishability of two quantum channels, we present the sufficient and necessary condition [DFY09] here:

Theorem 3.1. *Given two quantum channels $\mathcal{E}, \mathcal{F}$, described by the Choi-Kraus operators $\{E_i \in \mathcal{L}(\mathcal{H}_n, \mathcal{H}_m) : i = 1, \ldots, N_\mathcal{E}\}$ and $\{F_j \in \mathcal{L}(\mathcal{H}_n, \mathcal{H}_m) : j = 1, \ldots, N_\mathcal{F}\}$, respectively, then $\mathcal{E}$ and $\mathcal{F}$ are perfectly distinguishable if and only if:* (1) *$\mathcal{E}$ and $\mathcal{F}$ are disjoint;* (2) *$\mathrm{I}_n \notin \mathcal{S}_{\mathcal{E},\mathcal{F}}$.*

Thus, when we are focusing on the parallel distinguishability, the two quantum channels must priorly satisfy the condition in Theorem 3.1. Moreover, the parallel distinguishability can be formally defined as follows:

Definition 3.2. Given two quantum channels $\mathcal{E}, \mathcal{F}$, described by the Choi-Kraus operators $\{E_i \in \mathcal{L}(\mathcal{H}_n, \mathcal{H}_m) : i = 1, \ldots, N_\mathcal{E}\}$ and $\{F_j \in \mathcal{L}(\mathcal{H}_n, \mathcal{H}_m) : j = 1, \ldots, N_\mathcal{F}\}$, respectively, then $\mathcal{E}$ and $\mathcal{F}$ are said to be parallel distinguishable, if there exists $k \in \mathbb{N}$, such that there exists an input state $|\Psi\rangle_{\mathrm{RQ}} \in \mathcal{H}_{n_R}^{\otimes k} \otimes {\mathcal{H}_{n_Q}}^{\otimes k}$, $\mathcal{I}_\mathrm{R} \otimes \mathcal{E}_\mathrm{Q}^{\otimes k}(|\Psi\rangle\langle\Psi|_{\mathrm{RQ}}) \perp \mathcal{I}_\mathrm{R} \otimes \mathcal{F}_\mathrm{Q}^{\otimes k}(|\Psi\rangle\langle\Psi|_{\mathrm{RQ}})$.

The *numerical range* of a linear operator $B \in \mathcal{L}(\mathcal{H}_n)$ is defined as the set $W(B) := \{\langle\psi| B |\psi\rangle : \forall |\psi\rangle \in \mathcal{H}_n,\ \langle\psi|\psi\rangle = 1\}$. The celebrated *Toeplitz-Hausdorff theorem* indicates that $W(B)$ is convex for any linear operator B [HJ12]. the *angular numerical range* of B is defined as $\mathcal{W}(B) := \bigcup_{t>0} W(tB)$. By the convexity of $W(B)$, $\mathcal{W}(B)$ can be $\mathbb{C}$, a half space with a straight line passing through 0 as the boundary, or a pointed cone with 0 as the vertex. We can then define the *field angle* of A according to these cases as follows:

Definition 3.3. For a linear operator $B \in \mathcal{L}(\mathcal{H}_n)$, the *field angle* of B, denoted by $\Theta(B)$, is defined as follows:

- If $\mathcal{W}(B) = \mathbb{C}$, $\Theta(B) = 2\pi$;
- If $\mathcal{W}(B)$ is a half space, then $\Theta(B) = \pi$;
- If $\mathcal{W}(B)$ is a pointed cone, then $\Theta(B)$ is the angle between the two boundary rays of the cone.

3.3 Parallel Distinguishability of Quantum Channels

3.3.1 Characterizing the Parallel Distinguishability

We first derive an equivalent formulation of the parallel distinguishability of two quantum channels $\mathcal{E}, \mathcal{F}$. $\mathcal{E}$ and $\mathcal{F}$ are parallel distinguishable if and only if there exists an input state $|\Psi\rangle^{\mathrm{RQ}}$ such that $\mathcal{I}_{\mathrm{R}} \otimes \mathcal{E}_{\mathrm{Q}}(|\Psi\rangle\langle\Psi|^{\mathrm{RQ}}) \perp \mathcal{I}_{\mathrm{R}} \otimes \mathcal{F}_{\mathrm{Q}}(|\Psi\rangle\langle\Psi|^{\mathrm{RQ}})$. Equivalently, we have

$$\mathrm{Tr}(\mathcal{I}_{\mathrm{R}} \otimes \mathcal{E}_{\mathrm{Q}}(|\Psi\rangle\langle\Psi|^{\mathrm{RQ}})\mathcal{I}_{\mathrm{R}} \otimes \mathcal{F}_{\mathrm{Q}}(|\Psi\rangle\langle\Psi|^{\mathrm{RQ}})) = 0, \tag{3.1}$$

Note that we have $|\Psi\rangle^{\mathrm{RQ}} = (\mathrm{I}_n \otimes X)\,|\mathrm{EPR}_n\rangle_{\mathrm{RQ}}$, where $\mathrm{Tr}(X^\dagger X) = 1$ and $|\mathrm{EPR}_n\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} |i\rangle\,|i\rangle$ stands for the maximally entangled state. Replacing $\mathcal{E}$ and $\mathcal{F}$ by their Choi-Kraus operators $\{E_i :\ i = 1, \ldots, N_{\mathcal{E}}\}$ and $\{F_j :\ j = 1, \ldots, N_{\mathcal{F}}\}$, we obtain

$$\sum_{i,j} \mathrm{Tr}(E_i^\dagger F_j X X^\dagger)\mathrm{Tr}(F_j^\dagger E_i X X^\dagger) = 0. \tag{3.2}$$

We derive that $XX^\dagger$ needs to be orthogonal to $E_i^\dagger F_j$ for all possible i and j, i.e. $XX^\dagger \in \mathcal{S}_{\mathcal{E},\mathcal{F}}^{\perp}$. Note that the Choi-Kraus operator of $\mathcal{E}^{\otimes k}$ is $\{E_{i_1} \otimes \cdots \otimes E_{i_k} :\ \forall j = 1, \ldots, k,\ i_j = 1, \ldots, N_{\mathcal{E}}\}$, which leads to the following:

Proposition 3.4. *Given two quantum channels $\mathcal{E}, \mathcal{F} : \mathcal{L}(\mathcal{H}_n) \to \mathcal{L}(\mathcal{H}_m)$, described by the Choi-Kraus operators $\{E_i :\ i = 1, \ldots, N_{\mathcal{E}}\}$ and $\{F_j :\ j = 1, \ldots, N_{\mathcal{F}}\}$, respectively, then $\mathcal{E}$ and $\mathcal{F}$ are parallel distinguishable within k uses of the unknown quantum channel if and only if there exist nonzero positive semidefinite operators in the orthogonal complement of $S_{\mathcal{E},\mathcal{F}}^{\otimes k}$.*

Proposition 3.4 illustrates that the parallel distinguishability of $\mathcal{E}, \mathcal{F}$ is completely characterized by the operator space $\mathcal{S}_{\mathcal{E},\mathcal{F}} \leq \mathcal{L}(\mathcal{H}_n) \cong M(n, \mathbb{C})$. A natural question to ask is, given an arbitrary matrix space $\mathcal{S} \leq M(n, \mathbb{C})$, can we construct two quantum channels $\mathcal{E}, \mathcal{F}$ such that $\mathcal{S} = \mathcal{S}_{\mathcal{E},\mathcal{F}}$? The answer is affirmative. Assume $\mathcal{S}$ is spanned by a tuple of matrices $B_1, \ldots, B_N$, where $N \leq n^2$ is the dimension of $\mathcal{S}$. Without loss of generality, we assume $B_i^\dagger B_i \leq \mathrm{I}_n$ for $i = 1, \ldots, N$. We prove that for each B_i, there exist two isometries $U_i, V_i \in \mathcal{L}(\mathcal{H}_n, \mathcal{H}_{n'})$ where $n' \geq 2n$, such that $B_i = U_i^\dagger V_i$. To see this,

let the singular value decomposition of B_i be $\sum_{k=0}^{n-1} \sigma_{i,k} |\psi_{i,k}\rangle \langle\phi_{i,k}|$, where $0 \leq \sigma_{i,k} \leq 1$. Let $U_i = \sum_{k=1}^{n} |\alpha_{i,k}\rangle \langle\psi_{i,k}|$ and $V_i = \sum_{k=1}^{n} |\beta_{i,k}\rangle \langle\phi_{i,k}|$. For each i, $\{|\alpha_{i,0}\rangle, |\alpha_{i,n-1}\rangle\}$ and $\{|\beta_{i,0}\rangle, \ldots, |\beta_{i,n-1}\rangle\}$ are two sets of orthogonal states in $\mathcal{H}_n$, which need to be determined such that $B_i = U_i^\dagger V_i$. To achieve this, we need

$$\langle\alpha_{i,j}|\beta_{i,k}\rangle = 0 \;\; \langle\alpha_{i,k}|\beta_{i,k}\rangle = \sigma_{i,k}, \;\; \forall \; j,k \in \{0, \ldots n-1\}. \tag{3.3}$$

To achieve this, we choose n mutually orthogonal dimension-2 subspaces in $\mathcal{H}_{n'}$, denoted by $K_{i,k}$ for $k = 0, \ldots, n-1$. In each $K_{i,k}$ we can choose two quantum states $|\alpha_{i,k}\rangle, |\beta_{i,k}\rangle$ such that $\langle\alpha_{i,k}|\beta_{i,k}\rangle = \sigma_{i,k}$. This can be done since $0 \leq \sigma_{i,k} \leq 1$. (When $\sigma_{i,k} = 1$, $K_{i,k}$ can be reduced to a dimension 1 subspace.) For $n' \geq 2n$, these subspaces always exist. In the end, we can construct two quantum channels $\mathcal{E}$ and $\mathcal{F}$ with Choi-Kraus operators $\{\frac{1}{\sqrt{N}} U_i \otimes |i\rangle : i = 1, \ldots, N\}$ and $\{\frac{1}{\sqrt{N}} V_j \otimes |j\rangle : j = 1, \ldots, N\}$ such that $\mathcal{S} = \mathcal{S}_{\mathcal{E},\mathcal{F}}$.

Combine Proposition 3.4 and the above construction, the parallel distinguishability of quantum channels can be formulated as follows: Given a matrix space $\mathcal{S} \leq M(n, \mathbb{C})$, decide whether there exists $k \in \mathbb{N}$, such that the orthogonal complement of $\mathcal{S}^{\otimes k}$ contains nonzero positive semidefinite matrices. Similarly, we say $\mathcal{S}$ is k-positive-extendible, if $\mathcal{S}^{\otimes k}$ contains nonzero positive semidefinite matrices. Otherwise we say $\mathcal{S}$ is k-positive-unextendible. If for any $k \in \mathbb{N}$, $\mathcal{S}$ is k-positive-unextendible, we say $\mathcal{S}$ is strongly positive-unextendible. Note that Theorem 2.4 also implies a sufficient condition to determine whether a matrix space is strongly positive-unextendible:

Theorem 3.5. *If $\mathcal{S} \leq M(n, \mathbb{C})$ contains positive definite matrices, then $\mathcal{S}$ is strongly positive-unextendible.*

Exploiting this theorem, we can immediately obtain a necessary condition for parallel distinguishability:

Corollary 3.6. *Let $\mathcal{E}, \mathcal{F} : \mathcal{L}(\mathcal{H}_n) \to \mathcal{L}(\mathcal{H}_m)$ be two quantum channels, described by the Choi-Kraus operators $\{E_1, \ldots, E_{N_\mathcal{E}}\}$ and $\{F_1, \ldots, F_{N_\mathcal{F}}\}$, respectively. If $S_{\mathcal{E},\mathcal{F}}$ contains positive definite operators, then $\mathcal{E}$ and $\mathcal{F}$ are not parallel distinguishable.*

3.3.2 Determining the Parallel Distinguishability for Two Families of Quantum Channels

Corollary 3.6 turns out to be surprisingly useful. For instance, Harrow *et al.* [HHLW10] construct two quantum channels $\mathcal{E}, \mathcal{F}$ such that they are not parallel distinguishable, but can be distinguished perfectly by sequential schemes. In particular, the matrix space $\mathcal{S}_{\mathcal{E},\mathcal{F}}$ with respect to these channels contains positive definite matrices. In this subsection, we investigate when corollary 3.6 is also sufficient to determine the parallel distinguishability.

Theorem 3.7. *Given a matrix space $\mathcal{S} \in M(n, \mathbb{C})$, if $\mathcal{S}$ is self adjoint, i.e. $\mathcal{S}^\dagger = \{B^\dagger : B \in \mathcal{S}\}$ or $\dim(\mathcal{S}) = 1$, then $\mathcal{S}$ is strongly positive-unextendible* if and only if *$\mathcal{S}$ contains positive definite matrices. Moreover, if $\mathcal{S}$ contains no positive definite matrix, we can determine the minimum integer k where the orthogonal complement of $\mathcal{S}^{\otimes k}$ contains nonzero positive definite operator.*

Remark 3.8. Theorem 3.7 implies that two different unitary channels $\mathcal{U}, \mathcal{V}$ (with Choi-Kraus operators U and V, respectively) are parallel distinguishable, which reproves the result in [Acín01]. Since $\mathcal{S}_{\mathcal{U},\mathcal{V}} = \langle U^\dagger V \rangle$ is rank 1, and $U^\dagger V$ is not positive definite if and only if $U \neq V$.

Proof. We first deal with the case when $\mathcal{S}$ is self-adjoint. Then $\mathcal{S}$ admits a set of Hermitian basis, i.e. $\mathcal{S} = \langle B_1, \ldots, B_N \rangle$ where B_i is Hermitian for $i = 1, \ldots, N$ and $\langle \cdot \rangle$ denotes the linear span. By Farkas' lemma of semi-definite programming [Roc15], either

- There is a linear combination of $B_1, \ldots, B_N$ equal to a positive definite matrix; or
- There is a nonzero positive semidefinite matrix T such that $\mathrm{Tr}(B_i T) = 0$ for $i = 1, \ldots, N$.

The first statement is equivalent to the existence of a positive definite matrix in $\mathcal{S}$ and the second one is equivalent to $T \in \mathcal{S}^\perp$. Thus if there is no **positive definite** matrix in S, we can always find a nonzero positive semidefinite matrix in $\mathcal{S}^\perp$. This also indicates that, if $\mathcal{S}_{\mathcal{E},\mathcal{F}}$ is self-adjoint, they can be perfectly distinguished (with single use) if and only if $\mathcal{S}_{\mathcal{E},\mathcal{F}}$ contains no positive definite operator.

Now we consider the case when $\dim(\mathcal{S}) = 1$. Let $\mathcal{S} = \langle B \rangle$, where $B \in \mathcal{L}(\mathcal{H}_n)$. We shall prove that, if B is not positive definite, there exists $k \in \mathbb{N}$ such that we can find a nonzero positive semidefinite matrix P such that $\mathrm{Tr}(B^{\otimes k}P) = 0$. Note that it is equivalent to have $0 \in W(B^{\otimes k})$ for some $k \in \mathbb{N}$. If $0 \notin W(B)$, we know that $W(B) \not\subseteq e^{it}(0, \infty)$ for some $t \in \mathbb{R}$, otherwise $e^{it}B \in \mathcal{S}$ is positive definite. Thus, the angular numerical range $\mathcal{W}(B)$ is a cone in $\mathbb{C}$ with vertex 0 and contains $W(B)$. So, there are $\mu_1 = r_1 e^{i\theta_1}, \mu_2 = r_2 e^{i\theta_2} \in W(B)$ with $r_1, r_2 > 0$ and $\theta_1 < \theta_2 < \theta_1 + \pi$ so that $\theta_1 \leq \arg(\mu) \leq \theta_2$ for all $\mu \in W(B)$, where $\arg(\mu)$ denotes the argument of the complex number μ. This also indicates that the field angle of B equals $\theta_2 - \theta_1$. For simplicity, we may replace B by $e^{-i\frac{\theta_1+\theta_2}{2}}B$ and assume that

$$W(B) \subseteq \{\mu \in \mathbb{C} : -\frac{\Theta(B)}{2} \leq \arg(\mu) \leq \frac{\Theta(B)}{2}\}.$$

Now let $B = H + iG$, where H, G are Hermitian matrices and H need to be positive definite. In addition, we can find a unitary U such that

$$B_0 := U^\dagger H^{-1/2} B H^{-1/2} U = U^\dagger (\mathrm{I}_n + iH^{-1/2} G H^{-1/2}) U = \begin{bmatrix} 1 + a_1 i & & \\ & \ddots & \\ & & 1 + a_d i \end{bmatrix}, \quad (3.4)$$

where $a_1 \geq \cdots \geq a_d$. It is clear that $\mathcal{W}(B) = \mathcal{W}(B_0)$ and $a_1 = \tan\frac{\Theta(B)}{2}$ and $a_d = -\tan\frac{\Theta(B)}{2}$. Furthermore, we have $\mathcal{W}(B^{\otimes k}) = \mathcal{W}(B_0^{\otimes k})$. Since B_0 is diagonal, the field angle of $B_0^{\otimes k}$ equals $k\Theta(B_0) = k\Theta(B)$. Then $0 \in W(B^{\otimes k})$ if and only if $k \geq \frac{\pi}{\Theta(B)}$, which is always finite if $\Theta(B) > 0$.

In conclusion, if $B \in \mathcal{L}(\mathcal{H}_n)$ is not positive definite, than there exist nonzero positive semidefinite operators which are orthogonal to $B^{\otimes \lceil \frac{\pi}{\Theta(B)} \rceil 2}$. And for any $k < \frac{\pi}{\Theta(B)}$, there is no nonzero positive semidefinite operator which is orthogonal to $B^{\otimes k}$. □

3.3.3 A Counterexample for the Sufficiency of Corollary 3.6

Although corollary 3.6 appears to be useful in several interesting cases, we present the following simple matrix space $\mathcal{S}$ which not only contains no positive definite matrix, but

[2] $\lceil x \rceil$ equals the smallest integer which is larger than $x \in \mathbb{R}$.

also strongly positive-unextendible.

Theorem 3.9. *Let* $\mathcal{S} = \langle B_1, B_2 \rangle \leq M(n, \mathbb{C})$ *with* $B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 0 \end{bmatrix}$ *and* $B_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & i \end{bmatrix}$. *Then* $\mathcal{S}$ *contains no positive definite matrix and is unextendible.*

Proof. It is easy to verify that there is no positive definite matrix in $\mathcal{S}$. Now we use mathematical induction to show that for arbitrary $k \in \mathbb{N}$, there is no nonzero positive semidefinite matrix in the orthogonal complement of $S^{\otimes k}$. When $k = 1$, it is easy to verify $S^{\perp}$ contains no positive definite matrix. Assume for $k = N$, there is no nonzero positive semidefinite matrix in $(S^{\otimes N})^{\perp}$. For $k = N+1$, such a positive semidefinite matrix $T \in (S^{\otimes N+1})^{\perp}$ exists, we have

$$\mathrm{Tr}(T(B_1 \otimes M)) = 0, \quad \mathrm{Tr}(T(B_2 \otimes M)) = 0, \tag{3.5}$$

where $M \in \mathcal{S}^{\otimes N}$. Since B_1 and B_2 are diagonal, without loss of generality we assume $T = \begin{bmatrix} T_0 & 0 & 0 \\ 0 & T_1 & 0 \\ 0 & 0 & T_2 \end{bmatrix}$ where T_0, T_1 and T_2 are positive semidefinite and at least one of them is nonzero. Rewriting Equation (3.5), we obtain

$$\mathrm{Tr}(T_0 M) + i\mathrm{Tr}(T_1 M) = 0, \ -i\mathrm{Tr}(T_1 M) + \mathrm{Tr}(T_2 M) = 0. \tag{3.6}$$

If $T_0 + T_2 \neq 0$, let $T' = T_0 + T_2$ and $\mathrm{Tr}(T'M) = 0$ for all $M \in \mathcal{S}^{\otimes N}$. Thus T' is a nonzero positive semidefinite matrix in the orthogonal complement of $S^{\otimes N}$, which is a contradiction. Otherwise, we have $T_0 = T_2 = 0$, $T_1 \neq 0$ and $\mathrm{Tr}(T_1 M) = 0$ for all $M \in \mathcal{S}^{\otimes N}$, again a contradiction. Thus there is no nonzero positive semidefinite matrix in the orthogonal complement of $\mathcal{S}^{\otimes N+1}$. □

The matrix space in Theorem 3.9 can be applied to construct another pairs of quantum

channels which are not parallel distinguishable but perfect distinguishable. Let the Choi-Kraus operators of $\mathcal{E}, \mathcal{F} : \mathcal{H}_3 \to \mathcal{H}_8$ be

$$\begin{aligned} &\{\frac{1}{\sqrt{2}}(|0\rangle\langle 0| - |1\rangle\langle 1| + |3\rangle\langle 2|) \otimes |0\rangle\,, \frac{1}{\sqrt{2}}(|1\rangle\langle 1| - |2\rangle\langle 2| + |3\rangle\langle 0|) \otimes |1\rangle\}, \\ &\{\frac{1}{\sqrt{2}}(|0\rangle\langle 0| - i\,|1\rangle\langle 1| + |2\rangle\langle 2|) \otimes |0\rangle\,, \frac{1}{\sqrt{2}}(|1\rangle\langle 1| - i\,|2\rangle\langle 2| + |0\rangle\langle 0|) \otimes |1\rangle\}, \end{aligned} \tag{3.7}$$

respectively. Denote $\mathcal{S}_\mathcal{E}$ and $\mathcal{S}_\mathcal{F}$ as the operator space spanned by the Choi-Kraus operators of $\mathcal{E}$ and $\mathcal{F}$. Then it is easy to verify $\mathcal{S}_\mathcal{E} \cap \mathcal{S}_\mathcal{F} = \{0\}$. Following the procedure introduced in [DFY09], we can choose the 3×3 maximally entangled state $|\mathrm{EPR}_3\rangle_{\mathrm{RQ}} = \frac{1}{\sqrt{3}}(|00\rangle + |11\rangle + |22\rangle)$ such that $\mathrm{supp}(\mathcal{I}_\mathrm{R} \otimes \mathcal{E}_\mathrm{Q}(|\mathrm{EPR}_3\rangle\langle\mathrm{EPR}_3|)) \cap \mathrm{supp}(\mathcal{I}_\mathrm{R} \otimes \mathcal{F}_\mathrm{Q}(|\mathrm{EPR}_3\rangle\langle\mathrm{EPR}_3|)) = \{0\}$, i.e. $\mathcal{E}$ and $\mathcal{F}$ are disjoint. Since $\mathcal{S}_{\mathcal{E},\mathcal{F}}$ has no positive definite matrix, we know $\mathrm{I}_3 \notin \mathcal{S}_{\mathcal{E},\mathcal{F}}$. By Theorem 3.1, we know $\mathcal{E}$ and $\mathcal{F}$ can be distinguished perfectly.

3.4 Summary and Discussion

In this chapter, we have investigated the parallel distinguishability of quantum channels. The motivations of studying the parallel distinguishability of quantum channels is realistic: we would like to avoid the use of adaptive strategy and auxiliary quantum channels in channel discrimination tasks. We have derived a necessary condition to decide the parallel distinguishability of quantum channels, based on the characterization with respect to matrix spaces. In addition, we have proved that the necessary condition is also sufficient to determine the parallel distinguishability for two families of quantum channels, including unitary channels [Ací01]. We also exhibit two quantum channels which neither satisfy our necessary condition, nor can they be distinguished using parallel schemes.

To obtain our results, a key step is to derive the characterization of parallel distinguishability terms of matrix spaces. This opens up the possibilities to utilizing powerful mathematical tools in matrix analysis, such as the theory of numerical range and Farkas' lemma for semidefinite programming. It is also worth noting that, the structure of matrix spaces and their orthogonal complement (with respect to the Hilbert-Schmidt inner product) is much more complicated than that of vector spaces, which may deserve further study.

Chapter 4

Tripartite-to-Bipartite SLOCC Entanglement Transformation

In this chapter, we investigate the feasibility of transforming a tripartite pure state to a bipartite one via SLOCC, in both finite-copy and asymptotic settings. Notably, deciding the feasibility can be reduced to computing the *maximal rank* of a given matrix space, which has been studied in the context of computational complexity theory for decades. We first review the development of (tripartite-to-bipartite) SLOCC entanglement transformation and illustrate the closed connection with computational complexity theory in Section 4.1. Section 4.2 summarizes several previous results and mathematical tools. We study the tripartite-to-bipartite SLOCC entanglement transformation with multiple copies in Section 4.3, and focus on the asymptotic setting in Section 4.4. We close in Section 4.5 with a brief summary. This chapter is based on [LQWD18].

4.1 Introduction

Quantum entanglement is unarguably of great utility in quantum information processing and quantum computing. It is instrumental in quantum computational speed-up, quantum communication, quantum cryptography and so on. Exploring the structure of quantum entanglement is at the very heart of quantum information theory and one of the most fruitful branches is to discuss the possibility of transforming a pure entangled state into another one, using local operations and classical communication (LOCC).

The most seminal work in the study of *bipartite* entanglement transformation is by Nielsen [Nie99]. He proved that the feasibility of converting a pure entangled state $|\psi\rangle_{\mathrm{AB}}$, shared by Alice and Bob, to another pure entangled state $|\phi\rangle_{\mathrm{AB}}$ by LOCC (symbolically denoted by $|\psi\rangle_{\mathrm{AB}} \stackrel{\mathrm{LOCC}}{\longrightarrow} |\phi\rangle_{\mathrm{AB}}$) is completely characterized by the *majorization condition*: $|\psi\rangle_{\mathrm{AB}} \stackrel{\mathrm{LOCC}}{\longrightarrow} |\phi\rangle_{\mathrm{AB}}$ if and only if $(\lambda_0, \ldots, \lambda_{r-1})$ is *majorized* by $(\mu_0, \ldots, \mu_{r-1})$[1], where $\lambda_0 \geq \cdots \geq \lambda_{r-1}$ and $\mu_0 \geq \cdots \geq \mu_{r-1}$ are Schmidt coefficients of $|\psi\rangle_{\mathrm{AB}}$ and $|\phi\rangle_{\mathrm{AB}}$, respectively. However, the majorization condition also indicates that pairs of bipartite entangled states exist which cannot be converted to each other. Consequently, two surprising phenomena were subsequently explored: Jonathan and Plenio [JP99] observed that there exist bipartite entangled states $|\psi_1\rangle_{\mathrm{AB}}$, $|\psi_2\rangle_{\mathrm{AB}}$ and $|\phi\rangle_{\mathrm{AB}}$, such that $|\psi_1\rangle_{\mathrm{AB}} \stackrel{\mathrm{LOCC}}{\not\longrightarrow} |\psi_2\rangle_{\mathrm{AB}}$ but $|\psi_1\rangle_{\mathrm{AB}} \otimes |\phi\rangle_{\mathrm{AB}} \stackrel{\mathrm{LOCC}}{\longrightarrow} |\psi_2\rangle_{\mathrm{AB}} \otimes |\phi\rangle_{\mathrm{AB}}$. The role of $|\phi\rangle_{\mathrm{AB}}$ is similar to a *catalyst* in a chemical process and this type of LOCC transformation is called the *catalyst-assisted LOCC transformation*. On the other hand, Bandyopadhyay et al. [BRS02] discovered that there exist bipartite entangled states $|\psi_1\rangle_{\mathrm{AB}}$, $|\psi_2\rangle_{\mathrm{AB}}$ such that $|\psi_1\rangle_{\mathrm{AB}} \stackrel{\mathrm{LOCC}}{\not\longrightarrow} |\psi_2\rangle_{\mathrm{AB}}$ but $|\psi_1\rangle_{\mathrm{AB}}^{\otimes k} \stackrel{\mathrm{LOCC}}{\longrightarrow} |\psi_2\rangle_{\mathrm{AB}}^{\otimes k}$ for some positive integer k. The use of multiple copies in entanglement transformations is referred to as the *Multi-copy LOCC transformations*. Both types of transformations have been extensively studied, see [DFJY05, DFLY05b, DFLY05a, FDY06, SDY05, FDY06] and the references therein for a partial list.

Although Nielsen's condition is mathematically concise, it is hard to generalize for multipartite entanglement. Hitherto, characterizations for multipartite LOCC entanglement transformation have only been settled for specific families of multipartite states [XD07,

[1] $x = \{x_1, \ldots, x_n\}$ is majorized by $y = \{y_1, \ldots, y_n\}$ if for each $k = 1, \ldots, n$, $\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}$ with equality holding when $k = d$, where the $\downarrow$ indicates that elements are to be taken in descending order.

KT10, TGP10]. Meanwhile, the analysis of LOCC transformations does not allow us to classify entangled states into some *coarse grained classes*, which may be used to provide a rough but more transparent picture. This is due to the fact that continuous labels are needed to parametrize classes of entangled states which admits local unitary transformations, even in the bipartite case. To remedy this obstacle, one possible solution is to consider stochastic LOCC (SLOCC) transformations, i.e. seeking a LOCC protocol which transforms a multipartite entangled state to another with a non-vanishing probability. Such a paradigm successfully works in the bipartite case. Vidal [Vid99] proved that a pure entangled state $|\psi\rangle_{\mathrm{AB}}$ can be converted to another pure entangled state $|\phi\rangle_{\mathrm{AB}}$ by SLOCC if and only if the Schmidt rank of $|\psi\rangle_{\mathrm{AB}}$ is no less than that of $|\phi\rangle_{\mathrm{AB}}$. Furthermore, he derived a simple formula which gave the optimal probability for bipartite conversion (with a single copy). Afterwards, Feng, Duan and Ying illustrated that the catalyst may also boost the optimal probability in SLOCC bipartite entanglement transformations [FDY05]. They also proved that, catalyst-assisted and multi-copy LOCC transformations achieve the same optimal success probability, which established a surprising equivalence between these two phenomena [DFY05].

However, even in the SLOCC setting, the situation becomes much more complicated when the number of subsystems increases. Dür, Vidal and Cirac [DVC00] concluded that a three-qubit system can be partitioned into 6 equivalence classes (defined by SLOCC convertibility between states in the same class). Such a classification enables us to reduce the feasibility of SLOCC transformations to deciding whether two given three-qubit states belong to the same equivalent class. Unfortunately, the same approach fails for more than four qubits, as four-qubit systems contain an uncountable number of SLOCC inequivalent classes [VDDMV02, GW10]. In turn, we may need to seek relatively simple criterion for determining the convertibility of arbitrary multipartite states, like the Schmidt rank for bipartite SLOCC transformation. The notion of "simple" here can be made precise by using the language of *computational complexity theory*, which groups problems according to the amount of resources needed to solve them. For instance, determining the SLOCC transformations of bipartite pure states requires omputational resources to increase polynomially in the dimension of either subsystem undergoing the transformation. Thus, the bipartite SLOCC transformation problem belongs to the complexity class P, the set of

problems which admits a deterministic polynomial-time algorithm[2]. However, Chitambar, Duan and Shi proved that deciding the SLOCC transformation of tripartite pure states is NP-hard, which is widely believed to be *intractable*. Such a result rules out the possibility of the existence of simple criterion to decide the SLOCC convertibility of tripartite or multipartite states.

Although multipartite SLOCC transformation is hard in general, we may focus on special but realistic cases. One of the most famous one is to convert a specific bipartite state $|\varphi\rangle_{\mathrm{AB}}$, shared by Alice and Bob, from a tripartite state $|\Psi\rangle_{\mathrm{ABC}}$, shared by Alice, Bob and Charlie. In this circumstance, for Charlie's part, he could perform arbitrary local operations and communicates classical information to *assist* Alice and Bob. In the LOCC setting, such a transformation has been studied under the name of *entanglement of assistance* [DFM$^+$99, SVW05]. Such a model was then studied extensively in different settings, which introduced several new concepts and applications such as *localizable entanglement* [VPC04, PVMDC05], *concurrence of assistance* [GMS05], *random state entanglement* [FL07], *entanglement of collaboration* [GS06, Gou06] and *entanglement combing* [YE09].

In the SLOCC setting, Chitambar, Duan and Shi [CDS10] studied the multipartite-to-bipartite SLOCC transformations. They identified a surprising *algorithmic* connection: determining the multipartite-to-bipartite SLOCC transformation is equivalent to the celebrated *polynomial identity testing* (PIT) problem, which lies in the heart of (algebraic) complexity theory. Such an equivalence also ensures multipartite-to-bipartite SLOCC transformations to be tractable, as PIT admits a *randomized* polynomial-time algorithm using Schwartz-Zippel lemma [Sch80]. Notably, when we focus on tripartite states, the convertibility is completely determined by the *maximal rank* of tripartite states and the Schmidt rank of the bipartite states [CDS10]. Here, the maximal rank of a tripartite state $|\Psi\rangle_{\mathrm{ABC}}$, denoted as $mrk(\Psi_{\mathrm{ABC}})$ equals the largest Schmidt rank of bipartite states, which belong to the support of the reduced density operator $\mathrm{Tr}_{\mathrm{C}}(|\Psi\rangle\langle\Psi|_{\mathrm{ABC}})$. Utilizing the Choi-Jamiołkowski isomorphism between bipartite states and linear operators, computing the maximal rank of a given tripartite pure state is equivalent to computing the maximal rank

[2]In this chapter, all the algorithms which related to deciding the convertibility of entangled states have as input the *classical description* of the states.

of a given matrix space, which equals the largest rank of matrices in the matrix spaces. Note that in [CMW08], Cubbit, Montanaro and Winter have discussed the dimension of subspaces with bounded Schmidt rank, which then transformed into the study of *matrix spaces with bounded rank*.

Matrix Spaces, shrunk Subspaces and invariant Theory. As the tripartite case is our main interest, we take a short cut and introduce the development of computing the maximal rank of a matrix space. The maximal rank problem has been studied extensively in the community of computational complexity theory under the name of th *Edmonds' problem* [Edm67][3]. The decision version of the Edmonds' problem is to decide whether a given matrix space contains full rank matrices; and is better known as the *Symbolic Determinant Identity Testing* (SDIT) problem, which is equivalent to PIT for weakly-skew arithmetic circuits. The original motivation for studying Edmonds' problem is its applications to certain combinatorial problems, most notably the maximum matching problem on graphs, as exploited by Tutte [Tut47], Edmonds [Edm67] and Lovász [Lov89]. Since 2003, a major reason to study SDIT is based on implications to circuit lower bounds, as shown in the seminal work by Kabanets and Impagliazzo [KI04]. Note that SDIT admits a *randomized* polynomial-time algorithm [Lov79]. However, *derandomizing* SDIT could be extremely difficult, as Kabanets and Impagliazzo [KI04] proved that an explicit deterministic polynomial-time algorithm for SDIT would imply strong circuit lower bounds which seem beyond the current techniques.

An important concept related to the maximal rank of matrix spaces is *shrunk subspaces*: For $\mathcal{B} \leq M(n, \mathbb{C})$, subspace $U \leq \mathbb{C}^d$ satisfying $\dim(\mathcal{B}(U)) < \dim(U)$ ($\mathcal{B}(U) := \text{span}\{B(U) : B \in \mathcal{B}\}$) is called a shrunk subspace of $\mathcal{B}$. Shrunk subspaces emerge in several mathematical areas. The first appearance of shrunk subspaces seems to be in T. G. Room's treatise on determinants in the 1930s [Roo38]. An intuitive way to understand shrunk subspaces is to view them as *linear algebraic analog* of shrunk subsets as in *Hall's marriage theorem* [Hal35]. Recall that for a bipartite graph $G = (L \cup R, E)$ where $|L| = |R|$, Hall's marriage theorem states that G has a perfect matching if and only if G does not

[3]Edmonds' problem was originally stated with respect to symbolic matrices over integer field, and can be extended to matrix spaces over an arbitrary sufficiently large field in the literature (e.g. in complex field [Gur03]). We only point out the underlaying field in our statement when necessary, otherwise it is considered as the complex field.

have a shrunk subset, that is a subset $S \subseteq L$ such that $|S| > |N(S)|$ where $N(S)$ denotes the set of neighbors of S. Getting back to the matrix space setting, we postulate viewing matrix spaces as a *linear algebraic analog of the bipartite graphs*. We view vector spaces $U \cong V \cong \mathbb{C}^n$ as a linear algebraic analog of the left and right vertex sets of size n; we view the matrix space $\mathcal{B}$ of dimension m, of which the matrices epresents linear maps from U to V, as a linear algebraic analog of the edge set E of size m. In such a linear algebraic world, if the given matrix space contains full-rank matrices, the linear algebraic analog of bipartite graph has a perfect matching, which is essentially a bijection between the left and right vertex set. Similarly, the shrunk subspaces are naturally viewed as the linear algebraic analog of the shrunk subset. It is clear that if $\mathcal{B}$ possesses shrunk subspaces, then $\mathcal{B}$ contains no full-rank matrices. However, unlike in the classical setting, it is not true that any singular matrix space has shrunk subspaces. For instance, the 3×3 skew-symmetric matrix space

$$\left\langle \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \right\rangle \leq M(3, \mathbb{C}) \tag{4.1}$$

has neither a full-rank matrix nor a shrunk subspace [FR04].

Shrunk subspaces appear in *non-commutative algebra* as follows. Suppose $\mathcal{B} \leq M(n, \mathbb{F})$ is a matrix space over some field $\mathbb{F}$ and spanned by $\{B_1, \ldots, B_m\} \subseteq M(n, \mathbb{F})$. Let $\{x_1, \ldots, x_m\}$ be a set of *non-commuting* variables, and form a matrix $B = x_1 B_1 + \cdots + x_m B_m$ whose entries are linear forms in x_i's. Matrices of this type have been studied in non-commutative algebra in the context of the *free skew field* since the 1970s [Coh75]. The rank of such a matrix over the free skew field, which was named *non-commutative rank* and denoted by $ncrk(\cdot)$, was shown to be the minimum integer c such that there exists a subspace $U \leq \mathbb{F}^n$ with $\dim(U) - \dim(\mathcal{B}(U)) = n - c$ [FR04]. Thus, $ncrk(\mathcal{B}) < n$ if and only if $\mathcal{B}$ have shrunk subspaces.

Another way to reach the concept of shrunk subspaces is to consider matrix spaces with maximal ranks bounded from above [EH88]. Characterizing these matrix spaces is known to be a difficult problem; in fact, such matrix spaces basically correspond to certain torsion-free sheaves on projective spaces [EH88]. To make progress on this topic, one approach is to consider certain "witnesses" that can serve as an upper bound on the maximal rank.

As previously discussed, shrunk subspaces can be used as such witnesses. Specifically, if a matrix space $\mathcal{B} \leq M(n, \mathbb{F})$ has a shrunk subspace U with $\dim(U) - \dim(\mathcal{B}(U)) = c > 0$, then it is clear that $mrk(\mathcal{B}) \leq n - c$.

An important characterization of matrix spaces with shrunk subspaces comes from *invariant theory*. Consider the group action of $(A, C) \in \mathrm{SL}(n, \mathbb{F}) \times \mathrm{SL}(n, \mathbb{F})$ on $M(n, \mathbb{F})^{\oplus m}$ by sending $(B_1, \ldots, B_m)$ to $(AB_1C^{\mathrm{t}}, \ldots, AB_mC^{\mathrm{t}})$ (A^{t} denotes the transpose of A). The group action induces an action on the ring of polynomial functions on $M(n, \mathbb{F})^{\oplus m}$. Let $R(n, m)$ be the ring of those polynomials invariant under this action. $R(n, m)$ is called *the ring of matrix semi-invariants* (for matrices of size $n \times n$) [IQS17a, DM17]. The common zeros of the homogeneous polynomials of positive degrees in $R(n, m)$, denoted as $N(R(n, m))$, is referred to as the *nullcone* of this invariant ring in the invariant theory literature. The link to these matrix spaces which have shrunk subspaces is the following result from invariant theory, proved using the celebrated *Hilbert-Mumford* criterion.

Theorem 4.1 ([BD06, ANS07])**.** *$(B_1, \ldots, B_m) \in M(n, \mathbb{F})^{\oplus m}$ is in $N(R(n, m))$ if and only if $\langle B_1, \ldots, B_m \rangle$ has a shrunk subspace.*

Therefore, matrix spaces with shrunk subspaces are characterized by those polynomials in $R(n, m)$.

Invariant theory also helps to certify those matrix spaces with no shrunk subspaces. Given a matrix space $\mathcal{B} \leq M(n, \mathbb{F})$, if $\mathcal{B}$ does not have a shrunk subspace, we could present a short witness to certify this fact. For example, if $\mathcal{B}$ contains a full-rank matrix B, then exhibiting B is enough to certify that $\mathcal{B}$ does not contain shrunk subspaces. However, as mentioned, it is possible that a matrix space has neither a full-rank matrix nor a shrunk subspace. To resolve this difficulty, we first recall what polynomials in $R(n, m)$ look like. This task is usually resolved in the so-called *first fundamental theorem for $R(n, m)$*.

Theorem 4.2 ([DW00, SvdB01, DZ01, ANS07])**.** *Every nonzero homogeneous polynomial in $R(n, m)$ is of degree kn for some $k \in \mathbb{N}$, and is a linear combination of polynomials of the form $\det(X_1 \otimes A_1 + \cdots + X_m \otimes A_m)$ where the X_is are $n \times n$ variable matrices, and the A_is are $k \times k$ matrices over $\mathbb{F}$.*

Theorem 4.2 motivates the following definition. For a matrix space $\mathcal{B} \leq M(n, \mathbb{F})$, the kth *blow-up* of $\mathcal{B}$ is defined as $\mathcal{B}^{[k]} := \mathcal{B} \otimes M(k, \mathbb{F})$. If $\mathcal{B}$ possesses shrunk subspaces, then $\mathcal{B}^{[k]}$ has shrunk subspaces for any positive integer k. On the other hand, if $\mathcal{B} = \langle B_1, \ldots, B_m \rangle$ has no shrunk subspace, then it is not in the nullcone of $R(n, m)$ (Theorem 4.1). This implies that there exist $k \in \mathbb{N}$ and $A_1, \ldots, A_m \in M(k, \mathbb{F})$, such that $\det(A_1 \otimes B_1 + \cdots + A_m \otimes B_m) \neq 0$ (Theorem 4.2), which simply says that $\mathcal{B}^{[k]}$ contains full-rank matrices. To see that k is finite is classical: by *Hilbert's basis theorem*, $N(R(n, m))$ can be defined by finitely many polynomials, therefore k must be finite. Recently, exciting progress suggests that k can be taken to be no more than $n - 1$ as long as $|\mathbb{F}|$ is large enough [DM17]; see also [IQS17b] for a simpler proof of $k \leq n + 1$. Summarizing the above we have

Theorem 4.3 ([DM17, IQS17b])**.** *Suppose* $|\mathbb{F}| = n^{\Omega(1)}$[4]*. If* $\mathcal{B} \leq M(n, \mathbb{F})$ *does not have shrunk subspace, then for some* $k \leq n + 1$*,* $\mathcal{B} \otimes M(n, \mathbb{F})$ *contains a full-rank matrix.*

The full-rank matrices as in Theorem 4.3 serve as a witness for $\mathcal{B}$ to have no shrunk subspace. We can also easily formulate an algorithmic problem around shrunk subspaces, known as the non-commutative rank problem [FR04] or non-commutative Edmonds' problem in [IQS17a]. That is, given a matrix space $\mathcal{B} \leq M(n, \mathbb{F})$, decide whether $\mathcal{B}$ has a shrunk subspace $U \leq \mathbb{F}^n$. Recent advances imply that this problem can be solved *deterministically* in polynomial time.

Theorem 4.4 ([GGOW16, IQS17b])**.** *There exists a deterministic polynomial-time algorithm to decide whether a given matrix space* $\mathcal{B} \leq M(d, \mathbb{F})$ *has shrunk subspaces or not when* $|\mathbb{F}| = n^{\Omega(1)}$*.*

In this chapter, we first study tripartite-to-bipartite SLOCC entanglement transformation in the finite-copy setting. Our main contribution is to present illustrative examples of tripartite states where their maximal rank is *strictly super-multiplicative*, i.e. $mrk(|\Psi_1\rangle_{\mathrm{ABC}} \otimes |\Psi_2\rangle_{\mathrm{ABC}}) > mrk(|\Psi_1\rangle_{\mathrm{ABC}}) \times mrk(|\Psi_2\rangle_{\mathrm{ABC}})$. Meanwhile, given two copies of the tripartite states in our constructions, they can be converted into two copies of bipartite maximally entangled state via SLOCC, whereas a single copy of such a state cannot. In addition, we present a simple sufficient condition on $|\Psi_1\rangle_{\mathrm{ABC}}$ and $|\Psi_2\rangle_{\mathrm{ABC}}$ such that

[4] $\Omega(1)$ is asymptotic in n.

$mrk(|\Psi_1\rangle_{\mathrm{ABC}} \otimes |\Psi_2\rangle_{\mathrm{ABC}}) > mrk(|\Psi_1\rangle_{\mathrm{ABC}}) \times mrk(|\Psi_2\rangle_{\mathrm{ABC}})$ holds. These results can be easily proved by basic results of matrix spaces. Taking one step further, we investigate the information-theoretic limit of tripartite-to-bipartite SLOCC entanglement transformation, characterized by the *SLOCC entanglement transformation rate* [YGD14, VC15]. The SLOCC entanglement transformation rate between $|\Psi\rangle_{\mathrm{ABC}}$ and $|\psi\rangle_{\mathrm{AB}}$ is the largest ratio $m(n)/n$ when n goes to infinity, where $m(n)$ equals the number of copies of $|\psi\rangle_{\mathrm{AB}}$ one can obtain from $|\Psi\rangle_{\mathrm{ABC}}^{\otimes n}$. Due to the super-multiplicativity of the maximal rank, the asymptotic rate could be extremely difficult to compute. We exploit certain results of the structure of matrix spaces, including the study of matrix semi-invariants, to derive explicit formulas which compute the asymptotic rate for two families of tripartite states (with any bipartite states). Surprisingly, these formulas enable us to establish new connections between certain problems in algebraic complexity theory and asymptotic SLOCC transformation (see [CDS08, VC15] for other connections): determining the asymptotic SLOCC convertibility of a tripartite pure state to the bipartite maximally entangled state is equivalent to determining whether a given matrix space contains shrunk subspaces, the decision version of the non-commutative rank problem. Meanwhile, such an equivalence ensures that there exist deterministic polynomial-time algorithms to determine whether this condition holds for a given tripartite state [GGOW16, IQS17b].

4.2 Notations and Preliminaries

In the rest of this chapter, we focus on tripartite pure states, shared by Alice, Bob and Charlie, in the Hilbert space $\mathcal{H}_n \otimes \mathcal{H}_n \otimes \mathcal{H}_{n'}$, normally denoted by $|\Psi\rangle_{\mathrm{ABC}}$ or $|\Phi\rangle_{\mathrm{ABC}}$. Our aim is to decide the SLOCC convertibility of $|\Psi\rangle_{\mathrm{ABC}}$ to the tensor product of a bipartite pure state $|\psi\rangle_{\mathrm{AB}}$ shared by Alice and Bob, and an arbitrary state (say $|0\rangle_{\mathrm{C}}$) possessed by Charlie. To simplify the notations, we omit Charlie's state and write $|\Psi\rangle_{\mathrm{ABC}} \overset{\mathrm{SLOCC}}{\longrightarrow} |\phi\rangle_{\mathrm{AB}}$ to denote that the aforementioned SLOCC transformation. In particular, we use $|\mathrm{EPR}_n\rangle_{\mathrm{AB}} := \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} |i\rangle_{\mathrm{A}} |i\rangle_{\mathrm{B}}$ to denote the *maximally entangled state* in $\mathcal{H}_n \otimes \mathcal{H}_n$ (shared by Alice and Bob).

The *reduced density operator* of $|\Psi\rangle_{\mathrm{ABC}}$ (shared by Alice and Bob) is defined and denoted as $\rho_{\mathrm{AB}}^{\Psi} := \mathrm{Tr}_{\mathrm{C}} |\Psi\rangle\langle\Psi|_{\mathrm{ABC}}$. The *Choi-Jamiołkowski isomorphism* is the bijective linear map

$J : \mathcal{H}_m \otimes \mathcal{H}_n \to M(m \times n, \mathbb{C})$, defined as $J(|i\rangle \otimes |j\rangle) := |i\rangle\langle j|$ for all $i = 0, \ldots, m-1$ and $j = 0, \ldots, n-1$. In addition, the Choi-Jamiołkowski isomorphism ensures that the Schmidt rank of $|\psi\rangle_{\mathrm{AB}}$ equals $rk(J(|\psi\rangle_{\mathrm{AB}}))$, the *matrix rank* of $J(|\psi\rangle_{\mathrm{AB}})$. The matrix space associated with the tripartite pure state $|\Psi\rangle_{\mathrm{ABC}}$ (with respect to Alice and Bob), denoted by $M(\Psi_{\mathrm{ABC}})$, is obtained by applying the Choi-Jamiołkowski isomorphism to the support of $\rho_{\mathrm{AB}}^{\Psi}$. The maximal rank of $|\Psi\rangle_{\mathrm{ABC}}$ can be then defined and denoted as $mrk(\Psi_{\mathrm{ABC}}) := \max\{rk(B) : B \in M(\Psi_{\mathrm{ABC}})\}$. The tripartite-to-bipartite SLOCC convertibility can be characterized as follows:

Theorem 4.5 (Chitambar, Duan and Shi [CDS10]). $|\Psi\rangle_{\mathrm{ABC}} \overset{\mathrm{SLOCC}}{\longrightarrow} |\psi\rangle_{\mathrm{AB}}$ *if and only if* $mrk(\Psi_{\mathrm{ABC}}) \geq \mathrm{Sch}(\psi_{\mathrm{AB}})$.

The SLOCC protocol for Theorem 4.5, as proposed in [CDS10], takes the following form: Firstly, Charlie makes a measurement on his party and broadcasts the result to Alice and Bob via classical communication; then Alice and Bob propose an SLOCC protocol based on Charlie's results, which converts the state shared by themselves to the desired one. This "one-way" protocol coincide with the one exhibited in the entanglement of assistance [DFM$^+$99, SVW05]. There also exist "two-way" protocols, introduced in the entanglement of collaboration [GS06, Gou06]. This type of protocols allows Alice and Bob make measurements before Charlie make measurements, and broadcast their outcomes to Charlie as well. In the LOCC setting, such "two-way" protocols are necessary for some tripartite-to-bipartite entanglement transformations [GS06]. Here, in the SLOCC setting, we emphasize that the "one-way" protocols are sufficient [CDS10].

Note that Theorem 4.5 can be also applied when multiple copies of the tripartite states are provided. In this circumstances, the number of copies of the desired bipartite states would be our main concern. Taking one step further, it is the *super-multiplicativity* of maximal rank plays the role in the tripartite-to-bipartite SLOCC entanglement transformations. Note that the Schmidt ranks of bipartite states are *multiplicative*, i.e. $\mathrm{Sch}(\psi_1 \otimes \psi_2) = \mathrm{Sch}(\psi_1)\mathrm{Sch}(\psi_2)$ holds for arbitrary $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}_n \otimes \mathcal{H}_n$. For the maximal rank, by definition, we can easily verify that maximal rank is *super-multiplicative*, i.e. $mrk(\Psi_1 \otimes \Psi_2) \geq mrk(\Psi_1)mrk(\Psi_2)$ for $|\Psi_1\rangle, |\Psi_2\rangle \in \mathcal{H}_n \otimes \mathcal{H}_n \otimes \mathcal{H}_{n'}$. Clearly, if it is not strict, then the multiple-copy transformation has no difference with the single-copy transformation.

Otherwise, one could expect that the use of multiple copies can increase the conversion probability of some tripartite-to-bipartite transformations from zero to positive.

From an *information-theoretic* perspective, we can extend the multiple-copy setting to the asymptotic setting, i.e. let the number of available copies goes to infinity. The limitation of the convertibility of given tripartite state and bipartite state is characterized by the *SLOCC entanglement transformation rate* [YGD14, VC15], defined and denoted as

$$R(\Psi_{\mathrm{ABC}}, \psi_{\mathrm{AB}}) := \sup_{N\in\mathbb{N}} \left\{ \frac{1}{N} \max\{M : |\Psi\rangle_{\mathrm{ABC}}^{\otimes N} \overset{\mathrm{SLOCC}}{\longrightarrow} |\psi\rangle_{\mathrm{AB}}^{\otimes M}\} \right\}. \tag{4.2}$$

Note that $\max\{M : |\Psi\rangle_{\mathrm{ABC}}^{\otimes N} \overset{\mathrm{SLOCC}}{\longrightarrow} |\psi\rangle_{\mathrm{AB}}^{\otimes M}\} = \lfloor \log_{\mathrm{Sch}(\psi_{\mathrm{AB}})} mrk(|\Psi\rangle_{\mathrm{ABC}}^{\otimes N}) \rfloor$ for every fixed $N \in \mathbb{N}$. We are motivated to define the *asymptotic maximal rank* of $|\Psi\rangle_{\mathrm{ABC}}$ as

$$mrk^{\infty}(\Psi_{\mathrm{ABC}}) := \sup_{N\in\mathbb{N}} \sqrt[N]{mrk(\Psi_{\mathrm{ABC}}^{\otimes N})}. \tag{4.3}$$

If maximal rank is multiplicative, the asymptotic maximal rank equals the maximal rank itself; otherwise, it might be hard to compute the asymptotic quantities, due to the difficulty caused by the super-multiplicativity.

To make the asymptotic maximal rank more tractable, we point out that taking *supremum* "$\sup_{N\in\mathbb{N}}$" can be replaced by taking *limit* "$\lim_{N\to\infty}$":

Lemma 4.6. *$mrk^{\infty}(\Psi_{\mathrm{ABC}})$ is finite for all tripartite pure states $|\Psi\rangle_{\mathrm{ABC}} \in \mathcal{H}_n \otimes \mathcal{H}_n \otimes \mathcal{H}_{n'}$, and $mrk^{\infty}(\Psi_{\mathrm{ABC}}) := \lim_{N\to\infty} \sqrt[N]{mrk(\Psi_{\mathrm{ABC}}^{\otimes N})}$.*

Proof. We shall utilize the following lemma:

Lemma 4.7 ([BNS98])**.** *Suppose $c_1, c_2, \ldots, c_N, \ldots$ is a nonnegative sequence such that $c_N \leq kN$ for some $k \geq 0$, and $c_M + c_N \leq c_{M+N}$ for all $M, N \in \mathbb{N}$. Then $\lim_{N\to\infty} \frac{c_N}{N}$ exists and is finite.*

Let $c_N = \log_2 mrk(\Psi_{\mathrm{ABC}}^{\otimes N})$. It is easy to see $c_N \leq N \log_2 n$, as $mrk(\Psi_{\mathrm{ABC}}^{\otimes N}) \leq n^N$. Then we can choose $k = \log_2 n$. To prove $c_M + c_N \leq c_{M+N}$, note that $rk(A) \times rk(B) = rk(A \otimes B)$ holds for any matrices A, B. Then it is easy to derive the maximal rank function is super-multiplicative, i.e. $mrk(\Psi_{\mathrm{ABC}} \otimes \Phi_{\mathrm{ABC}}) \geq mrk(\Psi_{\mathrm{ABC}}) mrk(\Phi_{\mathrm{ABC}})$, which leads to $c_M +$

$c_N \leq c_{M+N}$. These ensure that $\lim_{N\to\infty} \frac{1}{N} mrk(\Psi_{\mathrm{ABC}}^{\otimes N})$ exists and is finite by Lemma 4.7. On the other hand, by Fekete's Lemma [Fek23], the condition that $c_M + c_N \leq c_{M+N}$ for all $M, N \in \mathbb{N}$ implies that $\sup_{N\in\mathbb{N}} \frac{c_N}{N} = \lim_{N\to\infty} \frac{c_N}{N}$. This concludes the proof. □

As we will explore maximal rank and asymptotic maximal rank in the rest of this chapter, it would be simpler to deal with matrix spaces rather than tripartite pure states. We introduce some background knowledge of matrix spaces which will be useful in our proofs. Firstly, the maximal rank and asymptotic maximal rank of matrix spaces can be defined in the same manner: For $\mathcal{B} \leq M(n, \mathbb{C})$, $mrk(\mathcal{B}) := \max\{rk(B) : B \in \mathcal{B}\}$ and $mrk^{\infty}(\mathcal{B}) := \sup_{N\in\mathbb{N}} \sqrt[N]{mrk(\mathcal{B}^{\otimes N})} = \lim_{N\to\infty} \sqrt[N]{mrk(\mathcal{B}^{\otimes N})}$. If $mrk(\mathcal{B}) < n$, we say $\mathcal{B}$ is *singular*. Note that maximal rank is invariant under the group action of $\mathrm{GL}(n, \mathbb{C}) \times \mathrm{GL}(n, \mathbb{C})$. Namely, for $(P, Q) \in \mathrm{GL}(n, \mathbb{C}) \times \mathrm{GL}(n, \mathbb{C})$, we have $mrk(\mathcal{B}) = mrk(P\mathcal{B}Q)$, where $P\mathcal{B}Q := \mathrm{span}\{PBQ : B \in \mathcal{B}\}$. We say $\mathcal{B}_1$ and $\mathcal{B}_2$ are equivalent, if $\exists (P, Q) \in \mathrm{GL}(n, \mathbb{C}) \times \mathrm{GL}(n, \mathbb{C})$ such that $\mathcal{B}_1 = P\mathcal{B}_2Q$. The image and kernel of $\mathcal{B} \in M(n, \mathbb{C})$ is defined as $\mathrm{Im}(\mathcal{B}) := \langle \bigcup_{B\in\mathcal{B}} \mathrm{Im}(B) \rangle$ and $\mathrm{Ker}(\mathcal{B}) := \bigcap_{B\in\mathcal{B}} \mathrm{Ker}(B)$. A singular matrix space $\mathcal{B}$ is called *image (kernel)-nondegenerate*, if $mrk(\mathcal{B}) < \dim(\mathrm{Im}(\mathcal{B}))$ ($mrk(\mathcal{B}) < n - \dim(\mathrm{Ker}(\mathcal{B}))$). $\mathcal{B}$ is called *non-degenerate*, if it is both image-degenerate and kernel-degenerate.

A vector space $U \leq \mathbb{C}^n$ is called a *shrunk subspace* of $\mathcal{B} \leq M(n, \mathbb{C})$, if $\dim(\mathcal{B}(U)) = \dim(\langle \bigcup_{B\in\mathcal{B}} B(U) \rangle) < \dim(U)$. Matrix spaces which possess shrunk subspaces are called *shrinking*. As we have mentioned before, this definition is reminiscent of the *shrunk subset* as in the famous *Hall's marriage theorem* [Hal35]. Although shrinking matrix spaces are manifestly singular, there exist matrix spaces, such as the *skew-symmetric matrix space with odd dimension*, which have neither full-rank matrix nor shrunk subspace. Nevertheless, the maximal rank of such matrix spaces cannot be too small, as proved by Fortin and Reutenauer:

Theorem 4.8 ([FR04])**.** *Let $\mathcal{B} \leq M(n, \mathbb{C})$ which has no shrunk subspace. Then $\frac{1}{2}n \leq mrk(\mathcal{B}) \leq n$.*

Shrinking matrix spaces admit a specific structural result. Let U be a shrunk subspace of $\mathcal{B} \leq M(n, \mathbb{C})$, satisfying $\dim(U) = n - q > \dim(\mathcal{B}(U)) = p$ for some $p, q \in \mathbb{N}$. Let

$\{|\alpha_q\rangle, \ldots, |\alpha_{n-1}\rangle\}$ and $\{|\beta_0\rangle, \ldots, |\beta_{p-1}\rangle\}$ be the linear bases of U and $\mathcal{B}(U)$, respectively. Extend them to full bases of $\mathbb{C}^n$, say $\mathbb{C}^n = \langle |\alpha_0\rangle, \ldots, |\alpha_{q-1}\rangle, |\alpha_q\rangle, \ldots, |\alpha_{n-1}\rangle\rangle = \langle |\beta_0\rangle, \ldots, |\beta_{p-1}\rangle, |\beta_p\rangle, \ldots, |\beta_{n-1}\rangle\rangle$. Let $P = \sum_{i=0}^{n-1} |\beta_i\rangle\langle i|$ and $Q = \sum_{j=0}^{n-1} |j\rangle\langle \alpha_j|$, where $\{|0\rangle, \ldots, |n-1\rangle\}$ is the original bases of matrices in $\mathcal{B}$. It is clear that $P, Q \in \mathrm{GL}(n, \mathbb{C})$. Take $\mathcal{B}' = P\mathcal{B}Q$. For $B' \in \mathcal{B}'$, divide B' into the following block form:

$$B' = \left[\begin{array}{c|c} B_{p\times q} & B_{p\times(n-q)} \\ \hline B_{(n-p)\times q} & B_{(n-p)\times(n-q)} \end{array}\right]_{n\times n}, \tag{4.4}$$

where $B_{p\times q} \in \mathrm{span}\{|\beta_i\rangle\langle\alpha_j| : 0 \le i \le p-1,\ 0 \le j \le q-1\}$, $B_{p\times(n-q)} \in \mathrm{span}\{|\beta_i\rangle\langle\alpha_j| : 0 \le i \le p-1,\ q \le j \le n-1\}$, $B_{(n-p)\times q} \in \mathrm{span}\{|\beta_i\rangle\langle\alpha_j| : p \le i \le n-1,\ 0 \le j \le q-1\}$ and $B_{(n-p)\times(n-q)} \in \mathrm{span}\{|\beta_i\rangle\langle\alpha_j| : p \le i \le n-1,\ q \le j \le n-1\}$. Note that $B_{(n-p)\times(n-q)} = 0$, since matrix $B \in \mathcal{B}$ always maps U into a subspace of $\mathcal{B}(U)$. Therefore, $B' \in \mathcal{B}'$ possesses the following form:

$$B' = \left[\begin{array}{c|c} B_{p\times q} & B_{p\times(n-q)} \\ \hline B_{(n-p)\times q} & 0 \end{array}\right]_{n\times n}. \tag{4.5}$$

Generally, every matrix space, of which the matrices therein are of the form 4.5 with parameter $p + q < n$, "compresses" some subspace into a lower-dimension one. These matrix spaces are called *compression matrix spaces* (with parameter $p + q < n$) in the literature [EH88]. Define

$$\mathcal{A}(p, q, n) := \langle \{|i\rangle\langle j| : 0 \le i \le p-1,\ 0 \le j \le n-1\} \bigcup \{|i\rangle\langle j| : p \le i \le n-1,\ 0 \le j \le q-1\} \rangle \tag{4.6}$$

to be the matrix space spanned by the elementary matrices of the first p rows and the first q columns. We call $\mathcal{A}(p, q, n)$ *maximal compression matrix space* if $p + q < n$. We can also define $\mathcal{A}(p, q, m, n) \le M(m \times n, \mathbb{C})$ to be the matrix space spanned by the elementary matrices of the first p rows and the first q columns, and call it maximal compression matrix space if $p + q < \min\{m, n\}$. It is not hard to see every compression matrix space with parameter $p + q < n$ is a subspace of $\mathcal{A}(p, q, n)$. Thus, we have the following:

Lemma 4.9. *Let $\mathcal{B} \le M(n, \mathbb{C})$ with shrunk subspace $U \le \mathbb{C}^n$. Assume $\dim(U) = n - q > \dim(\mathcal{B}(U)) = p$. Then $mrk(\mathcal{B}) < mrk(\mathcal{A}(p, q, n))$.*

4.3 Multi-Copy Transformation

In this section, we provide an affirmative answer to the strict super-multiplicativity of maximal rank. It then reveals that some tripartite-to-bipartite SLOCC entanglement transformation can be achieved by providing multiple copies (of the tripartite states). In this sense, we say $|\Psi\rangle_{\mathrm{ABC}}$ can be converted into $|\psi\rangle_{\mathrm{ABC}}$ by SLOCC with multiple copies, if there exist positive integer $k \geq 2$, such that $|\Psi\rangle_{\mathrm{ABC}}^{\otimes k} \overset{\mathrm{SLOCC}}{\longrightarrow} |\psi\rangle_{\mathrm{AB}}^{\otimes k}$ while $|\Psi\rangle_{\mathrm{ABC}} \overset{\mathrm{SLOCC}}{\nrightarrow} |\psi\rangle_{\mathrm{AB}}$. We first propose the following:

Theorem 4.10. *Let n be an odd number and*

$$|\Psi(n)\rangle_{\mathrm{ABC}} := \sqrt{\frac{2}{n(n-1)}} \sum_{0\leq i<j\leq n-1} (|i\rangle\,|j\rangle - |j\rangle\,|i\rangle)_{\mathrm{AB}} \otimes |ij\rangle_{\mathrm{C}}\,. \tag{4.7}$$

Then we have

$$mrk(\Psi(n)_{\mathrm{ABC}}) = n-1,\ mrk(\Psi(n)_{\mathrm{ABC}}^{\otimes 2}) = n^2 > (n-1)^2. \tag{4.8}$$

Equivalently, $|\Psi(n)\rangle_{\mathrm{ABC}} \overset{\mathrm{SLOCC}}{\nrightarrow} |\mathrm{EPR}_n\rangle_{\mathrm{AB}} \otimes |0\rangle_{\mathrm{C}}$ but $|\Psi(n)\rangle_{\mathrm{ABC}}^{\otimes 2} \overset{\mathrm{SLOCC}}{\longrightarrow} |\mathrm{EPR}_n\rangle_{\mathrm{AB}}^{\otimes 2} \otimes |0\rangle_{\mathrm{C}}^{\otimes 2}$. In fact, any tripartite state $|\Phi\rangle_{\mathrm{ABC}}$ with $M(\Phi_{\mathrm{ABC}}) = \mathrm{span}\{|i\rangle\langle j| - |j\rangle\langle i| : 0 \leq i < j \leq n-1\}$, which is the $n \times n$ skew-symmetric matrix space, satisfies Equation (4.8).

Proof. It is known that for odd n, the maximal rank of the $n \times n$ skew-symmetric matrix space is $n-1$ [FR04]. We will prove $mrk(M(\Psi(n)_{\mathrm{ABC}}^{\otimes 2})) = n^2$ by explicitly construct a full-rank matrix. Let $\{B_{i,j} = |i\rangle\langle j| - |j\rangle\langle i| : 0 \leq i < j \leq n-1\}$ be the linear bases of $M(\Psi(n)_{\mathrm{ABC}})$. We claim that $P := \sum_{0\leq i<j\leq n-1} B_{i,j} \otimes B_{i,j}$ has rank n^2, or equivalently, $\mathrm{Ker}(P) = \{0\}$. Note that P is in the block matrix form:

$$P = \begin{bmatrix} 0 & B_{0,1} & \cdots & B_{0,n-1} \\ -B_{0,1} & 0 & \cdots & B_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ -B_{0,n-1} & -B_{1,n-1} & \cdots & 0 \end{bmatrix}. \tag{4.9}$$

We consider the system of linear equations $P\,|x\rangle = 0$, where $|x\rangle = \sum_{i,j=0}^{n-1} x_i(j)\,|i\rangle\,|j\rangle$ and $x_i(j)$ are unknown variables. Denote $|x_i\rangle = \sum_{j=0}^{n-1} x_i(j)\,|j\rangle$. For $1 \leq k \leq n-2$, we can

rewrite the linear equations with respect to $|x_i\rangle$'s as

$$-\sum_{i=0}^{k-1} B_{i,k}\,|x_i\rangle + \sum_{i=k+1}^{n-1} B_{k,i}\,|x_i\rangle = 0. \tag{4.10}$$

For $k=0$, we have $\sum_{i=1}^{n-1} B_{0,i}\,|x_i\rangle = 0$; for $k=n-1$, we have $\sum_{i=0}^{n-2} B_{i,n-1}\,|x_i\rangle = 0$. Since $B_{j,k}\,|x_i\rangle = x_i(k)\,|j\rangle - x_i(j)\,|k\rangle$, we can derive the following from the above n equations:

$$\begin{gathered}\sum_{i=1}^{n-1}(x_i(0)\,|i\rangle - x_i(i)\,|0\rangle) = 0,\\ -\sum_{i=0}^{k-1}(x_i(k)\,|i\rangle - x_i(i)\,|k\rangle) + \sum_{i=k+1}^{n-1}(x_i(i)\,|k\rangle - x_i(k)\,|i\rangle) = 0, \quad k=1,\ldots,n-2\\ \sum_{i=0}^{n-2}(x_i(n-1)\,|i\rangle - x_i(i)\,|n-1\rangle) = 0.\end{gathered} \tag{4.11}$$

These equations implies $x_i(j) = 0$ for all $0 \le i \ne j \le n-1$ and $\sum_{i\ne k} x_i(i) = 0$ for $k = 0,\ldots,d-1$. Furthermore, we have $\sum_{i\ne k} x_i(i) - \sum_{i\ne k+1} x_i(i) = x_{k+1}(k+1) - x_k(k) = 0$ for $k = 0,\ldots,n-2$, and $\sum_{i\ne n-1} x_i(i) - \sum_{i\ne 0} x_i(i) = x_0(0) - x_{n-1}(n-1) = 0$ for $k = n-1$. These ensures that $x_k(k) = 0$ for $k = 0,\ldots,n-1$. Therefore, $|x\rangle = 0$ is the only solution for $P\,|x\rangle = 0$, which derives $mrk(M(\Psi(n)_{\mathrm{ABC}}^{\otimes 2})) = \mathrm{rank}(P) = n^2$. □

The above theorem implies the existence of multiple-copy tripartite-to-bipartite SLOCC entanglement transformation. For general multipartite states, the existence of multiple-copy SLOCC transformations is settled in [CCD+10]. Moreover, the state $|\Psi(n)\rangle_{\mathrm{ABC}}$ constructed in the above theorem also has the following property: with one single copy, $|\Psi(n)\rangle_{\mathrm{ABC}}$ cannot be transformed to $n\otimes n$ maximally entangled state by SLOCC, but can do so with two copies.

Taking one step further, we can further show that, given two singular matrix spaces, as long as they are nondegenerate, the maximal rank of their tensor product can be strictly larger than the product of their maximal ranks.

Theorem 4.11. *Given two tripartite states* $|\Psi\rangle_{\mathrm{ABC}}, |\Phi\rangle_{\mathrm{ABC}} \in \mathcal{H}_n \otimes \mathcal{H}_n \otimes \mathcal{H}_{n'}$, *let* $\mathcal{B}_1 = M(\Psi_{\mathrm{ABC}}) \le M(n,\mathbb{C})$ *and* $\mathcal{B}_2 = M(\Phi_{\mathrm{ABC}}) \le M(n,\mathbb{C})$. *If both* $\mathcal{B}_1$ *and* $\mathcal{B}_2$ *are singular, then* $mrk(\Psi_{\mathrm{ABC}}\otimes\Phi_{\mathrm{ABC}}) > mrk(\Psi_{\mathrm{ABC}})mrk(\Phi_{\mathrm{ABC}})$ *if*

- $\mathcal{B}_1$ *is image-nondegenerate and* $\mathcal{B}_2$ *is kernel-nondegenerate; or*
- $\mathcal{B}_2$ *is image-nondegenerate and* $\mathcal{B}_1$ *is kernel-nondegenerate.*

Proof. It is equivalent to prove $mrk(\mathcal{B}_1 \otimes \mathcal{B}_2) > mrk(\mathcal{B}_1)mrk(\mathcal{B}_2)$. The following observation from ref. [IKS10] would be useful.

Lemma 4.12 (Lemma 2.2 in ref. [IKS10])**.** *Given two matrices* $X, Y \in M(n, \mathbb{C})$*. If* $Y\mathrm{Ker}(X) \not\leq \mathrm{Im}(X)$*, then* $rk(X+rY) > rk(X)$ *except for at most* $rk(X)+1$ *elements* $r \in \mathbb{C}$*.*

For the necessary part, We focus on the first condition. The second condition holds by replacing $\mathcal{B}_1$ and $\mathcal{B}_2$ with each other. Choose $B_1 \in \mathcal{B}_1$ and $B_2 \in \mathcal{B}_2$ with the highest rank, i.e. $rk(B_1) = mrk(\mathcal{B}_1) < n$ and $rk(B_2) = mrk(\mathcal{B}_2) < n$. Define the following two matrix spaces:

$$\mathcal{X} := \{B \in \mathcal{B}_1 : \mathrm{Im}(B) \leq \mathrm{Im}(B_1)\} \leq \mathcal{B}_1,\ \mathcal{Y} := \{B_2 \in \mathcal{B}_2 : \mathrm{Ker}(B_2) \leq \mathrm{Ker}(B)\} \leq \mathcal{B}_2. \tag{4.12}$$

We claim that $\mathcal{X}$ and $\mathcal{Y}$ are two proper subspaces in $\mathcal{B}_1$ and $\mathcal{B}_2$, respectively. Otherwise, assuming $\mathcal{X} = \mathcal{B}_1$, we have $mrk(\mathcal{B}_1) = rk(B_1) = \dim(\mathrm{Im}(B_1)) = \dim(\mathrm{Im}(\mathcal{B}_1))$, which is a contradiction. If $\mathcal{Y} = \mathcal{B}_2$, we have $\dim(\mathrm{Ker}(\mathcal{B}_2)) = \dim(\mathrm{Ker}(B_2)) = n - rk(B_2) = n - mrk(\mathcal{B}_2)$, which is also a contradiction. We can choose $B_1' \in \mathcal{B}_1$ and $\mathcal{B}_2' \in \mathcal{B}_2$ such that $\mathrm{Im}(B_1') \not\leq \mathrm{Im}(B_1)$ and $\mathrm{Ker}(B_2) \not\leq \mathrm{Ker}(B_2')$.

Then we prove that there exists $r \in \mathbb{C}$ such that $rk(B_1 \otimes B_2 + rB_1' \otimes B_2') > rk(B_1 \otimes B_2)$ by Lemma 4.12. Note that $\mathrm{Ker}(B_1 \otimes B_2) = \mathrm{span}\{\mathrm{Ker}(B_1) \otimes \mathbb{C}^n, \mathbb{C}^n \otimes \mathrm{Ker}(B_2)\}$ and $\mathrm{Im}(B_1 \otimes B_2) = \mathrm{Im}(B_1) \otimes \mathrm{Im}(B_2)$. We only need to show $(B_1' \otimes B_2')(\mathbb{C}^n \otimes \mathrm{Ker}(B_2)) \not\leq \mathrm{Im}(B_1) \otimes \mathrm{Im}(B_2)$. Note that $B_2'\mathrm{Ker}(B_2) \leq \mathrm{Im}(B_2)$, as B_2 has the highest rank in $\mathcal{B}_2$ (by Lemma 4.12). While $B_2'\mathrm{Ker}(B_2) \neq \{0\}$ as $B_2' \notin \mathcal{Y}$. Therefore, we can find a nonzero vector $|u\rangle \in \mathrm{Ker}(B_2)$, such that $0 \neq B_2'|u\rangle \in \mathrm{Im}(B_2)$. On the other hand, since $B_1' \notin \mathcal{X}$, there exists $|v\rangle \in \mathbb{C}^n$ such that $B_1'|v\rangle \notin \mathrm{Im}(B_1)$. Thus, $|v\rangle \otimes |u\rangle \in \mathrm{Ker}(B_1 \otimes B_2)$ and $(B_1' \otimes B_2')|v\rangle \otimes |u\rangle \notin \mathrm{Im}(B_1) \otimes \mathrm{Im}(B_2)$. By Lemma 4.12, there exist $r \in \mathbb{C}$ such that $rk(B_1 \otimes B_2 + rB_1' \otimes B_2') > rk(B_1 \otimes B_2)$, which implies $mrk(\mathcal{B}_1 \otimes \mathcal{B}_2) > mrk(\mathcal{B}_1)mrk(\mathcal{B}_2)$. □

4.4 Asymptotic Transformation

To characterize the asymptotic convertibility of tripartite pure states to bipartite pure states, we need to evaluate the SLOCC entanglement transformation rate. More importantly, we are more interested in what kind of tripartite pure states can be used to obtain the bipartite maximally entangled states by SLOCC asymptotically. To resolve these problems, the main obstacle is the difficulty in computing the asymptotic maximal rank of the tripartite states. In the rest of this section, we derive explicit formulas to compute the asymptotic maximal rank of two large families of tripartite pure states in $\mathcal{H}_n \otimes \mathcal{H}_n \otimes \mathcal{H}_{n'}$, using powerful results from the theory of matrix spaces. Then we utilize these two formulas to answer the second question. Namely, we provide a complete characterization to those tripartite pure states which can be converted to bipartite maximally entangled states by SLOCC, asymptotically. We summarize our results with respect to matrix spaces in the following:

Theorem 4.13. *Given a tripartite pure state* $|\Psi\rangle_{\mathrm{ABC}} \in \mathcal{H}_n \otimes \mathcal{H}_n \otimes \mathcal{H}_{n'}$, *let* $\mathcal{B} = M(\Psi_{\mathrm{ABC}}) \leq M(n, \mathbb{C})$.

1. *If* $\mathcal{B}$ *has no shrunk subspace,* $mrk^{\infty}(\mathcal{B}) = n$;
2. *If* $\mathcal{B} = \mathcal{A}(p, q, n)$ *for some* $p + q < n$,

$$mrk^{\infty}(\mathcal{B}) = n \max\{2^{-D(1-\alpha||p')}, 2^{-D(\alpha||q')}\}, \tag{4.13}$$

where $p' = \frac{p}{n}$, $q' = \frac{q}{n}$, $\alpha = \frac{\log_2(n-q) - \log_2 p}{\log_2((n-p)(n-q)) - \log_2(pq)}$ *and* $D(a||b) := a \log_2 \frac{a}{b} + (1-a) \log_2 \frac{1-a}{1-b}$.

We shall split the proof of Theorem 4.13 into the next two subsections. For simplicity, we shall directly work with matrix spaces. Before this, we explain how these two formulas leads to a complete characterization of asymptotic convertibility to the maximally entangled state:

Theorem 4.14. *A tripartite pure state* $|\Psi\rangle_{\mathrm{ABC}} \in \mathcal{H}_n \otimes \mathcal{H}_n \otimes \mathcal{H}_{n'}$ *can be transformed to* $|\mathrm{EPR}_n\rangle_{\mathrm{AB}}$ *via SLOCC with rate* 1 *if and only if* $M(\Psi_{\mathrm{ABC}})$ *has no shrunk subspace.*

Proof. We shall utilize the two formulas in Theorem 4.13. Note that if $M(\Psi_{\mathrm{ABC}})$ has no shrunk subspace, Theorem 4.13 1 indicates the feasibility of asymptotic convertibility to the maximally entangled state. On the other hand, we show that if $M(\Psi_{\mathrm{ABC}})$ has shrunk subspaces, $mrk^{\infty}(\Psi_{\mathrm{ABC}}) < n$. By Lemma 4.9, we know there exist $p+q<n$ such that $mrk(M(\Psi_{\mathrm{ABC}})) \leq mrk(\mathcal{A}(p,q,n))$. Referring to Equation (4.13), $2^{-D(1-\alpha||p')} < 1$ and $2^{-D(\alpha||q')} < 1$ as $q' < \alpha < 1-p'$ (Lemma 4.22). Therefore, for any $p+q<n$, $mrk^{\infty}(\mathcal{A}(p,q,n)) < n$. This concludes the proof. □

Recall that deciding whether a matrix space contains shrunk subspaces is exactly the non-commutative SDIT problem. Exploiting Theorem 4.4, we have the following:

Corollary 4.15. *There exist deterministic polynomial-time algorithms to determine whether a tripartite pure state* $|\Psi\rangle_{\mathrm{ABC}} \in \mathcal{H}_n \otimes \mathcal{H}_n \otimes \mathcal{H}_{n'}$ *can be transformed to* $|\mathrm{EPR}_n\rangle_{\mathrm{AB}}$ *by SLOCC with rate* 1*, asymptotically.*

4.4.1 Asymptotic Maximal Rank of Matrix Spaces without Shrunk Subspace

We first compute the asymptotic maximal rank of matrix spaces without shrunk subspace. Clearly, if $\mathcal{B} \leq M(n,\mathbb{C})$ possesses full-rank matrices, $mrk^{\infty}(\mathcal{B}) = n$. Thus we can restrict ourselves to singular matrix spaces. Note that, Theorem 4.10 illustrates that two copies of skew-symmetric matrix space $\mathcal{B}_s$ with odd dimension will contain full-rank matrices. In the asymptotic setting, we can derive that the asymptotic maximal rank equals n, as there is a subsequence of the number series $\{\sqrt[N]{mrk(\mathcal{B}_s^{\otimes N})}\}_{N\in\mathbb{N}}$ which converges to n. To prove Theorem 4.13 1, we point out that the property of possessing no shrunk subspace is stable under tensor product.

Lemma 4.16. *If* $\mathcal{B}_1, \mathcal{B}_2 \leq M(n,\mathbb{C})$ *have no shrunk subspace, then* $\mathcal{B}_1 \otimes \mathcal{B}_2$ *have no shrunk subspace.*

Proof. By Theorem 4.1 and Theorem 4.2, there exist finite positive integers k_1, k_2, such that $\mathcal{B}_1 \otimes M(k_1,\mathbb{C})$ and $\mathcal{B}_2 \otimes M(k_2,\mathbb{C})$ are non-singular. Thus $\mathcal{B}_1 \otimes \mathcal{B}_2 \otimes M(k_1k_2,\mathbb{C})$ contains full-rank matrices. By Theorem 4.1 and Theorem 4.2 again, we know $\mathcal{B}_1 \otimes \mathcal{B}_2$ have no shrunk subspace. □

The above lemma ensures that, if $\mathcal{B} \leq M(n, \mathbb{C})$ has no shrunk subspaces, $\mathcal{B}^{\otimes N} \leq M(n^N, \mathbb{C})$ has no shrunk subspaces. By Theorem 4.8, we know $\frac{1}{2}n^N \leq mrk(\mathcal{B}^{\otimes N}) \leq n^N$. This derives $mrk^{\infty}(\mathcal{B}) = n$ as $\lim_{n\to\infty} \sqrt[N]{\frac{1}{2}n^N} = n$.

4.4.2 Asymptotic Maximal Rank of Maximal Compression Matrix Spaces

In this subsection, we deal with maximal compression matrix spaces $\mathcal{A}(p, q, n)$ with parameter $p+q < n$. To obtain the formula in Theorem 4.13 2, we exhibit an explicit formula to compute the maximal rank of $\mathcal{A}(p, q, n)^{\otimes N}$ with respect to p, q, n for arbitrary $N \in \mathbb{N}$. Then we prove that the *regularization* of the rank formula converges to Equation (4.13).

It would be convenient to deal with the *symbolic matrix* of $\mathcal{A}(p, q, m, n)$[5]. The entries of symbolic matrices are filled with 0 and $*$. More precisely, the (i, j)th entry of the symbolic matrix is $*$ if and only if there exist matrices in $\mathcal{A}(p, q, m, n)$ such that the (i, j)th entry of which is non-zero. Due to the structure of $\mathcal{A}(p, q, m, n)$, these $*$s can be chosen as arbitrary complex numbers independently. The rank of a symbolic matrix P, denoted by $rk(P)$, is defined as the largest rank it may achieve when replacing $*$'s with suitable complex numbers. Clearly, $rk(P)$ equals the maximal rank of its corresponding $\mathcal{A}(p, q, n)$. Moreover, it is easy to see that, the symbolic matrix of $\mathcal{A}(p_1, q_1, m_1, n_1) \otimes \cdots \otimes \mathcal{A}(p_N, q_N, m_N, n_N)$ is $P_1 \otimes \cdots \otimes P_N$, where P_k is the symbolic matrix of $\mathcal{A}(p_k, q_k, m_k, n_k)$ for $k = 1, \ldots, N$, and the multiplication rules of $\{0, *\}$ are $0 \times 0 = 0$, $0 \times * = * \times 0 = 0$, $* \times * = *$.

As a warm-up, we first show how to compute the maximal rank of the tensor product of two maximal compression matrix spaces.

Lemma 4.17. *Given two maximal-compression matrix spaces* $\mathcal{B}_1 = \mathcal{A}(p_1, q_1, m_1, n_1)$ *and* $\mathcal{B}_2 = \mathcal{A}(p_2, q_2, m_2, n_2)$ *with* $p_1 + q_1 < \min\{m_1, n_1\}$ *and* $p_2 + q_2 < \min\{m_2, n_2\}$*, we have:*

$$mrk(\mathcal{B}_1 \otimes \mathcal{B}_2) = p_1p_2 + \min\{(n_1 - q_1)q_2, p_1(m_2 - p_2)\} + \min\{(m_1 - p_1)p_2, q_1(n_2 - q_2)\} + q_1q_2. \tag{4.14}$$

[5]The symbolic matrix defined here is simplified from the normal definition, which can be found in [IQS17a, GGOW16]. Moreover, we do not require $p + q < \min\{m, n\}$, as our goal is to use symbolic matrices to simplify the notations when looking for suitable row and column exchanges.

Proof. Note that the symbolic matrix P of $\mathcal{B}_1 \otimes \mathcal{B}_2$ can be divided and indexed into the following form:

$$P = \begin{bmatrix} A_{1,1} & \cdots & A_{1,q_1} & A_{1,q_1+1} & \cdots & A_{1,n_1} \\ \vdots & P_0 & \vdots & \vdots & P_1 & \vdots \\ A_{p_1,1} & \cdots & A_{p_1,q_1} & A_{p_1,q_1+1} & \cdots & A_{p_1,n_1} \\ A_{p_1+1,1} & \cdots & A_{p_1+1,q_1} & & & \\ \vdots & P_2 & \vdots & & 0 & \\ A_{m_1,1} & \cdots & A_{m_1,q_1} & & & \end{bmatrix}_{m_1m_2 \times n_1n_2}, \tag{4.15}$$

where $A_{i,j}$ equals the symbolic matrix of $\mathcal{B}_2$ for all possible i and j, and the lower-right block of total size $(m_1 - p_1)m_2 \times (n_1 - q_1)n_2$ are all zero. Denote the upper-left block of size $p_1m_2 \times q_1n_2$ by P_0; the upper-right block of $p_1n_2 \times (n_1 - q_1)n_2$ by P_1 and the lower-left block of size $(m_1 - p_1)m_2 \times q_1n_2$ by P_2.

Rearranging rows and columns: We will show that, after properly rearranging rows and columns, P_0, P_1 and P_2 become symbolic matrices of maximal compression matrix spaces (and each parameters will be identified). This can be achieved as follows: In P_1, we move all columns with more than p_1p_2 $*$s to the left (by exchanging with the original ones on the left) and move all rows with more than $(n_1 - q_1)q_2$ $*$s to the top. In P_2, we move all rows with more than q_1q_2 $*$s to the top and move all columns with more than $(m_1 - p_1)p_2$ $*$s to the left. These row and column manipulations can be achieved by left and right multiplying with invertible matrices $Q_1 \in M(m_1m_2, \mathbb{C})$ and $Q_2 \in M(n_1n_2, \mathbb{C})$, respectively. More precisely, let $P' = \begin{bmatrix} P'_0 & P'_1 \\ P'_2 & 0 \end{bmatrix}$ be the symbolic (block) matrix of $\mathcal{A}' = Q_1(\mathcal{B}_1 \otimes \mathcal{B}_2)Q_2$. It can be verified that P'_1 is the symbolic matrix of $\mathcal{A}_1 = \mathcal{A}(p_1p_2, (n_1 - q_1)q_2, p_1m_2, (n_1 - q_1)n_2)$ and P'_2 is the symbolic matrix of $\mathcal{A}_2 = \mathcal{A}((m_1 - p_1)p_2, q_1q_2, (m_1 - p_1)m_2, q_1n_2)$. In fact, it is also easy to see that P'_0 is the symbolic matrix of $\mathcal{A}_0 = \mathcal{A}(p_1p_2, q_1q_2, p_1m_2, q_1n_2)$. Note that $\mathcal{A}'$ is equivalent to $\mathcal{B}_1 \otimes \mathcal{B}_2$. We can now focus on evaluating the maximal rank of $\mathcal{A}'$.

Proving $mrk(\mathcal{A}') = mrk(\mathcal{A}_1) + mrk(\mathcal{A}_2)$**:** Note that we can always find $E'' \in \mathcal{A}'$ with $rk(E'') = mrk(\mathcal{A}')$, $F' = \begin{bmatrix} F'_0 & F'_1 \\ F'_2 & 0 \end{bmatrix} \in \mathcal{A}'$ with $rk(F'_1) = mrk(\mathcal{B}'_1)$, and $F'' = \begin{bmatrix} F''_0 & F''_1 \\ F''_2 & 0 \end{bmatrix} \in$

$\mathcal{A}'$ with $rk(F_2'') = mrk(\mathcal{B}_2')$. We claim that there exist $\alpha, \beta, \gamma \in \mathbb{C}$, such that $E' = \alpha E'' + \beta F' + \gamma F'' = \begin{bmatrix} E_0' & E_1' \\ E_2' & 0 \end{bmatrix} \in \mathcal{A}'$ satisfies $rk(E') = mrk(\mathcal{A}')$, $rk(E_1') = mrk(\mathcal{A}_1)$ and $rk(E_2') = mrk(\mathcal{A}_2)$. To see this, consider the matrix $xE'' + yF' + zF''$, where x, y, z are variables. As $rk(E'') = mrk(\mathcal{A}') := r$, there exists at least one $r \times r$ submatrix of E'' with rank r. Let f_1 be the determinant of the corresponding submatrix in $xE'' + yF' + zF''$. f_1 is a nonzero homogeneous polynomial in $\mathbb{C}[x, y, z]$ of degree r. Similarly, let $s = mrk(\mathcal{B}_1')$ and $t = mrk(\mathcal{B}_2')$. Then there exists at least one $s \times s$ ($t \times t$) submatrix of $xE'' + yF' + zF''$ in the upper-right (lower-left) part, such that if we denote its determinant by f_2 (f_3), then f_2 (f_3) is a nonzero homogeneous polynomial in $\mathbb{C}[x, y, z]$ of degree s (t). Since $f = f_1 f_2 f_3$ is also a nonzero polynomial in $\mathbb{C}[x, y, z]$, there exists $(\alpha, \beta, \gamma) \in \mathbb{C}^3$ such that $f(\alpha, \beta, \gamma) \neq 0$. Such (α, β, γ) then translates to our desired conditions for $\alpha E'' + \beta F' + \gamma F''$.

Take such $E' = \begin{bmatrix} E_0' & E_1' \\ E_2' & 0 \end{bmatrix} \in \mathcal{A}'$. Since $p_1 + q_1 < \min\{m_1, n_1\}$ and $p_2 + q_2 < \min\{m_2, n_2\}$, we have $p_1 p_2 < (n_1 - q_1)(n_2 - q_2)$ and $q_1 q_2 < (m_1 - p_1)(m_2 - p_2)$. Then submatrix in the upper-right part of P_1' has full row rank $p_1 p_2$, and the lower-left part of P_2' has full column rank $q_1 q_2$. For any possible choice of $E_0' \in \mathcal{A}_0'$, we can use the upper-right part of P_1' to eliminate the first $p_1 p_2$ rows of P_0' without changing the rank of P'. Similarly, we can use the lower-left part of P_2' to eliminate the first $q_1 q_2$ columns of P_0' without changing the rank of P'. After these row and column operations (which can be achieved by left and right multiplying invertible matrices again), E' is transformed to $\begin{bmatrix} 0 & E_1' \\ E_2' & 0 \end{bmatrix}$, which implies that

$$\begin{aligned} &mrk(\mathcal{A}_1 \otimes \mathcal{A}_2) = rk(E_1') + rk(E_2') = mrk(\mathcal{A}_1) + mrk(\mathcal{A}_2) \\ =&mrk(\mathcal{A}(p_1 p_2, (n_1 - q_1)q_2, p_1 m_2, (n_1 - q_1)n_2)) + mrk(\mathcal{A}((m_1 - p_1)p_2, q_1 q_2, (m_1 - p_1)m_2, q_1 n_2)) \\ =&p_1 p_2 + \min\{(n_1 - q_1)q_2, p_1(m_2 - p_2)\} + \min\{(m_1 - p_1)p_2, q_1(n_2 - q_2)\} + q_1 q_2. \end{aligned} \tag{4.16}$$

□

Let us examine an example to illustrate the above procedure. Consider $\mathcal{B}_1 = \mathcal{B}_2 =$

$\mathcal{A}(1,1,3)$, where the symbolic matrix of $\mathcal{A}(1,1,3)$ is $P = \begin{bmatrix} * & * & * \\ * & 0 & 0 \\ * & 0 & 0 \end{bmatrix}$. It is easy to see $mrk(\mathcal{A}(1,1,3)) = 2$. Following the first step in Lemma 4.17, we exchange rows and columns to obtain an equivalent matrix space $\mathcal{A}'$ of $\mathcal{A}(1,1,3)^{\otimes 2}$. This can be done by choosing $Q = |00\rangle\langle 00| + |01\rangle\langle 01| + |02\rangle\langle 02| + |10\rangle\langle 10| + |11\rangle\langle 20| + |12\rangle\langle 12| + |20\rangle\langle 11| + |21\rangle\langle 21| + |22\rangle\langle 22|$ and let $\mathcal{A}' = Q\mathcal{A}(1,1,3)^{\otimes 2}Q$. More precisely, Multiplying Q with $P^{\otimes 2}$ form the left exchanges the 5th row with the 7th row of $P^{\otimes 2}$; multiplying Q with $P^{\otimes 2}$ from the right exchanges the 5th column with the 7th column.

$$P^{\otimes 2} = \begin{bmatrix} * & * & * & * & * & * & * & * & * \\ * & 0 & 0 & * & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & * & 0 & 0 & * & 0 & 0 \\ * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad QP^{\otimes 2}Q = \begin{bmatrix} * & * & * & * & * & * & * & * & * \\ * & 0 & 0 & * & * & 0 & 0 & 0 & 0 \\ * & 0 & 0 & * & * & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (4.17)$$

Let P_0' be the submatrix of size 3×3 in the upper-left corner (of $QP^{\otimes 2}Q$); P_1 be its submatrix of size 3×6 in the upper-right corner and P_2 be its submatrix of size 6×3 in the lower-left corner. It is easy to see that P_0 is the symbolic matrix of $\mathcal{A}(1,1,3)$, P_1 is the symbolic matrix of $\mathcal{A}(1,2,3,6)$, and P_2 is the symbolic matrix of $\mathcal{A}(2,1,6,3)$. We can compute $mrk(\mathcal{A}(1,1,3,3)^{\otimes 2}) = 1+2+2+1 = 6$ by looking at the form of P', which coincide with the one computed by Lemma 4.17.

This example also indicates that the smaller matrix spaces after dividing P', e.g. $\mathcal{A}(1,2,3,6)$ and $\mathcal{A}(2,1,6,3)$, may not be maximal-compression matrix spaces any more, as $1+2 = \min\{3,6\}$. (Recall that $\mathcal{A}(p,q,m,n)$ is a maximal-compression matrix space if $p+q < \min\{m,n\}$.) Thus, we cannot apply Lemma 4.17 recursively to capture a general formula to compute the maximal rank of $\mathcal{A}(p,q,m,n)^{\otimes N}$. Fortunately, for $m = n$, we have the following:

Lemma 4.18. *Given a maximal-compression matrix space $\mathcal{A}(p,q,n)$ and $N \in \mathbb{N}$, the maximal rank of $\mathcal{A}(p,q,n)^{\otimes N+1}$ equals*

$$\sum_{k=0}^{N} \binom{N}{k} \left(\min\{p^{N-k+1}(n-p)^k, q^k(n-q)^{N-k+1}\} + \min\{q^{k+1}(n-q)^{N-k}, p^{N-k}(n-p)^{k+1}\} \right). \tag{4.18}$$

Proof. The proof idea is as follows: First, we use induction to show that the symbolic matrix of $\mathcal{A}(p,q,n)^{\otimes N}$ can be transformed into an upper-anti-block-diagonal form, by appropriate row and column rearrangements. Note that all these block (symbolic) matrices either equals zero, or corresponds to $\mathcal{A}(p,q,m,n)$ with different parameters. Second, we explicitly compute the maximal rank of those anti-diagonal $\mathcal{A}(p,q,m,n)$s. Combining these two observations, we apply the similar techniques in Lemma 4.17 to obtain Equation (4.18).

For the first step, we will illustrate the following observation:

observation 4.19. *For $N \geq 1$, there exist invertible matrices $Q_1 \in M(n^N, \mathbb{C})$ and $Q_2 \in M(n^N, \mathbb{C})$, such that the symbolic matrix P' of $\mathcal{A}' = Q_1\mathcal{A}(p,q,n)^{\otimes N}Q_2$ is of upper-anti-block-diagonal form:*

$$P' = \begin{bmatrix} P_{0,2^{N-1}-1} & \cdots & P_{0,l} & \cdots & P_0 \\ \vdots & \iddots & \iddots & \iddots & 0 \\ P_{l,2^{N-1}-1} & \iddots & P_l & \iddots & 0 \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ P_{2^{N-1}-1} & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{4.19}$$

1. *For the anti-diagonal block matrices, label them as $P_0, \ldots, P_{2^{N-1}-1}$. Let $h(l)$ be the hamming weight of $l \in \{0, \ldots, 2^{N-1}-1\}$, i.e. the number of 1's in the binary expansion of l. Then P_l is the symbolic matrix of*

$$\mathcal{B}_l = \mathcal{A}(p^{N-h(l)}(n-p)^{h(l)}, q^{h(l)+1}(n-q)^{N-h(l)-1}, p^{N-h(l)-1}(n-p)^{h(l)}n, q^{h(l)}(n-q)^{N-h(l)-1}n). \tag{4.20}$$

2. *For the upper-left block matrices, label them as $P_{u,v}$ for $u, v \in \{0, 2^{N-1}-1\}$, where u is the label of the anti-diagonal block matrix P_u on the right of $P_{u,v}$ and v is the*

label of anti-diagonal block matrix P_v below $P_{u,v}$. If $h(u) \geq h(v)$, $P_{u,v} = 0$; otherwise $P_{u,v}$ is the symbolic matrix of

$$\mathcal{B}_{u,v} = \mathcal{A}(p^{N-h(u)}(n-p)^{h(u)}, q^{h(v)+1}(n-q)^{N-h(v)-1}, p^{N-h(u)-1}(n-p)^{h(u)}n, q^{h(v)}(n-q)^{N-h(v)-1}n). \tag{4.21}$$

Proof of Observation 4.19. We show the observation holds by induction on N. It holds for $N = 1$ trivially. Assume for $\mathcal{A}(p,q,n)^{\otimes N}$, observation 4.19 holds. Without loss of generality we assume $\mathcal{A}(p,q,n)^{\otimes N}$ is of the form in Equation (4.19). For, $\mathcal{A}(p,q,n)^{\otimes N+1} = \mathcal{A}(p,q,n)^{\otimes N} \otimes \mathcal{A}(p,q,n)$, let P_l be the symbolic matrix of the lth anti-diagonal block. It is sufficient to examine $P_l \otimes P$, where P is the symbolic matrix of $\mathcal{A}(p,q,n)$. Following the first step in Lemma 4.17, there exist two invertible matrices Q_1^l and Q_2^l, such that $Q_1^l P_l \otimes P Q_2^l = \begin{bmatrix} P_0^l & P_1^l \\ P_2^l & 0 \end{bmatrix}$, where P_1^l is the symbolic matrix of $\mathcal{B}_{l0} = \mathcal{A}(p^{N-h(l)+1}(n-p)^{h(l)}, q^{h(l)+1}(n-q)^{N-h(l)}, p^{N-h(l)}(n-p)^{h(l)}n, q^{h(l)}(n-q)^{N-h(l)}n)$, and P_2^l is the symbolic matrix of $\mathcal{B}_{l1} = \mathcal{A}(p^{N-h(l)}(n-p)^{h(l)+1}, q^{h(l)+2}(n-q)^{N-h(l)-1}, p^{N-h(l)-1}(n-p)^{h(l)+1}n, q^{h(l)+1}(n-q)^{N-h(l)-1}n)$, where $l0$ and $l1$ denote the N-bit strings in which the first $(N-1)$-bit strings equal the binary expansion of l. Moreover, we observe that $\mathcal{B}_{l0}$ and $\mathcal{B}_{l1}$ remain as the anti-diagonal blocks in $\overline{Q_1^l}\mathcal{A}(p,q,n)^{\otimes N+1}\overline{Q_2^l}$ ($\overline{Q_i^l}$ is the enlarged matrix of Q_i^l for $i = 1, 2$). Then the first fact in observation 4.19 follows since $h(l0) = h(l)$ and $h(l1) = h(l) + 1$.

For the second fact, for given $u, v \in \{0, \ldots, 2^{N-1}-1\}$, $u \neq v$ and $h(u) < h(v)$, we examine $P_{u,v} \otimes P$ where $P_{u,v}$ is the symbolic matrix of $\mathcal{B}_{u,v} = \mathcal{A}(p^{N-h(u)}(n-p)^{h(u)}, q^{h(v)+1}(n-q)^{N-h(v)-1}, p^{N-h(u)-1}(n-p)^{h(u)}n, q^{h(v)}(n-q)^{N-h(v)-1}n)$. Note that $P_{u,v}$ has the same "full" rows as that of P_u and has the same "full" columns as that of P_v. Here a "full" row (column) means the corresponding row (column) contains $*$ only. Denote the invertible matrix being responsible for the row rearrangements of $P_u \otimes P$ by Q_1^u and the invertible matrix being responsible for the column rearrangements of $P_v \otimes P$ by Q_2^v. These two matrices will also rearrange the rows and columns of $P_{u,v} \otimes P$, respectively. For simplicity, denote $\mathcal{B}_u = \mathcal{A}(p_1, q_1, m_1, n_1)$ and $\mathcal{B}_v = \mathcal{A}(p_2, q_2, m_2, n_2)$, then $\mathcal{B}_{u,v} = \mathcal{A}(p_1, q_2, m_1, n_2)$.

Write $P_{u,v}$ in the block matrix form

$$P_{u,v} = \begin{bmatrix} A_{1,1} & \cdots & A_{1,q_2} & A_{1,q_2+1} & \cdots & A_{1,n_2} \\ \vdots & P_0^{u,v} & \vdots & \vdots & P_1^{u,v} & \vdots \\ A_{p_1,1} & \cdots & A_{p_1,q_2} & A_{p_1,q_2+1} & \cdots & A_{p_1,n_2} \\ A_{p_1+1,1} & \cdots & A_{p_1+1,q_2} & & & \\ \vdots & P_2^{u,v} & \vdots & & 0 & \\ A_{m_1,1} & \cdots & A_{m_1,q_2} & & & \end{bmatrix}, \tag{4.22}$$

where $A_{i,j}$ are symbolic matrix of $\mathcal{A}(p,q,n)$ for all possible i and j. Q_1^u is responsible for moving all rows with more than $(n_2-q_2)q$ $*$s in $P_1^{u,v}$ and all rows with more than q_2q $*$s in $P_2^{u,v}$ to the top of them. To see this, note that all rows with more than $(n_2-q_2)q$ $*$s in $P_1^{u,v}$ is determined by those "full" rows in $P_{u,v}$, which are exact those rows in P_u; all those rows with more than q_2q $*$s in $P_2^{u,v}$ is determined by those "full" rows in P. Thus coincide with those rows in P_u. Similarly, Q_2^v is responsible to move all columns with more than p_1p $*$s in $P_1^{u,v}$ and all columns with more than $(m_1-p_1)p$ $*$s in $P_2^{u,v}$ to the left. Denote $P'_{u,v} = \begin{bmatrix} P_0^{u,v\prime} & P_1^{u,v\prime} \\ P_2^{u,v\prime} & 0 \end{bmatrix}$ be the symbolic (block) matrix after the aforementioned row and column rearrangement of $P_{u,v}$. We can then conclude that $P_1^{u,v\prime}$ is the symbolic matrix of $\mathcal{A}(p^{N-h(u)+1}(d-p)^{h(u)}, q^{h(v)+1}(d-q)^{N-h(v)}, p^{N-h(u)}(d-p)^{h(u)}d, q^{h(v)}(d-q)^{N-h(v)}d)$, and $P_2^{u,v\prime}$ is the symbolic matrix of $\mathcal{A}(p^{N-h(u)}(d-p)^{h(u)+1}, q^{h(v)+2}(d-q)^{N-h(v)-1}, p^{N-h(u)-1}(d-p)^{h(u)+1}d, q^{h(v)+1}(d-q)^{N-h(v)-1}d)$. In addition, we can verify that $P_0^{u,v\prime}$ is the symbolic matrix of $\mathcal{A}(p^{N-h(u)+1}(d-p)^{h(u)}, q^{h(v)+2}(d-q)^{N-h(v)-1}, p^{N-h(u)}(d-p)^{h(u)}d, q^{h(v)+1}(d-q)^{N-h(v)-1}d)$. In addition, $P_0^{u,v\prime}$, $P_1^{u,v\prime}$ and $P_2^{u,v\prime}$ in the symbolic matrix of $\mathcal{A}(p,q,n)^{\otimes N+1}$, after applying all these row and column rearrangements, will be relabeled as $P_{u0,v1}$, $P_{u0,v0}$ $P_{u1,v1}$ according to their corresponding anti-block terms (In this case, $P_{u1,v0}$ corresponds to the lower-right block, which is 0).

To see the second statement in observation 4.19 holds for $N+1$, we only need to show that if $h(u) \geq h(v)$, $P_{u,v} = 0$. For $u, v \in \{0, \ldots, 2^N-1\}$ and $u \neq v$, let $u = u'b$ and $v = v'c$, where $u', v' \in \{0, \ldots, 2^N-1\}$ equal the first $(N-1)$-bit strings of the binary expansion of u and v, and $b, c \in \{0,1\}$ are variables. If $h(u) \geq h(v)$ derives that either $h(u'0) \geq h(v'0)$, $h(u'0) \geq h(v'1) = h(v') + 1$ or $h(u'1) \geq h(v'1)$, it will imply that $h(u') \geq h(v')$. By

the induction hypothesis, $P_{u',v'} = 0$ and $P_{u,v} = 0$ holds automatically. Otherwise, if $h(u) \geq h(v)$ derives $h(u'1) \geq h(v'0)$ and $P_{u',v'}$ is nonzero, we can also observe that $P_{u'1,v'0}$ equals 0, as it is the lower-right part of $P'_{u,v}$. This concludes the proof. □

Now we focus on those anti-diagonal blocks. The following observation explicitly computes the maximal rank of $\mathcal{A}_l$ for $l = 0, \ldots, 2^{N-1} - 1$:

observation 4.20. *Let $h(l) = k$, the maximal rank of $\mathcal{B}_l = \mathcal{A}(p^{N-k+1}(n-p)^k, q^{k+1}(n-q)^{N-k}, p^{N-k}(n-p)^k n, q^k(n-q)^{N-k}n)$ equals*

$$\min\{p^{N-k+1}(n-p)^k, q^k(n-q)^{N-k+1}\} + \min\{q^{k+1}(n-q)^{N-k}, p^{N-k}(n-p)^{k+1}\}. \quad (4.23)$$

Proof. Note that the rank of $\mathcal{A}(p^{N-k+1}(n-p)^k, q^{k+1}(n-q)^{N-k}, p^{N-k}(n-p)^k n, q^k(n-q)^{N-k}n)$ equals

$$\min\{p^{N-k+1}(n-p)^k + q^{k+1}(n-q)^{N-k}, p^{N-k}(n-p)^k n, q^k(n-q)^{N-k}n\}. \quad (4.24)$$

If $p^{N-k}(n-p)^k \leq q^k(n-q)^{N-k}$, we only need to compare $p^{N-k+1}(n-p)^k + q^{k+1}(n-q)^{N-k}$ and $p^{N-k}(n-p)^k n$. Note that

$$p^{N-k}(n-p)^k n - (p^{N-k+1}(n-p)^k + q^{k+1}(n-q)^{N-k}) = p^{N-k}(n-p)^{k+1} - q^{k+1}(n-q)^{N-k}. \quad (4.25)$$

We further distinguish two cases. When $p^{N-k}(n-p)^{k+1} \geq q^{k+1}(n-q)^{N-k}$, we take $p^{N-k+1}(n-p)^k + q^{k+1}(n-q)^{N-k}$. When $p^{N-k}(n-p)^{k+1} < q^{k+1}(n-q)^{N-k}$, we take $p^{N-k}(n-p)^k n = p^{N-k+1}(n-p)^k + p^{N-k}(n-p)^{k+1}$. These two cases then can be unified in the following equation

$$\begin{aligned} &\min\{p^{N-k+1}(n-p)^k + q^{k+1}(n-q)^{N-k}, p^{N-k}(n-p)^k n, q^k(n-q)^{N-k}n\} \\ =&p^{N-k+1}(n-p)^k + \min\{q^{k+1}(n-q)^{N-k}, p^{N-k}(n-p)^{k+1}\}. \end{aligned} \quad (4.26)$$

Similarly, if $p^{N-k}(n-p)^k > q^k(n-q)^{N-k}$, we obtain

$$\begin{aligned} &\min\{p^{N-k+1}(n-p)^k + q^{k+1}(n-q)^{N-k}, p^{N-k}(n-p)^k n, q^k(n-q)^{N-k}n\} \\ =&\min\{p^{N-k+1}(n-p)^k, q^k(n-q)^{N-k+1}\} + q^{k+1}(n-q)^{N-k}. \end{aligned} \quad (4.27)$$

Note that, $p^{N-k}(n-p)^k \leq q^k(n-q)^{N-k}$ implies $p^{N-k+1}(n-p)^k \leq pq^k(n-q)^{N-k} < q^k(n-q)^{N-k+1}$, where the second inequality uses $p+q<n$ as $\mathcal{A}(p,q,n)$ is maximal-compression. Similarly, $p^{N-k}(n-p)^k > q^k(n-q)^{N-k}$ implies $q^{k+1}(n-q)^{N-k} < qp^{N-k}(n-p)^k < p^{N-k}(n-p)^{k+1}$. This observation allows us to combine Equation (4.26) and Equation (4.27) to obtain

$$mrk(\mathcal{B}_l) = \min\{p^{N-k+1}(n-p)^k, q^k(n-q)^{N-k+1}\} + \min\{q^{k+1}(n-q)^{N-k}, p^{N-k}(n-p)^{k+1}\}. \tag{4.28}$$

□

Finally, we combine observations 4.19 and 4.20 to prove that $mrk(\mathcal{A}(p,q,n)^{\otimes N+1}) = \sum_{l=0}^{2^N-1} mrk(\mathcal{B}_l)$. Then Equation (4.18) follows. Without loss of generality we assume the symbolic matrix of $\mathcal{A}(p,q,n)^{\otimes N+1}$ equals the one in Equation (4.9). Adapting the similar argument in the second step in the proof of Lemma 4.17. We can find $B \in \mathcal{A}(p,q,n)^{\otimes N+1}$ with $rank(B) = mrk(\mathcal{A}')$, and if write B into the upper-anti-block-diagonal form as shown in Equation (4.9), we can assume $rk(B_l) = rk(P_l)$ for $0 \leq l \leq 2^N-1$, where B_l is the block matrix corresponding to the symbolic matrix P_l in Equation (4.9) with same size and location. Let $\lambda = \log_2 \frac{n-p}{q}$, $\mu = \log_2 \frac{n-q}{p}$, $\alpha = \frac{\mu}{\lambda+\mu}$. We can derive

$$k \leq \lfloor \alpha N + \alpha - 1 \rfloor \Leftrightarrow p^{N-k}(n-p)^{k+1} \leq q^{k+1}(n-q)^{N-k} \tag{4.29}$$

and

$$k \leq \lfloor \alpha N + \alpha \rfloor \Leftrightarrow p^{N-k+1}(n-p)^k \leq q^k(n-q)^{N-k+1}. \tag{4.30}$$

Choose $N' = \lfloor \alpha N + \alpha \rfloor = \lfloor \alpha N + \alpha - 1 \rfloor + 1$. For any $l \in \{l : h(l) \leq N'-1\}$, B_l can be chosen to have full row rank. For any $l \in \{l : h(l) \geq N'+1\}$, B_l can be chosen to have full column rank. Now we claim that, for $B_{u,v}$, where $u \neq v$ and $u, v \in \{2^N-1\}$, we can use the anti-block matrices B_u and B_v to eliminate $B_{u,v}$. By observation 4.19, we only need to consider those $P_{u,v}$ satisfying $h(u) < h(v)$. In this case, either $h(u) \leq N'-1$, or $h(v) \geq N'+1$. If $h(u) \leq N'-1$, we can use B_u to clear $B_{u,v}$, since B_u has full row rank. The other case is similar. These yield that $mrk(\mathcal{A}(p,q,n)^{\otimes N+1}) = \sum_{l=0}^{2^N-1} mrk(\mathcal{B}_l)$, which, together with Equation (4.28), allow us to conclude the proof. □

Now we are ready to compute the asymptotic maximal rank for maximal-compression matrix spaces. We restate Theorem 4.13 2 here:

Theorem 4.13 2, restated. Let $\mathcal{B} = \mathcal{A}(p, q, n)$ where $p + q < n$.

$$mrk^{\infty}(\mathcal{B}) = n \max\{2^{-D(1-\alpha||p')}, 2^{-D(\alpha||q')}\}, \tag{4.31}$$

where $p' = \frac{p}{n}$, $q' = \frac{q}{n}$, $\alpha = \frac{\log_2(n-q)-\log_2 p}{\log_2((n-p)(n-q))-\log_2(pq)}$ and $D(a||b) := a\log_2\frac{a}{b} + (1-a)\log_2\frac{1-a}{1-b}$.

Proof. Let $\lambda = \log_2\frac{n-p}{q}$, $\mu = \log_2\frac{n-q}{p}$, $\alpha = \frac{\log_2(n-q)-\log_2 p}{\log_2((n-p)(n-q))-\log_2(pq)} = \frac{\mu}{\lambda+\mu}$ and $N' = \lfloor \alpha N + \alpha \rfloor$ as discussed in Lemma 4.18. We can rewrite Equation (4.18) explicitly as the following:

$$\begin{aligned} rk(\mathcal{A}(p,q,n)^{\otimes(N+1)}) &= \sum_{k=0}^{N'-1}\binom{N}{k}p^{N-k}(n-p)^k n + \sum_{k=N'+1}^{N}\binom{N}{k}q^k(n-q)^{N-k}n \\ &+ \binom{N}{N'}\left(p^{N-N'+1}(n-p)^{N'} + q^{N'+1}(n-q)^{N-N'}\right) \\ &= \sum_{k=0}^{N'}\binom{N}{k}p^{N-k}(n-p)^k n + \sum_{k=N'}^{N}\binom{N}{k}q^k(n-q)^{N-k}n \\ &- \binom{N}{N'}\left(p^{N-N'}(n-p)^{N'+1} + q^{N'}(n-q)^{N-N'+1}\right). \end{aligned} \tag{4.32}$$

Let $p' = \frac{p}{n}$ and $q' = \frac{q}{n}$, we have $p' + q' < 1$. The above quantity is upper and lower bounded by

$$rk(\mathcal{A}(p,q,n)^{\otimes(N+1)}) \leq n^{N+1}\Big(\sum_{k=0}^{N'}\binom{N}{k}{p'}^{N-k}(1-p')^k + \sum_{k=0}^{N-N'}\binom{N}{k}{q'}^{N-k}(1-q')^k\Big); \tag{4.33}$$

$$rk(\mathcal{A}(p,q,n)^{\otimes(N+1)}) \geq n^{N+1}\Big(\sum_{k=0}^{N'-1}\binom{N}{k}{p'}^{N-k}(1-p')^k + \sum_{k=0}^{N-N'-1}\binom{N}{k}{q'}^{N-k}(1-q')^k\Big). \tag{4.34}$$

We shall use the following inequalities:

Lemma 4.21 (Lemma 4.7.2 in Ref. [Ash90])**.** *For $N' < Np$, we have:*

$$\frac{1}{\sqrt{2N}}2^{-ND(\frac{N'}{N}||p)} \leq \sum_{k=0}^{N'}\binom{N}{k}p^k(1-p)^{N-k} \leq 2^{-ND(\frac{N'}{N}||p)}. \tag{4.35}$$

To apply Lemma 4.21 to prove Equation (4.13), we need $Nq' < N' < N(1-p')$ holds for sufficiently large N. We first prove the following:

Lemma 4.22. *Let p', q' and α be defined as above. We have $q' < \alpha < 1 - p'$.*

Proof. By expressing α explicitly with respect to p' and q', we need to prove

$$q' < \frac{\log_2 \frac{1-q'}{p'}}{\log_2 \frac{1-q'}{p'} + \log_2 \frac{1-p'}{q'}}, \quad 1 - p' > \frac{\log_2 \frac{1-q'}{p'}}{\log_2 \frac{1-q'}{p'} + \log_2 \frac{1-p'}{q'}}. \tag{4.36}$$

This is equivalent to show

$$(1-q')^{1-q'} q'^{q'} > p'^{1-q'} (1-p')^{q'}, \quad (1-p')^{1-p'} p'^{p'} > (1-q')^{p'} q'^{1-p'}. \tag{4.37}$$

Consider the function $f(x,y) = x^y (1-x)^{1-y}$ with $x, y \in (0,1)$. The partial derivative in x is

$$\frac{\partial}{\partial x} f(x,y) = \frac{x^{y-1}(y-x)}{(1-x)^y}. \tag{4.38}$$

For any fixed y, $\max_{x \in (0,1)} f(x,y) = f(y,y)$. Then inequality (4.37) holds by choosing $x = 1-p', y = q'$ and $x = 1-q', y = p'$. □

Recall $N' = \lfloor \alpha N + \alpha \rfloor$. To ensure that $Nq' < N' < N(1-p')$, it is sufficient to satisfy that $\alpha + \frac{\alpha}{N} < 1 - p'$ and $q' < \alpha - \frac{1-\alpha}{N}$. Since α, p', and q' are fixed, these can be achieved as long as $N > \max\{\frac{\alpha}{1-p'-\alpha}, \frac{1-\alpha}{\alpha-q'}\} > 0$. Now, applying the upper bound in Lemma 4.21 to inequality (4.33), we obtain

$$rk(\mathcal{A}(p,q,n)^{\otimes(N+1)}) \leq n^{N+1} (2^{-ND(\frac{N'}{N}||1-p')} + 2^{-ND(1-\frac{N'}{N}||1-q')}). \tag{4.39}$$

Note that $-D(a||p)$ is increasing for $0 < a < p$, and $\alpha(N+1) - 1 \leq \lfloor \alpha(N+1) \rfloor \leq \alpha(N+1)$. We can replace $\frac{N'}{N}$ by $\alpha + \frac{\alpha}{N}$, and $1 - \frac{N'}{N}$ by $1 - \alpha + \frac{1-\alpha}{N}$, which gives that

$$rk(\mathcal{A}(p,q,n)^{\otimes(N+1)}) \leq n^{N+1} (2^{-ND(\alpha+\frac{\alpha}{N}||1-p')} + 2^{-ND(1-\alpha+\frac{1-\alpha}{N}||1-q')}). \tag{4.40}$$

Let N go to infinity. Since L^p norm converges L^∞ norm when $p \to +\infty$, we have

$$rk^\infty(\mathcal{A}(p,q,n)) \leq n \max\{2^{-D(\alpha||1-p')}, 2^{-D(1-\alpha||1-q')}\}. \tag{4.41}$$

Similarly, applying the lower bound in Lemma 4.21 to inequality (4.34), we obtain

$$\begin{aligned} rk(\mathcal{A}(p,q,n)^{\otimes(N+1)}) &\geq \frac{n^{N+1}}{(N+1)^2}(2^{-ND(\frac{N'-1}{N}||1-p')} + 2^{-ND(1-\frac{N'-1}{N}||1-q')}) \\ &\geq \frac{d^{N+1}}{(N+1)^2}(2^{-ND(\alpha+\frac{\alpha-2}{N}||1-p')} + 2^{-ND(1-\alpha+\frac{1-\alpha}{N}||1-q')}). \end{aligned} \tag{4.42}$$

The second inequality holds since $\frac{N'-1}{N} \geq \alpha + \frac{\alpha-2}{N}$ and $1 - \frac{N'-1}{N} \geq 1 - \alpha + \frac{1-\alpha}{N}$. Thus we have

$$rk^{\infty}(\mathcal{A}(p,q,n)) \geq n \max\{2^{-D(\alpha||1-p')}, 2^{-D(1-\alpha||1-q')}\}. \tag{4.43}$$

Since $D(a||b) = D(1-a||1-b)$, combining inequalities (4.41) and (4.43), we have

$$rk^{\infty}(\mathcal{A}(p,q,n)) = n \max\{2^{-D(1-\alpha||p')}, 2^{-D(\alpha||q')}\}. \tag{4.44}$$

□

4.5 Summary and Discussion

In this chapter, we have systematically studied the tripartite-to-bipartite SLOCC entanglement transformations in multiple-copy and asymptotic settings. We have constructed tripartite pure states which cannot be transformed to the bipartite maximally entangled state by SLOCC with a single copy, but can do so with two copies. Such an interesting phenomenon not only reveals that maximal rank can be strictly super-multiplicative, but also illustrates the existence of multiple-copy SLOCC entanglement transformations in the tripartite-to-bipartite setting. Meanwhile, we have exhibited a simple condition which can be used to construct and verify tripartite states whose maximal ranks are strictly super-multiplicative. This condition also implies that, except for the degenerated case, the strict super-multiplicativity holds for most tripartite states of which their maximal ranks are not full.

In the asymptotic setting, we have derived explicit formulas to compute the tripartite-to-bipartite entanglement transformation rate of two families of tripartite states with respect to their asymptotic maximal ranks. Surprisingly, these formulas lead to a complete

characterization of the asymptotic convertibility of tripartite pure state and the bipartite maximally entangled state, i.e. a tripartite pure states can be transformed to the bipartite maximally entangled state by SLOCC in the asymptotic setting if and only if its corresponding matrix space contains no shrunk subspace. Note that the latter problem is known as the non-commutative SDIT problem. Furthermore, based on the recent progress on the non-commutative rank problem [GGOW16, IQS17b], there exist deterministic polynomial-time algorithms to decide whether a tripartite state can be transformed to the maximally entangled state by SLOCC, asymptotically. Most prominently, our characterization not only provide another connection between certain problems in *algebraic complexity theory* and questions regarding asymptotic SLOCC entanglement transformations, but also provide another relation of the commutative and non-commutative SDIT problem.

Through our investigation, powerful results from the theory of matrix spaces serve as our main tools. As we have shown that the structure of matrix spaces is crucial in the study of tripartite-to-bipartite SLOCC entanglement transformations. Advanced results, such as matrix semi-invariants, can also be exploited to derive beautiful results. In particular, we would like to further discuss the postulate: *"Matrix spaces can be viewed and studied as a linear algebraic analog of bipartite graphs."* Although this viewpoint is not original from our work, it enables us to import *combinatorial* ideas for graphs into the study of *algebraic* structures of matrix spaces. As we have mentioned before, the existence of full-rank matrices in a given matrix space can be viewed as the "bipartite perfect matching" in the linear algebraic analog of bipartite graphs (Section 4.1). Although there exist matrix spaces which have no full-rank matrices or shrunk subspaces, properties of graphs may not always be generalized into the "linear algebraic world" in a good manner. However, such a viewpoint is helpful to come up with new insights from a combinatorial perspective. For instance, if the matrix space is promised to be spanned by rank-1 matrices or triangularizable matrices, Ivanyos et al. [IKQS15] devised an deterministic polynomial-time algorithm which outputs a matrix of maximal rank of the given matrix space. Their framework is based on a generalization of Wong *sequences*, which can be viewed as a linear algebraic analog of *augmenting paths*, introduced for devising classical bipartite perfect matching algorithms.

It is worth noting that we apply the techniques used in this chapter about manipulating matrix spaces to study *rank-critical matrix spaces*. This type of matrix spaces satisfy that any matrix space which properly contains it has the maximal rank strictly greater than that of it. Rank-critical matrix spaces have wide applications and interpretations in both algebraic geometry and computational complexity theory. In [LQ17b], we derive a complete characterization from the complexity perspective, which recovers a previous sufficient condition obtained by Draisma [DRA06] from a geometric perspective. We also investigate rank-critical spaces in the context of compression and primitive matrix spaces and derives several interesting structural results.

Chapter 5

Testing Isometry between Alternating Matrix Spaces

In this chapter, we study the algorithmic problem of testing isometry between two alternating matrix spaces. It is known that solving such a problem in time polynomial in the size of the underlying vector space is equivalent to testing isomorphism of p-groups of class 2 and exponent p in time polynomial in the group order – the widely believed bottleneck case of the group isomorphism problem. We propose a venue of attack for the alternating matrix space isometry problem by viewing it as a linear algebraic analog of the graph isomorphism problem. We first revisit the development of the alternating matrix space isometry problem, including its connection with graph and group isomorphism problems, in Section 5.1. Then we describe the outline of our main algorithm in Section 5.2, and provide a detailed proof in Section 5.4. Furthermore, we apply Luks' *dynamic programming technique* for GraphIso to slightly improve the worst-case time complexity of AltMatSpIso in Section 5.5. We summarize our results in Section 5.6. This chapter is based on [LQ17a].

5.1 Introduction

Let $\mathbb{F}_q$ be the finite field with q elements. An $n \times n$ matrix B over $\mathbb{F}_q$ is *alternating* if for every $u \in \mathbb{F}_q^n$, $u^{\mathrm{t}}Bu = 0$. $\Lambda(n, q)$ denotes the linear space of $n \times n$ alternating matrices over $\mathbb{F}_q$, and a dimension-m subspace of $\Lambda(n, q)$ is called an m-alternating (matrix) space. $\mathrm{GL}(n, q)$ denotes the general linear group of degree n over $\mathbb{F}_q$. We study the following problem.

Problem 5.1 (Alternating matrix space isometry problem, AltMatSpIso)**.** *Given the linear bases of two m-alternating spaces $\mathcal{G}, \mathcal{H}$ in $\Lambda(n, q)$, decide whether there exists $A \in \mathrm{GL}(n, q)$, such that $A^{\mathrm{t}}\mathcal{G}A := \{A^{\mathrm{t}}BA : B \in \mathcal{G}\} = \mathcal{H}$ (as subspaces).*

If such a T exists, we say $\mathcal{G}$ and $\mathcal{H}$ are *isometric*. As will be explained later, AltMatSpIso has been studied, mostly under other names, for decades. It lies at the heart of the *group isomorphism problem* (GroupIso), and has an intimate relationship with the celebrated *graph isomorphism problem* (GraphIso). As a problem in NP$\cap$coAM, its worst-case time complexity has barely been improved over the brute-force algorithm, which runs in time $q^{\Theta(n^2)} \cdot \mathrm{poly}(n, m, \log q)$. In fact, a $q^{O(n+m)}$-time algorithm is already regarded as very difficult.

Let us recall one formulation of GraphIso. For $n \in \mathbb{N}$, let $[n] = \{1, 2, \ldots, n\}$, and S_n denotes the symmetric group on $[n]$. A simple undirected graph is just a subset of $\Lambda_n := \{\{i, j\} : i, j \in [n], i \neq j\}$. A permutation $\sigma \in S_n$ induces a natural action on Λ_n. The following formulation of GraphIso as an instance of the *setwise transporter problem* is well-known [Luk82].

Problem 5.2 (GraphIso, group-theoretic definition)**.** *Given two subsets G, H of Λ_n, decide whether there exists $\sigma \in S_n$, such that $G^\sigma := \{\{\sigma(i), \sigma(j)\} : \{i, j\} \in G\} = H$ (as subsets).*

The formulations of AltMatSpIso and GraphIso as in Problem 5.1 and Problem 5.2 are quite similar. As a related note, it reminds us the formulations of the bipartite perfect matching problem and the Edmonds' problem, introduced in [IKQS15, IQS17a, IQS17b, GGOW16] and mentioned in Section 4.1. It is natural for us to view and

study AltMatSpIso as a linear algebraic analog of GraphIso. On the other hand, AltMatSpIso has been studied for decades as an instance, in fact, the long-believed bottleneck case, of the GroupIso. This problem also has an intricate relationship with the GraphIso. In the next two subsections, we briefly review for two the connections.

5.1.1 Relation with the Group Isomorphism Problem

GroupIso asks to decide whether two finite groups (of the same order n) are isomorphic or not. The difficulty of this problem depends crucially on how we represent the groups in the algorithms. If our goal is to obtain an algorithm running in time *polynomial in* n, then we may assume that we have at our disposal the *Cayley (multiplication) table* of the group, as we can recover the Cayley table from most reasonable models for computing with finite groups. Therefore, in the main text we restrict our discussion to this very redundant model, which is meaningful mainly because we do not know of a poly(n)-time or even an $n^{o(\log n)}$-time algorithm [Wil14] (log to the base 2), despite that a simple $n^{\log n+O(1)}$-time algorithm has been known for decades [FN70, Mil78]. The past few years have witnessed a resurgence of activity on algorithms for this problem with worst-case analysis with respect to the group order; we refer the readers to [GQ17a] which contains a survey of these algorithms.

It is long believed that p-groups form the bottleneck case for GroupIso. In fact, the decades-old quest for a polynomial-time algorithm has focused on class-2 p-groups, with little success. Even if we restrict further our study to consider p-groups of class 2 and exponent p, the problem is still difficult. A group G is a p-group of exponent p if every element in G has order p. A group G is of class 2 if the commutator subgroup $[G, G] := \{g^{-1}h^{-1}gh : g, h \in G\}$ is normal in G. Recent works [LW12, BW12, BMW17, IQ18] solve some nontrivial subclasses of this group class, and have led to substantial improvement in practical algorithms. But the methods in these works do not seem helpful enough to lead to any improvement for the worst-case time complexity of the general class.

By a classical result of Baer [Bae38] (see also [Wil09]), testing isomorphism of p-groups of class 2 and exponent p in time polynomial in the group order reduces to solving AltMatSpIso over $\mathbb{F}_p$ in time $p^{O(m+n)}$. We revisit the reduction here: Suppose we

are given two p-groups of class 2 and exponent p. G_1 and G_2 of order p^ℓ. For G_i, let $b_i : G_i/[G_i, G_i] \times G_i/[G_i, G_i] \to [G_i, G_i]$ be the commutator map, defined by

$$b_i(g_1[G_i, G_i], g_2[G_i, G_i]) = [g_1, g_2], \ \forall \ g_1, g_2 \in G_i. \tag{5.1}$$

By the class 2 and exponent p assumption, $G_i/[G_i, G_i]$ are elementary Abelian groups of exponent p. For G_1 and G_2 to be isomorphic it is necessary that $[G_1, G_1] \cong [G_2, G_2] \cong \mathbb{Z}_p^m$ and $G_1/[G_1, G_1] \cong G_2/[G_2, G_2] \cong \mathbb{Z}_p^n$ for some $m, n \in \mathbb{N}$ satisfying $m+n = \ell$. Furthermore, it is easy to see b_is are alternating bilinear maps. So we have alternating bilinear maps $b_i : \mathbb{F}_p^n \times \mathbb{F}_p^n \to \mathbb{F}_p^m$. G_1 and G_2 are isomorphic if and only if there exist $A \in \mathrm{GL}(n, p)$ and $D \in \mathrm{GL}(m, p)$ such that for every $u, v \in \mathbb{F}_p^n$, $b_1(A(u), A(v)) = D(b_2(u, v))$. Representing b_i as a tuple of alternating matrices $\mathbb{B}_i = (B_1, \ldots, B_m) \in \Lambda(n, p)^m$, it translates to asking whether $A^{\mathrm{t}}\mathbb{B}_1 A = \mathbb{B}_2^D$. Letting $\mathcal{B}_i$ be the linear span of $\mathbb{B}_i$, this becomes an instance of AltMatSpIso with respect to $\mathcal{B}_1$ and $\mathcal{B}_2$.

When $p > 2$, we can reduce AltMatSpIso to the isomorphism testing of p-groups of class 2 and exponent p using the following construction. Starting from $\mathbb{G} \in \Lambda(n, p)^m$ representing $\mathcal{G}$, $\mathbb{G}$ can be viewed as representing a bilinear map $b : \mathbb{F}_p^n \times \mathbb{F}_p^n \to \mathbb{F}_p^m$. Define a group G with operation $\circ$ over the set $\mathbb{F}_p^m \times \mathbb{F}_p^n$ as

$$(v_1, u_1) \circ (v_2, u_2) = (v_1 + v_2 + \frac{1}{2}b(u_1, u_2), u_1 + u_2). \tag{5.2}$$

It can be verified that G is a p-group of class 2 and exponent p, and it is then easy to verify that two such groups G_1 and G_2 built from $\mathbb{G}_1$ and $\mathbb{G}_2$ are isomorphic if and only if $\mathcal{G}_1$ and $\mathcal{G}_2$ are isometric.

When working with groups in the Cayley table model and working with AltMatSpIso in time $p^{O(m+n)}$, the above procedures can be performed efficiently. Because of these reductions and the current status of GroupIso, we see that AltMatSpIso lies at the heart of GroupIso, and solving AltMatSpIso in $q^{O(m+n)}$ is already very difficult.

5.1.2 Relation with the Graph Isomorphism Problem

The celebrated graph isomorphism problem asks to decide whether two undirected simple graphs are isomorphic. The relation between AltMatSpIso and GraphIso is very delicate. Roughly speaking, the two time-complexity measures of AltMatSpIso, $q^{O(n+m)}$ and $\text{poly}(n, m, q)$, sandwiches GraphIso in an interesting way. For one direction, solving AltMatSpIso in time $q^{O(n+m)}$ can be reduced to solving GraphIso for graphs of size $q^{O(n+m)}$ by first reducing it to solve GroupIso for groups of order $q^{O(n+m)}$ as above, and then to solve GraphIso for graphs of size $q^{O(n+m)}$ by the reduction from GroupIso to GraphIso [KST93]. Therefore, a polynomial-time algorithm for GraphIso implies an algorithm for AltMatSpIso in time $q^{O(n+m)}$. For the other direction, Grochow and Qiao [GQ17b] showed that solving GraphIso in polynomial time reduces to solving AltMatSpIso over $\mathbb{F}_q$ in time $\text{poly}(n, m, q)$ with $q = \text{poly}(n)$.

It is reasonable to examine whether the recent breakthrough of Babai [Bab16a, Bab16b], a quasipolynomial-time algorithm for GraphIso, helps with reducing the time complexity of AltMatSpIso. This seems *unlikely*. One indication is that the brute-force algorithm for AltMatSpIso is already quasipolynomial with respect to $q^{O(n+m)}$. Other evidence is that Babai noted that his algorithm did not seem to be helpful for improving GroupIso, and posed GroupIso as one roadblock for putting GraphIso in P [Bab16a, Section 13.2]. Since AltMatSpIso captures the long-believed bottleneck case for GroupIso, the current results for GraphIso are unlikely to improve the time complexity to $q^{O(n+m)}$. There is also an explanation from the technical viewpoint [GR16]. Roughly speaking, the barrier in the group-theoretic framework for GraphIso is dealing with large alternating groups, as other composition factors like projective special linear groups can be handled by brute-force in quasipolynomial time, so for the purpose of a quasipolynomial-time algorithm these groups are not a concern. On the other hand, for AltMatSpIso it is exactly the projective special linear groups that form a bottleneck.

5.1.3 Current Status and ALgorithmic Results

It is not hard to show that AltMatSpIso $\in$ NP $\cap$ coAM, so it is unlikely to be NP-complete unless the polynomial hierarchy collapse to the second level [GMW86]. As to the worst-case time complexity, the brute-force algorithm for AltMatSpIso runs in time $q^{n^2} \cdot \mathrm{poly}(m, n, \log q)$, by simply enumerating the elements in $\mathrm{GL}(n, q)$ and verifying whether it is an isometry. Another analyzed algorithm for AltMatSpIso offers a running time of $q^{\frac{1}{4}(n+m)^2+O(n+m)}$ when $q = p$ is a prime, by first reducing to testing isomorphism of class-2 and exponent-p p-groups of order p^{n+m}, and then applying Rosenbaum's $N^{\frac{1}{4}\log_p N+O(1)}$-time algorithm for p-groups of order N [Ros13]. This is only better than the brute-force one when $m < n$.[1] It is somewhat embarrassing that for a problem in NP $\cap$ coAM, we are barely able to achieve an improvement over the brute-force algorithm in a limited range of parameters.

On the other hand, practical algorithms for AltMatSpIso have been implemented. As far as we know, the currently implemented algorithms for AltMatSpIso can handle the case when $m + n \approx 20$ and $p \approx 13$, but absolutely not the case if $m + n \approx 200$, though for $m + n \approx 200$ and say $p \approx 13$ the input can be stored in a few megabytes.[2] For GraphIso, the programs Nauty and Traces [MP14] can test isomorphism of graphs stored in gigabytes in a reasonable amount of time. Therefore, unlike GraphIso, AltMatSpIso seems difficult even in the practical sense.

From the discussion above, we see that solving AltMatSpIso with a worst-case time complexity $q^{O(n+m)}$ is already a difficult target. In a very true sense, our current understanding of the worst-case time complexity of AltMatSpIso is like the situation for GraphIso in the 1970s. In the development of algorithms for GraphIso, one breakthrough was to develop *average-case efficient* algorithms: for almost all graphs in certain random models, isomorphism testing can be performed efficiently. The first of such algorithms was devised by Babai, Erdős and Selkow [BES80], which tests isomorphism for all but $o(2^{\binom{n}{2}})$

[1] As pointed out in [BMW17], there are numerous unanalyzed algorithms [O'B94, ELGO02] which may lead to some improvement, but $q^{cn^2} \cdot \mathrm{poly}(n, m, \log q)$ for some constant $0 < c < 1$ is a reasonable over estimate of the best bound by today's method.

[2] We thank James B. Wilson, who maintains a suite of algorithms for p-group isomorphism testing, for communicating his hands-on experience to us. We take the responsibility for any possible misunderstanding or not knowing of the performance of other implemented algorithms.

of the $2^{\binom{n}{2}}$ graphs on n vertices in time $O(n^2)$. This algorithm was then improved by Lipton [Lip78], Karp [Kar79], and Babai and Kučera [BK79]. The random models used in the average-case analysis is the celebrated Erdős-Rényi model [ER59, ER63, Bol01], which is the uniform probability distribution over the set of size-m subsets of Λ_n. That is, each subset is endowed with probability $1/\binom{\binom{n}{2}}{m}$.

Recall the "linear algebraic" idea: Vectors in $\mathbb{F}_q^n$ are viewed as vertices; matrices in an m-alternating space are viewed as edges. The $q^{O(n+m)}$ measure can be thought of as polynomial in the number of "vertices" and the number of "edges". Based on the "linear algebraic" viewpoint, we could expect to derive an *average-case efficient* algorithm for AltMatSpIso runs in time $q^{O(m+n)}$. Simultaneously, we may also formulate a model of random alternating matrix space over $\mathbb{F}_q$ as follows: Let $\left[\ \right]_q$ be the Gaussian binomial coefficient with base q. The linear algebraic Erdős-Rényi model, LinER(n, m, q), is defined as the uniform probability distribution over the set of dimension-m subspaces of $\Lambda(n, q)$. That is, each subspace is endowed with probability $1/\left[\binom{n}{2}\right]_q$.

Here, the parameter m comes into the theme. because q^m, while no more than $q^{\binom{n}{2}}$, is not necessarily bounded by a polynomial in q^n. This is in contrast to GraphIso, where the edge number is at most quadratic in the vertex number. In particular, when $m = \Omega(n^2)$, the brute-force algorithm is already in time $q^{O(n+m)}$. On the other hand, when m is very small compared to n, say $m = O(1)$, we can enumerate all elements in $\mathrm{GL}(m, q)$ in time $q^{O(1)}$, and apply the isometry testing for *alternating matrix tuples* from [IQ18] which runs in randomized time $\mathrm{poly}(n, m, q)$. Therefore, the $q^{O(n+m)}$-time measure makes most sense when m is comparable with n, in particular when $m = \Theta(n)$.

The above discussion, as well as the one related to the non-commutative rank problem and the linear algebraic analog of bipartite graphs, indicate that realities in the combinatorial world and the linear algebraic world can be quite different. So meaningful results cannot be obtained by adapting the results for graphs to alternating matrix spaces in a straightforward fashion. One purpose is to provide further evidence that, despite potential technical difficulties, certain ideas that have been developed for GraphIso can be adapted to work with AltMatSpIso as well. We take a short cut, by presenting our main algorithmic result:

Theorem 5.1 (Main result)**.** *Suppose $m = cn$ for some constant c. There is an algorithm which, for almost but at most $1/q^{\Omega(n)}$ fraction of alternating matrix spaces $\mathcal{G}$ in* $\textsc{LinER}(n, m, q)$*, tests any alternating matrix space $\mathcal{H}$ for isometry to $\mathcal{G}$ in time $q^{O(n)}$.*

An important ingredient in Theorem 5.1, the utility of which should go beyond the average-case setting, is an adaptation of the *individualization* technique for GraphIso to AltMatSpIso. We also realize a reformulation of the *refinement* technique for GraphIso as used in [BES80] in the AltMatSpIso setting. Individualization and refinement are very influential combinatorial ideas for GraphIso and have been crucial in the progress of the worst-case time complexity of GraphIso, including Babai's recent breakthrough [Bab16a, Bab16b], but were missing in the GroupIso context.

In addition, for an m-alternating space $\mathcal{G}$ in $\Lambda(n, q)$, we define the *autometry group* of $\mathcal{G}$, $\mathrm{Aut}(\mathcal{G})$ as $A \in \mathrm{GL}(n, q) : A^{\mathrm{t}}\mathcal{G}A = \mathcal{G}$. The proof of Theorem 5.1 implies the following, which can be viewed as a weaker correspondence of the classical result that *most graphs have trivial automorphism groups* [ER63].

Corollary 5.2. *Suppose $m = cn$ for some constant c. All but $1/q^{\Omega(n)}$ fraction of alternating matrix spaces in* $\textsc{LinER}(n, m, q)$ *have autometry groups of size $q^O(n)$.*

Another piece of evidence which supports the usefulness of the "linear algebraic" viewpoint is by adapting Luks' dynamic programming technique for GraphIso [Luk99] to AltMatSpIso. In the GraphIso setting, this technique improves the naive $n! \cdot \mathrm{poly}(n)$ time bound to the $2^{O(n)}$ time bound, which can be understood as replacing the number of permutations $n!$ with the number of subsets 2^n. In the linear algebraic setting, the analog would be replacing $\Theta(q^{n^2})$, the number of invertible matrices over $\mathbb{F}_q$, with the number of subspaces in $\mathbb{F}_q^n$ (which is roughly $q^{\frac{1}{4}n^2+O(n)}$). We show that this is indeed possible.

Theorem 5.3. AltMatSpIso *can be solved in time $q^{\frac{1}{4}(m^2+n^2)+O(m+n)}$.*

Note that the quadratic term on the exponent of the algorithm in Theorem 5.3 is $\frac{1}{4}(m^2 + n^2)$, slightly better than the one based on Rosenbaum's result [Ros13], which is $\frac{1}{4}(m+n)^2$.

5.2 Towards the Main Algorithm

We now describe the outline of the algorithm for Theorem 5.1, which is inspired by the first average-case efficient algorithm for GraphIso by Babai, Erdős, and Selkow [BES80]. We will recall the individualization and refinement technique therein. We define a linear algebraic individualization, and propose a reformulation of the refinement step. Then we present an outline of the main algorithm. During the procedure we will also see how the "linear algebraic" viewpoint guides the generalizations here.

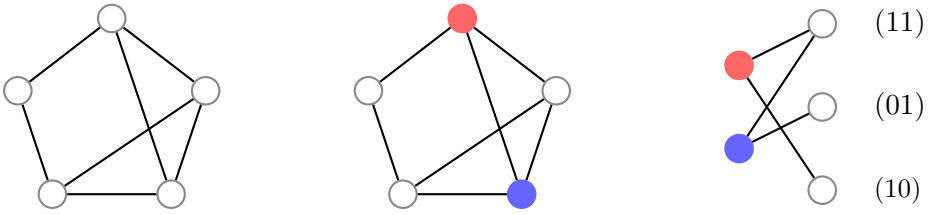
5.2.1 A Variant of the Naive Refinement Algorithm

Two properties of random graphs are used in the average-case analysis of the algorithm in [BES80]. The first property is that *most graphs have the first* $\lceil 3 \log n \rceil$ *largest degrees distinct*. The second property, which is relevant to us, is the following.

Let $G = ([n], E)$ be a simple and undirected graph. Let $r = \lceil 3 \log n \rceil$, $S = [r]$ and $T = [n] \backslash [r]$. Define $B = (S \bigcup T, F)$ as the bipartite graph induced by the cut $[r] \cup \{r+1, \ldots, n\}$, where $F = \{(i, j) : \ i \in S, \ j \in T, \ \{i, j\} \in E\}$. For each $j \in T$, assign a length-r bit string f_j as follows: $f_j \in \{0,1\}^r$ such that $f_j(k) = 1$ if and only if $(k, j) \in F$ for $k = [r]$. It is easy to verify that, *all but at most* $\frac{1}{O(n)}$ *fraction of graphs satisfy that* f_j*'s are distinct over* $j \in T$.

Let us see how the above property alone, together with the individualization and refinement heuristic, give an average-case algorithm in $n^{O(\log n)}$ for GraphIso. Suppose G satisfies the property stated in the last paragraph, and we would like to test isomorphism between $G = ([n], E)$ and an arbitrary graph $H = ([n], E')$. Let $S_G \subseteq \{0,1\}^r$ be the set of r bit strings obtained in the procedure above, with respect to the cut $[r] \cup \{r+1, \ldots, n\}$. The above property guarantees that $|S_G| = n - r$. In the *individualization* step, we enumerate all r-tuple of vertices in H with a multiplicative cost at most n^r. For a fixed r-tuple $(i_1, \ldots, i_r) \in [n]^r$, we perform the *refinement* step, that is, label the remaining vertices in H according to their adjacency relations with $(i_1, \ldots, i_r)$ to obtain another set of bit-strings S_H. If $S_G \neq S_H$ we neglect this r-tuple. If $S_G = S_H$, then we can form a bijective map between $[n]$ and $[n]$, by mapping j to i_j for $j \in [r]$, and the rest according to their

labels. Finally check whether this bijective map induces an isomorphism. See Figure 5.1 for an illustrative example about how the individualization and refinement procedures work.



(A) Graph G (B) Individualization (C) Refinement

FIGURE 5.1: For a given graph G. Individualize the top (red) and lower left (blue) vertices. We obtain the induced bipartite graphs and label the rest of vertices based on their adjacency relations with the individualized vertices.

The above algorithm runs in time $n^{O(\log n)}$, which tests isomorphism between G and H given that G satisfies the required property. In particular, this algorithm also implies that for such G, $|\mathrm{Aut}(G)| \leq n^{O(\log n)}$. To recover the algorithm in [BES80], assuming that the largest r degrees are distinct (the first property mentioned in the beginning), one can canonicalize the choice of the r-tuples by choosing the one with largest r degrees for both G and H.

5.2.2 Individualization and Refinement in the AltMatSpIso Setting

We aim to generalize the above idea to the setting of AltMatSpIso. To do this, we first make sense of what individualization means in the alternating space setting. We then discuss how the refinement step may be generalized, and indicate how we follow an alternative formulation of it.

Let $\mathcal{G}, \mathcal{H} \leq \Lambda(n, q)$ be two m-alternating spaces for which we want to test isometry. As the case in subSection 5.2.1, we will look for properties of $\mathcal{G}$ which enable the average-case analysis, and perform individualization on $\mathcal{H}$ side. For $i \in [n]$, let e_i denotes the ith standard basis vector of $\mathbb{F}_q^n$.

Individualization. In the graph setting, individualizing r vertices in H can be understand as follows. First we fix a size-r subset L of $[n]$ as well as an order on the elements

in L. The result is a tuple of distinct vertices $(i_1, \ldots, i_r) \in [n]^r$. Enumerating such tuples incurs a multiplicative cost of at most n^r.

In the alternating space setting, we simply view vectors in $\mathbb{F}_q^n$ as vertices, connected by matrices in $\mathcal{H}$. Consider the following procedure: First fix a dimension-r subspace L of $\mathbb{F}_q^n$ with an ordered basis, which is represented by a tuple of linearly independent vectors $(v_1, \ldots, v_r) \in (\mathbb{F}_q^n)^r$. Enumerating all such dimension-r subspaces incurs a multiplicative cost of at most q^{rn}. Up to this point, this is in complete analogy with the graph setting. We may stop here and say that an r-individualization amounts to fix an r-tuple of linearly independent vectors.

We can go a bit further though. As will be clear in the following, it is beneficial if we also fix a complement subspace R of L, i.e. $R \leq \mathbb{F}_q^n$ satisfying $L \cap R = \{0\}$ and $\langle L \cup R \rangle = \mathbb{F}_q^n$. This adds another multiplicative cost of $q^{r(n-r)}$, which is the number of complement subspaces of a fixed dimension-r subspace in $\mathbb{F}_q^n$. In the graph setting, this step is not necessary, because for any $L \subseteq [n]$ there exists a unique complement subset $R = [n] \setminus L$. To summarize, by an *r-individualization*, we mean choosing a direct sum decomposition $\mathbb{F}_q^n = L \oplus R$ where $\dim(L) = r$ and $\dim(R) = n - r$, together with an ordered basis $(v_1, \ldots, v_r)$ of L. Enumerating all r-individualizations incurs a total multiplicative cost of at most q^{2rn-r^2}.

Towards a refinement step as in [BES80]. In the GraphIso setting, individualizing r vertices which gives $(i_1, \ldots, i_r) \in [n]^r$ allows us to focus on isomorphisms that respect this individualization, which are those isomorphism ϕ between two graphs G and H such that $\phi(j) = i_j$ for $j \in [r]$. There are at most $(n-r)!$ such isomorphisms. Since r is usually set as $O(\log n)$, just naively trying all such permutations does not help. Therefore the individualization is usually accompanied with a refinement type technique. Specifically, For an individualization $L = (i_1, \ldots, i_r)$, set $R = [n] \setminus L$. The refinement step as in [BES80] assigns every $v \in R$ a label according to its adjacency relation with respect to $(i_1, \ldots, i_r)$, an ordered set of vertices in L. Each label in fact corresponds to a *subset* of L, and an individualization-respecting isomorphism has to preserve this adjacency relation for every vertex in R. Such a restriction turns out to be quite severe for most graphs: as mentioned in subSection 5.2.1, for most graphs G, the adjacency relations between $(1, 2, \ldots, r)$ and

$j \in [n] \setminus [r]$ are completely different over j. For such a graph G and any individualization of H, this means that there is at most one way to extend $\phi(j) = i_j$ for $j \in [r]$ to an isomorphism between G and H.

In the AltMatSpIso setting, an r-individualization also allows us to focus on isometries which respect the decomposition $L \oplus R$ and the ordered basis $(v_1, \ldots, v_r)$ of L, namely those ϕ such that $\phi(e_i) = v_i$ for $i \in [r]$, and $\phi(\langle e_{r+1}, \ldots, e_n \rangle) = R$. There are at most $q^{(n-r)^2}$ such isometries. Since r will be also set to be very small - in fact a constant here - we also need some refinement type argument.

We may simply generalize the refinement step in [BES80]. For $u \in R$, we can record its "adjacency relation" with respect to an r-individualization $\mathbf{v} := (v_1, \ldots, v_r)$ as a *subspace* of $L \cong \mathbb{F}_q^r$ as follows. For $H \in \mathcal{H} \leq \Lambda(n, q)$, define $H(\mathbf{v}, u) := (v_1^{\mathrm{t}} H u, \ldots, v_r^{\mathrm{t}} H u)^{\mathrm{t}} \in \mathbb{F}_q^r$, and $\mathcal{H}(\mathbf{v}, u) := \{H(\mathbf{v}, u) : Q \in \mathcal{H}\}$. $\mathcal{H}(\mathbf{v}, u)$is a subspace in $\mathbb{F}_q^r$ which records the "adjacency relation" between $(v_1, \ldots, v_r)$ and u under $\mathcal{H}$. It can be verified that an *individualization-respecting* isometry has to preserve this adjacency relation. It is tempting to check then on the $\mathcal{G}$ side, where we individualize the first r standard basis (which produces $(e_1, \ldots, e_r)$ and $\langle e_{r+1}, \ldots, e_n \rangle$), whether for most $\mathcal{G}$'s it is the case that every $v \in \langle e_{r+1}, \ldots, e_n \rangle$ gets a unique label. If this is so, then the number of individualization-respecting isomorphisms can also be significantly reduced. However, this cannot be the case when r is small, as there are $q^{(n-r)^2}$ vectors in R but there are at most q^{r^2} subspaces in $\mathbb{F}_q^r$.

Since we are looking for linear maps from $\langle e_{r+1}, \ldots, e_n \rangle$ to R, the above counting argument does not make much sense, as it mostly concerns *setwise maps* from $\langle e_{r+1}, \ldots, e_n \rangle$ to R. It is indeed the case, and we further note that the map from $u \in R$ to $\mathcal{H}(\mathbf{v}, u) \leq \mathbb{F}_q^r$ defines a sheaf over the projective space $\mathbb{P}(R)$. So, such labels have some nontrivial relation to glue together to form a sheaf. (See the related concept of kernel sheaves as in [KV12].) It may be possible to use these observations to define a reasonable refinement step in the alternating matrix space setting as follows.

A reformulation of the refinement step. To resolve the above problem, we reformulate the idea in the GraphIso setting as follows. On the G side we start with the standard individualization $[r] \cup \{r+1, \ldots, n\}$ with an order on $[r]$ as $(1, \ldots, r)$, which induce the bipartite graph $B = (S \cup T, F)$ where $S = [r]$, $T = [n] \setminus [r]$, and the edge set F is induced from G. On

the H side, a fixed individualization, say $\mathbb{F}_q^n = L \cup R$, $L = (i_1, \ldots, i_r) \subseteq [n]^r$, also induces a bipartite graph $C = (L \cup R, F')$, where F' is induced from H. A bijective $\psi : T \to R$ is a *right-side isomorphism* between B and C if it is not only an isomorphism between T and R, but also induces an isomorphism between B and C as bipartite graphs. With respect to the chosen individualizations on G and H sides, let $\mathrm{RIso}(B, C)$ be the set of right-side isomorphisms, $\mathrm{IndIso}(G, H)$ be the set of individualization-respecting isomorphisms from G to H. Note that both $\mathrm{RIso}(B, C)$ and $\mathrm{IndIso}(G, H)$ can be embedded to the set of bijective maps between T and R. The key observation is that an individualization-respecting isomorphism has to be a right-side isomorphism between B and C, i.e. $\mathrm{IndIso}(G, H) \subseteq \mathrm{RIso}(B, C)$. Also note that either $|\mathrm{RIso}(B, C)| = 0$ (e.g. when B and C are not right-isomorphic), or $|\mathrm{RIso}(B, C)| = |\mathrm{RAut}(B)|$ where $\mathrm{RAut}(B) := \mathrm{RIso}(B, B)$. The refinement step as in subSection 5.2.1 achieves two goals. Firstly on the G side, most G's in $\mathrm{ER}(n, m)$ have the corresponding induced bipartite graph B with $|\mathrm{RAut}(B)| = 1$. This means that $|\mathrm{RIso}(B, C)| \leq 1$. Secondly, given H with a fixed individualization which induce the bipartite graph C, there is an efficient procedure to decide whether B and C are right-isomorphic (e.g. by comparing the labels), and if they do, enumerate all right-isomorphisms (actually unique).

In the AltMatSpIso setting, on the $\mathcal{G}$ side we start with the standard individualization $S = \langle e_1, \ldots, e_r \rangle$, $T = \langle e_{r+1}, \ldots, e_n \rangle$ with the ordered basis $(e_1, \ldots, e_r)$ of S. We can also define a correspondence of the bipartite graph in this setting, which is the matrix space

$$\mathcal{B}' = \{ \begin{bmatrix} e_1 & \cdots & e_r \end{bmatrix}^{\mathrm{t}} G \begin{bmatrix} e_{r+1} & \cdots & e_n \end{bmatrix} : G \in \mathcal{G} \} \leq M(r \times (n-r), q), \tag{5.3}$$

where $\begin{bmatrix} e_1 & \cdots & e_r \end{bmatrix}$ and $\begin{bmatrix} e_{r+1} & \cdots & e_n \end{bmatrix}$ denotes the $n \times r$ and $n \times (n-r)$ matrices listing the column vectors $\{e_1, \ldots, e_r\}$ and $\{e_{r+1}, \ldots, e_n\}$, respectively. $\begin{bmatrix} e_1 & \cdots & e_r \end{bmatrix}^{\mathrm{t}} G \begin{bmatrix} e_{r+1} & \cdots & e_n \end{bmatrix}$ stands for the upper-right $r \times (n-r)$ submatrix of G (recall that $G \in \mathcal{G}$ is represented with respect to the standard basis). Similarly, the individualization on the $\mathcal{H}$ side yields $L \oplus R$ with an ordered basis of L, $(v_1, \ldots, v_r)$ where $v_i \in \mathbb{F}_q^n$. Take any basis of $R = \langle v_{r+1}, \ldots, v_n \rangle$. We can construct

$$\mathcal{C}' = \{ \begin{bmatrix} v_1 & \cdots & v_r \end{bmatrix}^{\mathrm{t}} H \begin{bmatrix} v_{r+1} & \cdots & v_n \end{bmatrix} : H \in \mathcal{H} \} \leq M(r \times (n-r), q). \tag{5.4}$$

We say $A \in \mathrm{GL}(n-r,q)$ is a *right-side equivalence* between $\mathcal{B}'$ and $\mathcal{C}'$ if $\mathcal{B}'A := \{B'A : B' \in \mathcal{B}'\} = \mathcal{C}'$ (as subspaces). Let $\mathrm{RIso}(\mathcal{B}', \mathcal{C}')$ be the set of right-side equivalences between $\mathcal{B}'$ and $\mathcal{C}'$, and $\mathrm{IndIso}(\mathcal{G}, \mathcal{H})$ the set of *individualization-respecting isometries* between $\mathcal{G}$ and $\mathcal{H}$. Similarly, both $\mathrm{RIso}(\mathcal{B}', \mathcal{C}')$ and $\mathrm{IndIso}(\mathcal{G}, \mathcal{H})$ can be embedded in the set of invertible linear maps from T to R (which is isomorphic to $\mathrm{GL}(n-r,q)$), and we have $\mathrm{IndIso}(\mathcal{G}, \mathcal{H}) \subseteq \mathrm{RIso}(\mathcal{B}', \mathcal{C}')$. Furthermore, $\mathrm{RIso}(\mathcal{B}', \mathcal{C}')$ is either empty (e.g. $\mathcal{B}'$ and $\mathcal{C}'$ are not right-side equivalent), or a coset of $\mathrm{RAut}(\mathcal{B}') := \mathrm{RIso}(\mathcal{B}', \mathcal{B}')$. So in analogy with the GraphIso setting, for our purpose the goals become:

1. for most m-alternating space $\mathcal{G} \leq \Lambda(n,q)$ (as discussed in subSection 5.1.3, we assume $m = cn$ for some constant c), setting r to be some constant, we have $|\mathrm{RAut}(\mathcal{B}')| \leq q^{O(n)}$, and
2. for $\mathcal{G}$'s satisfying 1, $\mathrm{RIso}(\mathcal{B}', \mathcal{C}')$ can be enumerated efficiently.

5.2.3 Algorithm Outline

Now we outline our algorithm which tests isometry between two m-alternating spaces $\mathcal{G}, \mathcal{H} \in \Lambda(n,q)$, given their linear basis $\mathcal{G} = \langle G_1, \ldots, G_m \rangle$ and $\mathcal{H} = \langle H_1, \ldots, H_m \rangle$. In the outline we assume $r = 4$ and $m = n - 4$, and deal with the general case in Section 5.4.

We first define the property on $\mathcal{G}$ for the sake of average-case analysis. Given those $G_k \in \Lambda(n,q)$ linearly spanning $\mathcal{G}$, we can form a 3-tensor $\mathbb{G} \in \mathbb{F}_q^{n \times n \times m}$ where $\mathbb{G}(i,j,k)$ denotes the (i,j)th entry of G_k. Let $\mathbb{B}'$ be the upper-right $r \times (n-r) \times m$ subtensor of $\mathbb{G}$, with B'_k being the corresponding corner in G_k. Let $\mathcal{B}' = \langle B'_1, \ldots, B'_m \rangle$ as discussed above. $A \in \mathrm{RAut}(\mathcal{B}') \leq \mathrm{GL}(n-r,q)$ if and only if there exists $D \in \mathrm{GL}(m,q)$ such that $\forall i \in [m]$, $\sum_{j \in [m]} d_{i,j} B'_j = B'_i A$, where $d_{i,j}$ is the (i,j)th entry of D. It will be more convenient if we flip $\mathbb{B}'$, which is of size $r \times (n-r) \times m$, to the $\mathbb{B}$, which is of size $(n-r) \times m \times r$ (Figure 5.2). Slicing $\mathbb{B}$ along the third index, we obtain an r-tuple of $(n-r) \times m$ matrices, denoted by $B_1, \ldots, B_r$ (figure 5.3).

Define the set of equivalences of $\mathbb{B}$ as

$$\mathrm{Aut}(\mathbb{B}) := \{(A, D) \in \mathrm{GL}(n-r,q) \times \mathrm{GL}(m,q) : \forall i \in [r], AB_iD^{-1} = B_i\}. \tag{5.5}$$

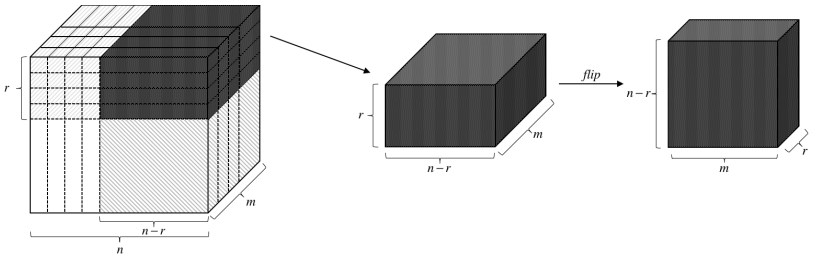


FIGURE 5.2: The 3-tensor $\mathbb{G}$, and flipping $\mathbb{B}'$ to get $\mathbb{B}$.

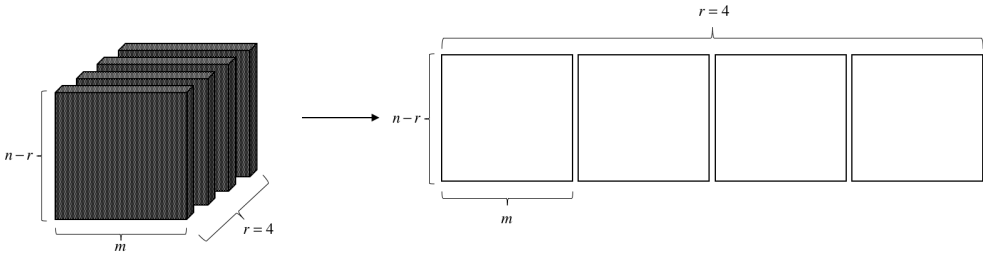


FIGURE 5.3: Slicing $\mathbb{B}$.

Note that $\mathrm{RAut}(\mathcal{B}')$ is the projection of $\mathrm{Aut}(\mathbb{B})$ to the first component. We define the adjoint algebra of $\mathbb{B}$ as

$$\mathrm{Adj}(\mathbb{B}) := \{(A, D) \in M(n-r, q) \oplus M(m, q) : \forall i \in [r], AB_i = B_i D\}. \tag{5.6}$$

$(A, D) \in M(n-r, q) \oplus M(m, q)$ is called invertible, if both A and D are invertible. Clearly, $\mathrm{Aut}(\mathbb{B})$ consists of the invertible elements in $\mathrm{Adj}(\mathbb{B})$. Recall that we focus on the case that $r = 4$ and $m = n - r = n - 4$. It can be shown that the *adjoint algebra of* 4 *random matrices in* $M(m, q)$ *is of size* $q^{O(n)}$ *with probability* $1 - 1/q^{\Omega(n)}$. The key to prove this statement is the stable notion from geometric invariant theory [MFK94] in the context of the *left-right action of* $\mathrm{GL}(m, q) \times \mathrm{GL}(m, q)$ *on matrix tuples* $M(m, q)^r$. In this context, a matrix tuple $(B_1, \ldots, B_r) \in M(m, q)^r$ is *stable*, if for every nontrivial subspace $U \leq \mathbb{F}_q^n$, $\dim(\langle \cup_{i \in [r]} B_i(U) \rangle) > \dim(U)$. An upper bound on $|\mathrm{Adj}(\mathbb{B})|$ can be obtained by analyzing this notion using some classical algebraic results and elementary probability calculations. The property we impose on $\mathcal{G}$ is that *the corresponding* $|\mathrm{Adj}(\mathbb{B})| \leq q^{O(n)}$. It can be verified that this property does not depend on the choices of bases of $\mathcal{G}$.

Now we have achieved our first goal: defining a good property satisfied by most $\mathcal{G}$'s. Let us see how this property enables an algorithm for such $\mathcal{G}$'s. For an arbitrary $\mathcal{H} \leq \Lambda(n, q)$,

at a multiplicative cost of $q^{2rn-r^2} \in q^{O(n)}$ (recall that $r = 4$) we can enumerate all r-individualizations of $\mathcal{H}$. Consider a fixed one, say $\mathbb{F}_q^n = L \oplus R$ with an ordered basis $(v_1, \ldots, v_r)$ of L and $R = \langle v_{r+1}, \ldots, v_n \rangle$ represented by an arbitrary basis $\{v_{r+1}, \ldots, v_n\}$. We construct $\mathbb{C}'$ with respect to the chosen individualization, flip to get $\mathbb{C}$, and slice $\mathbb{C}$ into r $m \times m$ matrices $(C_1, \ldots, C_r)$ as we have done to $\mathcal{B}'$. The task then becomes to compute

$$\mathrm{Adj}(\mathbb{B}, \mathbb{C}) := \{(A, D) \in M(n-r, q) \oplus M(m, q) : \forall i \in [r], AB_i = C_i D\}. \tag{5.7}$$

Viewing A and D as variable matrices, $AB_i = C_i D$ are linear equations on A and D, so the solution set can be computed efficiently. Note that $|\mathrm{Adj}(\mathbb{B})| \leq q^{O(n)}$. For $\mathrm{Adj}(\mathbb{B}, \mathbb{C})$ to contain an invertible element, it must be that $|\mathrm{Adj}(\mathbb{B}, \mathbb{C})| = |\mathrm{Adj}(\mathbb{B})| \leq q^{O(n)}$. In this case, we can enumerate all elements in $\mathrm{Adj}(\mathbb{B}, \mathbb{C})$ in time $q^{O(n)}$. For each element $(A, D) \in \mathrm{Adj}(\mathbb{B}, \mathbb{C})$, test whether it is invertible, and if so, test whether the A in that solution induces an isometry together with the individualization. This completes a high-level description of the algorithm. In particular, the above procedure implies that if $\mathcal{G}$ satisfies $|\mathrm{Adj}(\mathbb{B})| \leq q^{O(n)}$, then $|\mathrm{Aut}(\mathcal{G})| \leq q^{O(n)}$.

5.3 Preliminaries

We collect and restate some notation and definitions for our proofs. q is reserved for prime powers, and p for primes. For $n \in \mathbb{N}$, $[n] := \{1, \ldots, n\}$. $\mathbb{F}_q$ denotes the field of size q. For $i \in [n]$, e_i denotes the ith standard basis vector of $\mathbb{F}_q^n$. $M(s \times t, q)$ denotes the linear space of matrices of size $s \times t$ over $\mathbb{F}_q$, and $M(s, q) := M(s \times s, q)$. I_s denotes the $s \times s$ identity matrix. For $A \in M(s \times t, q)$, A^{t} denotes the transpose of A. $\mathrm{GL}(n, q)$ is the general linear group of degree n over $\mathbb{F}_q$, which consists of all $n \times n$ invertible matrices over $\mathbb{F}_q$. $\Lambda(n, q)$ is the linear space of alternating matrices of size $n \times n$ over $\mathbb{F}_q$. We denote $N_n := \binom{n}{2} = \dim(\Lambda(n, q))$, or just N if n is obvious from the context. We use $\begin{bmatrix} n \\ k \end{bmatrix}_q$ for the Gaussian binomial coefficient with base q, and $\binom{n}{k}$ for the ordinary binomial coefficient. For $N \in \mathbb{N}$ and $m \in [N] \cup \{0\}$,

$$\begin{bmatrix} N \\ m \end{bmatrix}_q = \frac{(1-q^N)(1-q^{N-1}) \cdots (1-q^{N-m+1})}{(1-q)(1-q^2) \cdots (1-q^m)} \tag{5.8}$$

counts the number of dimension-m subspaces in $\mathbb{F}_q^N$.

By a random vector in $\mathbb{F}_q^N$, we mean a vector of length N where each entry is chosen independently and uniformly at random from $\mathbb{F}_q$. By a random matrix in $M(s \times t, q)$, we mean a matrix of size $s \times t$ where each entry is chosen independently and uniformly at random from $\mathbb{F}_q$. By a random alternating matrix in $\Lambda(n, q)$, we mean an alternating matrix of size n where each entry in the strictly upper triangular part is chosen independently and uniformly at random from $\mathbb{F}_q$. Then the diagonal entries are set to 0, and the lower triangular entries are set in accordance with the corresponding upper triangular ones.

Lemma 5.4. *Let $N \in \mathbb{N}$ and $m \in [N] \cup \{0\}$.*

1. *For a fixed subspace U in $\mathbb{F}_q^N$ of dimension m, the number of complements of U in $\mathbb{F}_q^N$ is $q^{m(N-m)}$. Here we say a subspace V is a complement of U, if $V \cap U = \{0\}$ and $\langle V \cup U \rangle = \mathbb{F}_q^N$.*
2. *A random matrix $A \in M(N \times m, q)$ is of rank m with probability $\geq 1 - m/q^{N-m+1}$. Moreover, this probability is greater than $\frac{1}{4}$.*

Proof. For 1. Let $U = \langle u_1, \ldots, u_m \rangle$. We first count the number of $(N-m)$-tuple of vectors $(v_1, \ldots, v_{N-m})$ which are linear independent, and does not contain in U. This can be done recursively. We first choose a vector as v_1 which is not of the form $\lambda_1 u_1 + \ldots \lambda_m u_m$, where $\lambda_1, \ldots, \lambda_m \in \mathbb{F}_q$. The number of possible choice are $q^N - q^m$. To choose v_{k+1}, the number of possible choice is $q^N - q^{m+k}$, as v_k is not of the form $\lambda_1 u_1 + \ldots \lambda_m u_m + \mu_1 v_1 + \cdots + \mu_k v_k$, where $\lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_{k-1} \in \mathbb{F}_q$. Thus, the total number of $(N-m)$-tuple of vectors is $(q^N - q^m)(q^N - q^{m+1}) \cdots (q^N - q^{N-1})$.

Note that each subspace spanned by the above chosen $(N-m)$-tuple of vectors is a complement of U. While different tuples of vectors could span the same subspaces. Note that different dimension-$(N-m)$ subspaces have the same number of $(N-m)$-tuple of vectors as their linear bases, as they are one-to-one correspondences and connected by the change-of-bases transformations. Count the number of $(N-m)$-tuple of vectors which generate a fixed dimension-$(N-m)$ subspaces, which equals $(q^{N-m} - 1)(q^{N-m} - q) \cdots (q^{N-m} - q^{N-m-1})$. To conclude, the number of complement subspaces of a fixed

dimension-m subspace U in $\mathbb{F}_q^N$ can be computed by

$$\frac{(q^N - q^m)(q^N - q^{m+1}) \cdots (q^N - q^{N-1})}{(q^{N-m} - 1)(q^{N-m} - q) \cdots (q^{N-m} - q^{N-m-1})} = q^{m(N-m)}. \tag{5.9}$$

For 2, we have

$$\begin{aligned} \Pr[rk(A) = m | A \in M(N \times m, q)] &= (1 - 1/q^N)(1 - 1/q^{N-1}) \ldots (1 - 1/q^{N-m+1}) \\ &\geq 1 - (1/q^N + 1/q^{N-1} + \cdots + 1/q^{N-m+1}) \\ &\geq 1 - m/q^{N-m+1}. \end{aligned} \tag{5.10}$$

Meanwhile, $\Pr[rk(A) = m | A \in M(N \times m, q)] = (1 - 1/q^N)(1 - 1/q^{N-1}) \ldots (1 - 1/q^{N-m+1}) > (1 - 1/2^N)(1 - 1/2^{N-1}) \ldots (1 - 1/2^{N-m+1}) > \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{7}{8} \cdot \cdots \approx 0.288788 > 1/4.$ □

5.3.1 Matrix Tuples and Matrix Spaces

An r-matrix tuple of size $s \times t$ over $\mathbb{F}_q$ is an element in $M(s \times t, q)^r$. An r-matrix space of size $s \times t$ over $\mathbb{F}_q$ is a dimension-r subspace in $M(s \times t, q)$. An m-alternating (matrix) tuple of size n over $\mathbb{F}_q$ is an element from $\Lambda(n, q)^m$. An m-alternating (matrix) space of size n over $\mathbb{F}_q$ is a dimension-m subspace in $\Lambda(n, q)$. We employ $\mathcal{G}, \mathcal{H}, \ldots$ to denote alternating spaces, and $\mathbb{G}, \mathbb{H}, \ldots$ to denote alternating tuples. $\mathcal{B}, \mathcal{C}, \ldots$ are for (not necessarily alternating nor square) matrix spaces, and $\mathbb{B}, \mathbb{C}, \ldots$ for (not necessarily alternating nor square) matrix tuples[3]. We say that a matrix tuple $\mathbb{B}$ represents a matrix space $\mathcal{B}$, if the matrices in $\mathbb{B}$ form a spanning set (not necessarily a basis) of $\mathcal{B}$. Given $A \in M(s, q)$, $D \in M(t, q)$, and $\mathbb{B} = (B_1, \ldots, B_r) \in M(s \times t, q)^r$, $A\mathbb{B}D$ is the tuple $(AB_1D, \ldots, AB_rD)$. For $D \in M(r, q)$, $\mathbb{B}^D := (\sum_{i \in [r]} d_{1,i} B_i, \sum_{i \in [r]} d_{2,i} B_i, \ldots, \sum_{i \in [r]} d_{r,i} B_i)$, where $d_{i,j}$ is the (i, j)th entry of D.

Two alternating tuples $\mathbb{G} = (G_1, \ldots, G_m)$ and $\mathbb{H} = (H_1, \ldots, H_m)$ in $\Lambda(n, q)^m$ are *isometric*, if there exists $A \in \mathrm{GL}(n, q)$, $A^{\mathrm{t}}\mathbb{G}A = \mathbb{H}$. Two alternating spaces $\mathcal{G}$ and $\mathcal{H}$ in $\Lambda(n, q)$ are *isometric*, if there exists $A \in \mathrm{GL}(n, q)$, such that $A^{\mathrm{t}}\mathcal{G}A = \mathcal{H}$ (equal as subspaces). Given alternating tuples $\mathbb{G} \in \Lambda(n, q)^m$ and $\mathbb{H} \in \Lambda(n, q)^m$ representing $\mathcal{G}$ and $\mathcal{H}$ respectively, $\mathcal{G}$ and $\mathcal{H}$ are isometric, if and only if there exists $D \in \mathrm{GL}(m, q)$ such that $\mathbb{G}$ and $\mathbb{H}^D$

[3]To clarify, in other chapters we use $\mathbb{C}$ to denote the complex field.

are isometric – in other words, there exist $A \in \mathrm{GL}(n,q)$ and $D \in \mathrm{GL}(m,q)$, such that $A^{\mathrm{t}}\mathbb{G}A = \mathbb{H}^D$. We use $\mathrm{Iso}(\mathcal{G},\mathcal{H}) \subseteq \mathrm{GL}(n,q)$ to denote the set of isometries between $\mathcal{G}$ and $\mathcal{H}$. When $\mathcal{G} = \mathcal{H}$, the isometries between $\mathcal{G}$ and $\mathcal{G}$ are also called *autometries*. The set of all autometries forms a *matrix group*, i.e. $\mathrm{Aut}(\mathcal{G}) = \mathrm{Iso}(\mathcal{G},\mathcal{G}) \leq \mathrm{GL}(n,q)$. $\mathrm{Iso}(\mathcal{G},\mathcal{H})$ is either empty or a right coset of $\mathrm{Aut}(\mathcal{G})$. Analogously, we can define the corresponding concepts for tuples, such as $\mathrm{Iso}(\mathbb{G},\mathbb{H})$ and $\mathrm{Aut}(\mathbb{G})$.

Two matrix tuples $\mathbb{B} = (B_1,\ldots,B_r)$ and $\mathbb{C} = (C_1,\ldots,C_r)$ in $M(s\times t,q)^r$ are *equivalent*, if there exist $A \in \mathrm{GL}(s,q)$ and $D \in \mathrm{GL}(t,q)$, such that $A\mathbb{B} = \mathbb{C}D$. Two matrix spaces $\mathcal{B}$ and $\mathcal{C}$ in $M(s\times t,q)$ are *equivalent*, if there exist $A \in \mathrm{GL}(s,q)$ and $D \in \mathrm{GL}(t,q)$, such that $A\mathcal{B} = \mathcal{C}D$ (equal as subspaces). By abuse of notation, we use $\mathrm{Iso}(\mathcal{B},\mathcal{C}) \leq \mathrm{GL}(s,q)\times\mathrm{GL}(t,q)$ to denote the set of equivalences between $\mathcal{B}$ and $\mathcal{C}$, and let $\mathrm{Aut}(\mathcal{B}) = \mathrm{Iso}(\mathcal{B},\mathcal{B})$. $\mathrm{Iso}(\mathcal{B},\mathcal{C})$ is either empty or a left coset of $\mathrm{Aut}(\mathcal{B})$. Similarly we can define $\mathrm{Iso}(\mathbb{B},\mathbb{C})$ and $\mathrm{Aut}(\mathbb{B})$. A trivial but useful observation is that $\mathrm{Iso}(\mathbb{B},\mathbb{C})$ and $\mathrm{Aut}(\mathbb{B})$ are naturally contained in certain subspaces of $M(s,q)\oplus M(t,q)$. Following [Wil09], we define the *adjoint algebra* of $\mathbb{B} \in M(s\times t,q)^r$ as

$$\mathrm{Adj}(\mathbb{B}) := \{(A,D) \in M(s,q)\oplus M(t,q) : A\mathbb{B} = \mathbb{B}D\}. \tag{5.11}$$

The adjoint algebra (of matrix tuples) is a classical concept, and has been recently studied in the context of p-group isomorphism testing by Wilson et al. [Wil09, LW12, BW12, BMW17]. We further define the *adjoint space* between $\mathbb{B}$ and $\mathbb{C}$ in $M(s\times t,q)^r$ as

$$\mathrm{Adj}(\mathbb{B},\mathbb{C}) := \{(A,D) \in M(s,q)\oplus M(t,q) : A\mathbb{B} = \mathbb{C}D\}. \tag{5.12}$$

$(A,D) \in M(s,q)\oplus M(t,q)$ is called invertible if both A and D are invertible. Then $\mathrm{Aut}(\mathbb{B})$ ($\mathrm{Iso}(\mathbb{B},\mathbb{C})$) consists of invertible elements in $\mathrm{Adj}(\mathbb{B})$ ($\mathrm{Adj}(\mathbb{B},\mathbb{C})$). An easy observation is that, if $\mathbb{B}$ and $\mathbb{C}$ are isometric, then any isometry between $\mathbb{B}$ and $\mathbb{C}$ defines a bijection between $\mathrm{Adj}(\mathbb{B},\mathbb{C})$ and $\mathrm{Adj}(\mathbb{B})$. Note that the $\mathrm{Adj}(\mathbb{B},\mathbb{C})$ and $\mathrm{Adj}\mathbb{B}$ can be viewed as *linearizations* of $\mathrm{Iso}(\mathbb{B},\mathbb{C})$ and $\mathrm{Aut}(\mathbb{B})$, respectively, which allows us to decide whether $\mathbb{B}$ and $\mathbb{C}$ are equivalent, and compute a generating set of $\mathrm{Aut}(\mathbb{B})$. These two tasks can be performed efficiently by using (sometimes with a little twist) existing algorithms for

testing *module isomorphism* [CIK97, BL08, IKS10] and computing the unit group in a matrix algebra [BO08].[4]

Given $\mathbb{B} = (B_1, \ldots, B_r) \in M(s \times t, q)^r$, let $\mathrm{Im}(\mathbb{B}) := \langle \cup_{i\in[m]} \mathrm{Im}(B_i) \rangle$ and $\mathrm{Ker}(\mathbb{B}) := \cap_{i\in[m]} \mathrm{Ker}(B_i)$. We say $\mathbb{B}$ is *image-nondegenerate* (*kernel-nondegenerate*), if $\mathrm{Im}(\mathbb{B}) = \mathbb{F}_q^s$ ($\mathrm{Ker}(\mathbb{B}) = \{0\}$). If $\mathbb{B}$ is an alternating tuple in $\Lambda(n, q)^m$, then $\mathbb{B}$ is image-nondegenerate if and only if it is kernel-nondegenerate, as $\mathrm{Im}(\mathbb{B})$ and $\mathrm{Ker}(\mathbb{B})$ are orthogonal. $\mathbb{B}$ is *nondegenerate* if it is both image-nondegenerate and kernel-nondegenerate. It is easy to see that, if $\mathbb{B}$ is image-nondegenerate (kernel-nondegenerate), then the projection of $\mathrm{Adj}(\mathbb{B})$ to the first (second) component along the second (first) component is injective.

For a matrix tuple $\mathbb{B} = (B_1, \ldots, B_r) \in M(s \times t, q)^r$ and a (vector) subspace $U \leq \mathbb{F}_q^t$, the image of U under $\mathbb{B}$ is $\mathbb{B}(U) := \langle \cup_{i\in[m]} B_i(U) \rangle$. It is easy to verify that, $(A\mathbb{B})(U) = A(\mathbb{B}(U))$, and $(\mathbb{B}D)(U) = \mathbb{B}(D(U))$. $U \leq \mathbb{F}_q^t$ is trivial if $U = \{0\}$ or $U = \mathbb{F}_q^t$.

Definition 5.5. $\mathbb{B} \in M(s \times t, q)^r$ is *stable*, if $\mathbb{B}$ is nondegenerate, and for every nontrivial subspace $U \leq \mathbb{F}_q^t$, $\frac{\dim(\mathbb{B}(U))}{\dim(U)} > \frac{s}{t}$.

Remark 5.6. In Definition 5.5, we can replace nondegenerate with image-nondegenerate, as the second condition already implies kernel-nondegenerate.

Stable matrix tuples admits interesting properties, and we shall utilizing the following lemma:

Lemma 5.7. *If $\mathbb{B}$ is stable, then any nonzero $(A, D) \in \mathrm{Adj}(\mathbb{B})$ is invertible.*

Proof. Take any $(A, D) \in \mathrm{Adj}(\mathbb{B})$. If $D = 0$, then $A\mathbb{B} = \mathbb{B}D = 0$. Since $\mathbb{B}$ is image-nondegenerate, A has to be 0.

Suppose now that D is not invertible nor 0, so $\mathrm{Ker}(D)$ is not $\{0\}$ nor $\mathbb{F}_q^t$. By $A\mathbb{B}(\mathrm{Ker}(D)) = \mathbb{B}D(\mathrm{Ker}(D)) = 0$, we know $A(\mathbb{B}(\mathrm{Ker}(D))) = 0$, and $\mathbb{B}(\mathrm{Ker}(D)) \leq \mathrm{Ker}(A)$. As $\mathbb{B}$ is stable, we have $\dim(\mathbb{B}(\mathrm{Ker}(D))) > \frac{s}{t}\dim(\mathrm{Ker}(D))$, thus $\dim(\mathrm{Ker}(A)) > \frac{s}{t}\dim(\mathrm{Ker}(D))$. On the other hand, $A\mathbb{B}(\mathbb{F}_q^t) = \mathbb{B}D(\mathbb{F}_q^t)$. Since $\mathbb{B}$ is image-nondegenerate, $\mathbb{B}(\mathbb{F}_q^t) = \mathrm{Im}(\mathbb{B}) =$

[4] On the other hand, $\mathrm{Iso}(\mathbb{G}, \mathbb{H})$ and $\mathrm{Aut}(\mathbb{G})$ for alternating tuples do not permit such easy linearization. Therefore testing isometry between $\mathbb{G}$ and $\mathbb{H}$ [IQ18] and computing a generating set for $\mathrm{Aut}(\mathbb{G})$ [BW12] requires new ideas, including exploiting the $*$-algebra structure.

$\mathbb{F}_q^s$. So $A\mathbb{B}(\mathbb{F}_q^t) = \mathrm{Im}(A) = \mathbb{B}(\mathrm{Im}(D))$. As $\mathbb{B}$ is stable, $\dim(\mathrm{Im}(A)) > \frac{s}{t}\dim(\mathrm{Im}(D))$ if $\mathrm{Im}(A)$ is non-trivial. It follows that $s = \dim(\mathrm{Im}(A)) + \dim(\mathrm{Ker}(A)) > \frac{s}{t}(\dim(\mathrm{Im}(D)) + \dim(\mathrm{Ker}(D))) = \frac{s}{t} \cdot t = s$. This is a contradiction, so D has to be invertible or $\mathrm{Im}(A)$ is trivial. If $\mathrm{Im}(A) = \{0\}$, by the similar discussion we can derive $D = 0$ as well, which is a contradiction. If $\mathrm{Im}(A) = \mathbb{F}_q^s$, $A\mathbb{B}$ is kernel-nondegenerate, so D has to be invertible, as otherwise $\mathbb{B}D$ would not be kernel-nondegenerate.

If D is invertible, then $\mathbb{B}D$ is image-nondegenerate. A has to be invertible, as otherwise $A\mathbb{B}$ would not be image-nondegenerate. □

Remark 5.8. Briefly speaking, the stable concept corresponds to the concept of *simple* as in the *representation theory of associative algebras*. Lemma 5.7 is an analog of the *Schur's lemma* there. Both the stable concept and the simple concept are special cases of the stable concept in geometric invariant theory [MFK94, KIN94], specialized to the left-right action of $\mathrm{GL}(s,q) \times \mathrm{GL}(t,q)$ on $M(s \times t, q)^r$, and the conjugation action of $\mathrm{GL}(s,q)$ on $M(s,q)^r$, respectively.

Specifically, consider a tuple of square matrices $\mathbb{B} \in M(s,q)^r$, which can be understood as a representation of an associative algebra with r generators. This representation is simple if and only if it does not have a non-trivial invariant subspace, that is $U \leq \mathbb{F}_q^s$, such that $\mathbb{B}(U) \leq U$. This amounts to say that there does not exist $A \in \mathrm{GL}(s,q)$ such that every B in $A\mathbb{B}A^{-1}$ is of the block-matrix form $\begin{bmatrix} B_1 & B_2 \\ 0 & B_3 \end{bmatrix}$, where $B_1 \in M(s',q)$, $1 \leq s' \leq s-1$. On the other hand, the stable concept can be rephrased as the following. $\mathbb{B} \in M(s \times t, q)^r$ is stable, if there do not exist $A \in \mathrm{GL}(s,q)$ and $D \in \mathrm{GL}(t,q)$ such that every $B \in A\mathbb{B}D^{-1}$ is of the form $\begin{bmatrix} B_1 & B_2 \\ 0 & B_3 \end{bmatrix}$ where B_1 is of size $s' \times t'$, $1 \leq t' \leq t-1$, such that $\frac{s'}{t'} \leq \frac{s}{t}$.

Lemma 5.7 can be understood as an analog of Schur's lemma, which states that if $\mathbb{B} \in M(s,q)^r$ is simple then a nonzero homomorphism $A \in M(s,q)$ of $\mathbb{B}$ (e.g. $A\mathbb{B}A^{-1} = \mathbb{B}$) has to be invertible.

Lemma 5.9. *Let $\mathfrak{A} \subseteq M(n,q)$ be a field containing λI_n, $\lambda \in \mathbb{F}_q$. Then $|\mathfrak{A}| \leq q^n$.*

Proof. (Communicated by G. Ivanyos.) Let $\mathfrak{A}$ be an extension field of $\mathbb{F}_q$ with extension degree d. Then $\mathbb{F}_q^n$ is an $\mathfrak{A}$-module, or in other words, a vector space over $\mathfrak{A}$. So $\mathbb{F}_q^n \cong \mathfrak{A}^m$

as vector spaces over $\mathfrak{A}$ for some $m \in \mathbb{N}$. Considering them as $\mathbb{F}_q$ vector spaces, we have $n = md$ so d divides n. It follows that $|\mathfrak{A}| = q^d \leq q^n$. □

By Lemma 5.7 and 5.9, we have the following.

Proposition 5.10. *If $\mathbb{B} \leq M(s \times t, q)^r$ is stable, then $|\mathrm{Adj}(\mathbb{B})| \leq q^s$.*

Proof. As $\mathbb{B}$ is stable, it is nondegenerate, so the projection of $\mathrm{Adj}(\mathbb{B}) \leq M(s,q) \oplus M(t,q)$ to $M(s,q)$ (naturally embedded in $M(s,q) \oplus M(t,q)$) along $M(t,q)$ is injective. By Lemma 5.7, the image of the projection is a *finite division algebra* over $\mathbb{F}_q$ containing $\lambda \mathrm{I}_s$. So by *Wedderburn's little theorem*, it is a field. By Lemma 5.9, the result follows. □

The following proposition about stable matrix spaces is also useful in our proof.

Proposition 5.11. *Given $\mathbb{B} = (B_1, \ldots, B_r) \in M(s \times t, q)^r$, let $\mathbb{B}^{\mathrm{t}} = (B_1^{\mathrm{t}}, \ldots, B_r^{\mathrm{t}}) \in M(t \times s, q)^r$. Then $\mathbb{B}$ is stable if and only if $\mathbb{B}^{\mathrm{t}}$ is stable.*

Proof. First we consider the nondegenerate part. If $u \in \mathbb{F}_q^s$ satisfies $\mathbb{B}(u) = \{0\}$, then it is easy to verify that $\mathbb{B}^{\mathrm{t}}(\mathbb{F}_q^s)$ is contained in the hyperplane defined by u, e.g. $u^{\mathrm{t}}(\mathbb{B}^{\mathrm{t}}(\mathbb{F}_q^s)) = 0$. If $\mathbb{B}(\mathbb{F}_q^t) \neq \mathbb{F}_q^s$, then there exists some $u \in \mathbb{F}_q^s$ such that $u^{\mathrm{t}}(\mathbb{B}(\mathbb{F}_q^t)) = 0$, so $u \in \mathrm{Ker}(\mathbb{B}^{\mathrm{t}})$. Therefore $\mathbb{B}$ is nondegenerate if and only if $\mathbb{B}^{\mathrm{t}}$ is nondegenerate.

In the following we assume that $\mathbb{B}$ is nondegenerate, and check nontrivial subspaces to show that $\mathbb{B}$ is not stable if and only if $\mathbb{B}^{\mathrm{t}}$ is not stable. This can be seen easily from the discussion in Remark 5.8. Assume $\mathbb{B}$ is not stable. There exist $A \in \mathrm{GL}(s,q)$ and $D \in \mathrm{GL}(t,q)$ such that every $B \in A\mathbb{B}D^{-1}$ is of the form $\begin{bmatrix} B_1 & B_2 \\ 0 & B_3 \end{bmatrix}$, where B_1 is of size $s' \times t'$, $1 \leq t' \leq t-1$, such that $\frac{s'}{t'} \leq \frac{s}{t}$. Note that $s' > 0$; otherwise $\mathbb{B}$ is degenerate, so $1 \leq s' \leq \frac{s}{t} \cdot t' < s$. Now consider $(A\mathbb{B}D^{-1})^{\mathrm{t}}$, the elements in which is of the form $\begin{bmatrix} B_1^{\mathrm{t}} & 0 \\ B_2^{\mathrm{t}} & B_3^{\mathrm{t}} \end{bmatrix}$. Note that B_3^{t} is of size $(t-t') \times (s-s')$ where $1 \leq s-s' \leq s-1$, $1 \leq t-t' \leq t-1$, and $\frac{t-t'}{s-s'} \leq \frac{t}{s}$. It indicates that $(A\mathbb{B}D^{-1})^{\mathrm{t}}$ is not stable, and $\mathbb{B}^{\mathrm{t}}$ being not stable follows. This concludes the proof. □

5.3.2 Random Alternating Matrix Spaces

We formally define the random models which will be used in the average-case analysis. As we have mentioned before, the linear algebraic Erdős-Rényi model is defined as follows:

Definition 5.12 (linear algebraic Erdős-Rényi model)**.** The linear algebraic Erdős-Rényi model, $\textsc{LinER}(n, m, q)$, is the uniform probability distribution over the set of dimension-m subspaces of $\Lambda(n, q)$, where each subspace is endowed with probability $1/\begin{bmatrix} N \\ m \end{bmatrix}_q$.

We also introduce the following model, which is much more useful in our analysis.

Definition 5.13 (Naive models for matrix tuples and matrix spaces)**.** The naive model for alternating tuples, $\textsc{NaiT}(n, m, q)$, is the probability distribution over the set of all m-tuples of $n \times n$ alternating matrices, where each tuple is endowed with probability $1/q^{Nm}$.

The naive model for alternating spaces, $\textsc{NaiS}(n, m, q)$, is the probability distribution over the set of alternating spaces in $\Lambda(n, q)$ of dimension no larger than m, where the probability at some $\mathcal{G} \leq \Lambda(n, q)$ of dimension $0 \leq d \leq m$ equals the number of m-tuples of $n \times n$ alternating tuples that represent $\mathcal{G}$, divided by q^{Nm}.

We now justify that working with the naive model suffices for the analysis even in the linear algebraic Erdős-Rényi model. Consider the following setting. Suppose we have $E(n, m, q)$, a property of dimension-m alternating spaces in $\Lambda(n, q)$, and wish to show that $E(n, m, q)$ holds with high probability in $\textsc{LinER}(n, m, q)$. $E(n, m, q)$ naturally induces $E'(n, m, q)$, a property of alternating tuples in $\Lambda(n, q)^m$ that span dimension-m alternating spaces. It is usually the case that there exists a property $F(n, m, q)$ of all m-alternating tuples in $\Lambda(n, q)^m$, so that $F(n, m, q)$ and $E'(n, m, q)$ coincide when restricting to those alternating tuples spanning dimension-m matrix spaces. If we could prove that $F(n, m, q)$ holds with high probability, then since a nontrivial fraction of m-tuples do span dimension-m spaces, we would get that $E(n, m, q)$ holds with high probability as well. The following proposition summarizes and makes precise the above discussion.

Proposition 5.14. *Let $E(n, m, q)$ and $F(n, m, q)$ be defined as above. Suppose $F(n, m, q)$ happens with probability $\geq 1 - f(n, m, q)$ in $\textsc{NaiT}(n, m, q)$, where $0 \leq f(n, m, q) < 1/4$. Then in $\textsc{LinER}(n, m, q)$, $E(n, m, q)$ happens with probability $> 1 - 4 \cdot f(n, m, q)$.*

Proof. The number of tuples for which $F(n,m,q)$ fails is no larger than $f(n,m,q)\cdot q^{Nm}$. Clearly the worst case for $E'(n,m,q)$ is when each of them spans an m-alternating space, so we focus on this case. Recall that $E'(n,m,q)$ is induced from a property of m-alternating spaces. That is, if two tuples span the same m-alternating space, then either both of them satisfy $E'(n,m,q)$, or neither of them satisfies $E'(n,m,q)$. Since the number of m-alternating spaces for which $E(n,m,q)$ fails is $\leq f(n,m,q)\cdot\frac{q^{Nm}}{(q^m-1)(q^m-q)\ldots(q^m-q^{m-1})}$. The fraction of m-alternating spaces for which $E(n,m,q)$ fails is no larger than $f(n,m,q)\cdot\frac{q^{Nm}}{(q^N-1)(q^N-q)\ldots(q^N-q^{N-m+1})}<4\cdot f(n,m,q)$, where 4 comes from Lemma 5.4 2. □

Random matrix spaces. For $s,t,r\in\mathbb{N}$, we can define the Erdős-Rényi model for bipartite graphs on the vertex set $[s]\times[t]$ with edge set size r by taking every subset of $[s]\times[t]$ of size r with probability $\binom{st}{r}^{-1}$. Analogously, we can define the following in the matrix space and matrix tuple setting.

Definition 5.15. 1. The *bipartite linear algebraic Erdős-Rényi model*, BipLinER$(s\times t,r,q)$, is the probability distribution over the set of all r-matrix space in $M(s\times t,q)$, where each matrix space is endowed with probability $1/\begin{bmatrix}st\\r\end{bmatrix}_q$.

2. The *bipartite naive model for matrix tuples*, BipNaiT$(s\times t,r,q)$, is the probability distribution over the set of all r-matrix tuple in $M(s\times t,q)^r$, where each matrix space is endowed with probability $1/q^{str}$.

5.4 Proof of the Main Algorithm

We first define the property $F(n,m,q,r)$ of m-alternating tuples in $\Lambda(n,q)^m$ for the average-case analysis, where r is the parameter for individualization. To lower bound the probability, we will in turn work with a stronger property $F'(n,m,q,r)$, which will be also defined. Utilizing $F(n,m,q,r)$, we then expand the high-level idea displayed in subSection 5.2.3 into a rigorous algorithm.

5.4.1 Properties of Alternating Spaces and Alternating Tuples

An m-alternating space $\mathcal{G} \leq \Lambda(n, q)$ induces

$$\mathcal{B}'_{\mathcal{G}} := \{\begin{bmatrix} e_1 & \cdots & e_r \end{bmatrix}^{\mathrm{t}} G \begin{bmatrix} e_{r+1} & \cdots & e_n \end{bmatrix} : G \in \mathcal{G}\} \leq M(r \times (n-r), q), \tag{5.13}$$

of which the dimension is at most m. Define the right-side equivalence of $\mathcal{B}'_{\mathcal{G}}$

$$\mathrm{RAut}(\mathcal{B}'_{\mathcal{G}}) := \{A \in \mathrm{GL}(n-r, q) : \mathcal{B}'_{\mathcal{G}} A = \mathcal{B}'_{\mathcal{G}}\}. \tag{5.14}$$

An element in $\mathrm{RAut}(\mathcal{B}'_{\mathcal{G}})$ is called a right-side equivalence of $\mathcal{B}'_{\mathcal{G}}$.

Definition 5.16. Let $E'(n, m, q, r)$ be a property of m-alternating spaces in $\Lambda(n, q)$, defined as follows. Given an m-alternating space $\mathcal{G}$ in $\Lambda(n, q)$, let $\mathcal{B}'_{\mathcal{G}}$ be the matrix space in $M(r \times (n-r), q)$ defined as in Equation (5.13). $\mathcal{G} \in E'(n, m, q, r)$ if and only if $|\mathrm{RAut}(\mathcal{B}')| \leq q^{n-r}$.

Right-side equivalence is a useful concept which leads to our algorithm (as seen in Section 5.2.2), but what we actually need is the following linearization of $\mathrm{RAut}(\mathcal{B}')$.

Definition 5.17. Let $E(n, m, q, r)$ be a property of m-alternating spaces in $\Lambda(n, q)$, defined as follows. Given an m-alternating space $\mathcal{G}$ in $\Lambda(n, q)$, let $\mathcal{B}'_{\mathcal{G}}$ be the matrix space in $M(r \times (n-r), q)$ defined as in Equation (5.13). $\mathcal{G} \in E(n, m, q, r)$, if and only if $|\{A \in M(n-r, q) : \mathcal{B}'_{\mathcal{G}} A \leq \mathcal{B}'_{\mathcal{G}}\}| \leq q^{n-r}$.

Clearly, $\mathcal{G} \in E'(n, m, q, r)$ implies $\mathcal{G} \in E(n, m, q, r)$. Thus, lower bound on the probability of $\mathcal{G} \notin E(n, m, q, r)$ implies lower bound on the probability of $\mathcal{G} \notin E'(n, m, q, r)$.

We define a property $F(n, m, q, r)$ for alternating tuples that corresponds to $E(n, m, q, r)$. Given $\mathbb{G} = (G_1, \ldots, G_m) \in \Lambda(n, q)^m$, we can analogously construct a matrix tuple

$$\mathbb{B}'_{\mathbb{G}} := ([e_1, \ldots, e_r]^t G_1 [e_{r+1}, \ldots, e_n], \ldots, [e_1, \ldots, e_r]^t G_m [e_{r+1}, \ldots, e_n]) \in M(r \times (n-r), q)^m. \tag{5.15}$$

Definition 5.18. Let $F(n, m, q, r)$ be a property of m-alternating tuples in $\Lambda(n, q)^m$, defined as follows. Given an m-alternating tuple $\mathbb{G}$ in $\Lambda(n, q)^m$, let $\mathbb{B}'_{\mathbb{G}}$ be the m-matrix

tuple in $M(r \times (n-r), q)^m$ defined as in Equation (5.15). $\mathbb{G} \in F(n, m, q, r)$, if and only if $|\{A \in M(n-r, q) : \exists D \in M(m, q),\ \mathbb{B}'_{\mathbb{G}} A = \mathbb{B}'^{D}_{\mathbb{G}}\}| \leq q^{n-r}$.

It is not hard to see that $F(n, m, q, r)$ is a proper extension of $E(n, m, q, r)$.

Proposition 5.19. *Suppose $\mathbb{G} \in \Lambda(n, q)^m$ represents an m-alternating space $\mathcal{G} \leq \Lambda(n, q)$. Then $\mathbb{G}$ is in $F(n, m, q, r)$ if and only if $\mathcal{G}$ is in $E(n, m, q, r)$.*

Proof. Let $\mathcal{B}'_{\mathcal{G}}$ and $\mathbb{B}'_{\mathbb{G}}$ be the matrix space and matrix tuple defined as above for $\mathcal{G}$ and $\mathbb{G}$, respectively. Clearly $\mathbb{B}'_{\mathbb{G}}$ represents $\mathcal{B}'_{\mathcal{G}}$, so $\mathbb{B}'_{\mathbb{G}} A$ represents $\mathcal{B}'_{\mathcal{G}} A$. Finally note that $\mathbb{B}'_{\mathbb{G}} A = \mathbb{B}'^{D}_{\mathbb{G}}$ for some $D \in M(n, q)$ if and only if the linear span of $\mathbb{B}'_{\mathbb{G}} A$ is contained in the linear span of $\mathbb{B}'_{\mathbb{G}}$, that is $\mathcal{B}'_{\mathcal{G}} A \leq \mathcal{B}'_{\mathcal{G}}$. □

Instead of working with $\mathbb{B}'_{\mathbb{G}}$ and $\{A \in \mathrm{GL}(n-r, q) : \exists D \in \mathrm{GL}(m, q), \mathbb{B}'_{\mathbb{G}} A = \mathbb{B}'^{D}_{\mathbb{G}}\}$, it is more convenient to flip $\mathbb{B}'_{\mathbb{G}}$, an m-matrix tuple of size $r \times (n-r)$, to get $\mathbb{B}_{\mathbb{G}}$, an r-matrix tuple of size $(n-r) \times m$. Then the collection $\{A \in M(n-r, q) : \exists D \in M(m, q), \mathbb{B}'_{\mathbb{G}} A = \mathbb{B}'^{D}_{\mathbb{G}}\}$ is exactly $\{A \in M(n-r, q) : \exists D \in M(m, q), A\mathbb{B}_{\mathbb{G}} = \mathbb{B}_{\mathbb{G}} D\}$, which is closely related to the adjoint algebra concept for matrix tuples as defined in Equation (5.11). Let $\pi_1 : M(n-r, q) \oplus M(m, q) \to M(n-r, q)$ be the projection to the first component along the second. $\{A \in M(n-r, q) : \exists D \in M(m, q), A\mathbb{B}_{\mathbb{G}} = \mathbb{B}_{\mathbb{G}} D\}$ is then just $\pi_1(\mathrm{Adj}(\mathbb{B}_{\mathbb{G}}))$. We may reformulate definition 5.18 as the following.

Lemma 5.20 (Definition 5.18, alternative formulation.)**.** *$F(n, m, q, r)$ is a property of m-alternating tuples in $\Lambda(n, q)^m$, defined as follows. Given an m-alternating tuple $\mathbb{G}$ in $\Lambda(n, q)^m$, let $\mathbb{B}_{\mathbb{G}}$ be the r-matrix tuple in $M((n-r) \times m, q)^r$ defined as above. $\mathbb{G} \in F(n, m, q, r)$ if and only if $|\pi_1(\mathrm{Adj}(\mathbb{B}_{\mathbb{G}}))| \leq q^{n-r}$.*

Our algorithm will be based on the property $F(n, m, q, r)$. To show that $F(n, m, q, r)$ holds with high probability though, we turn to study the following stronger property.

Definition 5.21. Let $F'(n, m, q, r)$ be a property of m-alternating tuples in $\Lambda(n, q)^m$, defined as follows. Given an m-alternating tuple $\mathbb{G}$ in $\Lambda(n, q)^m$, let $\mathbb{B}_{\mathbb{G}}$ be the r-matrix tuple in $M((n-r) \times m, q)^r$ defined as above. $\mathbb{G} \in F'(n, m, q, r)$ if and only if $|\mathrm{Adj}(\mathbb{B}_{\mathbb{G}})| \leq q^{n-r}$.

Clearly $\mathbb{G} \in F'(n,m,q,r)$ implies $\mathbb{G} \in F(n,m,q,r)$. To show that $\mathbb{G} \in F'(n,m,q,r)$ holds with high probability, Proposition 5.10 immediately implies the following, which directs us to make use of the stable property.

Proposition 5.22. *Let $\mathbb{G}$ and $\mathbb{B}_{\mathbb{G}}$ be defined as above. If $\mathbb{B}$ is stable, then $\mathbb{G} \in F'(n,m,q,r)$.*

5.4.2 Estimate the Probability of $\mathbb{G} \in F'(n,m,q,r)$

We now show that $\mathbb{G} \in F'(n,m,q,r)$ holds with high probability in $\textsc{NaiT}(n,m,q)$, where $m = cn$ for some positive constant c. The integer r is chosen so that $r \geq 4 \cdot \frac{n-r}{m}$ if $n-r \geq m$, and $r \geq 4 \cdot \frac{m}{n-r}$ if $m \geq n-r$. When n is large enough this is always possible. Let $s := n-r$ and $t := m$. By Proposition 5.22, to show $F'(n,m,q,r)$ holds with high probability, we can alternatively show that for most $\mathbb{G} \in \textsc{NaiT}(n,m,q)$, the corresponding $\mathbb{B}_{\mathbb{G}}$ in $M(s \times t, q)^r$ is stable. A simple observation is that $\textsc{NaiT}(n,m,q)$ induces $\textsc{BipNaiT}(s \times t, r, q)$ obtained by flipping the upper right $s \times t$ corners of the alternating matrices (see Figure 5.2). So we only need to estimate the probability of $\mathbb{B} \in \textsc{BipNaiT}(s \times t, r, q)$ being stable.

By our choice of r, we obtain an r-matrix tuple in $M(s \times t, q)$ with $r \geq 4 \cdot \frac{\max(s,t)}{\min(s,t)}$. By Proposition 5.11, we know $\Pr[\mathbb{B} \in \textsc{BipNaiT}(s \times t, r, q) \text{ is stable}] = \Pr[\mathbb{C} \in \textsc{BipNaiT}(t \times s, r, q) \text{ is stable}]$ via the transpose map. So it is enough to consider the case when $s \geq t$.

Proposition 5.23. *Give positive integers s, t, and r such that $s \geq t \geq 16$, $\frac{s}{t} := b \geq 1$ (b is a constant), and $r \geq 4b$. Then $\mathbb{B}$ is stable with probability $1 - \frac{1}{q^{\Omega(t)}}$ in $\textsc{BipNaiT}(s \times t, r, q)$, where $\Omega(t)$ hides a positive constant depending on b.*

Proof. We will upper bound the probability of $\mathbb{B}$ being *not* stable in $\textsc{BipNaiT}(s \times t, r, q)$:

$$P := \Pr[\mathbb{B} \text{ is degenerate, or } \exists U \leq \mathbb{F}_q^t, U \text{ non-trivial}, \frac{\dim(\mathbb{B}(U))}{\dim(U)} \leq \frac{s}{t}]. \tag{5.16}$$

By the union bound, we have:

$$P \leq \sum_{\substack{U \leq \mathbb{F}_q^t, \\ 1 \leq \dim(U) \leq t-1}} \Pr[\frac{\dim(\mathbb{B}(U))}{\dim(U)} \leq \frac{s}{t}] + \Pr[\mathbb{B} \text{ is degenerate}]. \tag{5.17}$$

About $\mathbb{B}$ being degenerate. By remark 5.6, we only need to bound the image-degenerate case. Note that the columns of B_i's form a linear basis of $\mathrm{Im}(\mathbb{B})$. By forming an $s \times rt$ matrix $A = \begin{bmatrix} B_1 & B_2 & \cdots & B_r \end{bmatrix}$, this amounts to upper bound the probability that $\Pr[rk(A) < s | A \in M(s \times rt, q)]$. As $rt \geq 4bt = 4s$,

$$\Pr[rk(A) = s | A \in M(s \times rt, q)] \geq \Pr[rk(A) = s | A \in M(s \times 4s, q)] \geq 1 - s/q^{3s+1},$$

where the last inequality is from Lemma 5.4 2. So we have $\Pr[\mathbb{B} \text{ is image-degenerate}] \leq 1/q^{\Omega(t)}$ since $s = bt$.

Reduce to work with nontrivial subspaces with the same dimension. Now we focus on upper bound $\sum_{U \leq \mathbb{F}_q^t, 1 \leq \dim(U) \leq t-1} \Pr[\frac{\dim(\mathbb{B}(U))}{\dim(U)} \leq \frac{s}{t}]$. For a nontrivial subspace $U \leq \mathbb{F}_q^t$, let

$$B_U := \{\mathbb{B} \in M(s \times t, q)^r : \frac{\dim(\mathbb{B}(U))}{\dim(U)} \leq \frac{s}{t}\}. \tag{5.18}$$

For two subspace $U_1, U_2 \leq \mathbb{F}_q^t$ of the same dimension $d \in \{1, \ldots, t-1\}$, we claim that $|B_{U_1}| = |B_{U_2}|$. Let $X \in \mathrm{GL}(t, q)$ be any invertible matrix such that $X(U_2) = U_1$, and consider the map $T_X : M(s \times t, q)^r \to M(s \times t, q)^r$ defined by sending $\mathbb{B}$ to $\mathbb{B}X$. It is easy to see that T_X is a bijection between B_{U_1} and B_{U_2}. The claim then follows and we have

$$\Pr_{\mathbb{B}}[\frac{\dim(\mathbb{B}(U_1))}{\dim(U_1)} \leq \frac{s}{t}] = \Pr_{\mathbb{B}}[\frac{\dim(\mathbb{B}(U_2))}{\dim(U_2)} \leq \frac{s}{t}]. \tag{5.19}$$

Setting $U_d = \langle e_1, \ldots, e_d \rangle$, we apply union bound again and derive

$$\sum_{\substack{U \leq \mathbb{F}_q^t, \\ 1 \leq \dim(U) \leq t-1}} \Pr[\frac{\dim(\mathbb{B}(U))}{\dim(U)} \leq \frac{s}{t}] = \sum_{1 \leq d \leq t-1} \begin{bmatrix} t \\ d \end{bmatrix}_q \Pr[\dim(\mathbb{B}(U_d)) \leq \frac{s}{t} \cdot d]. \tag{5.20}$$

Upper bound $\begin{bmatrix} t \\ d \end{bmatrix}_q \Pr[\dim(\mathbb{B}(U_d)) \leq \frac{s}{t} \cdot d]$. For $d = 1, \ldots, t-1$, denote $P_d := \Pr[\dim(\mathbb{B}(U_d)) \leq \frac{s}{t} \cdot d]$. Note that for any matrix $B \in M(s \times t, q)$, $B(U_d)$ is spanned by the first d column vectors of B. So for $\mathbb{B} = (B_1, \ldots, B_r) \in M(s \times t, q)^r$, $\mathbb{B}(U_d)$ is spanned by the first d columns of B_i's. Collect those columns to form a matrix $A \in M(s \times rd, q)$, we can derive

$$P_d = \Pr[\dim(\mathbb{B}(U_d)) \leq bd] = \Pr[rk(A) \leq \lfloor bd \rfloor | A \in M(s \times rd, q)]. \tag{5.21}$$

Note that we substituted bd with $\lfloor bd \rfloor$ as that does not change the probability. Equation (5.21) suggests the following to upper bound of P_d. For A to be of rank $\leq \lfloor bd \rfloor$, there must exist $\lfloor bd \rfloor$ columns such that other columns are linear combinations of them. So we enumerate all subsets of $\{1, \ldots, rd\}$ of size $\lfloor bd \rfloor$ to locate the $\lfloor bd \rfloor$ linear independent columns. Fill in these columns arbitrarily, and fill the other columns as linear combinations of them to construct the matrix A. This procedure suggests an upper bound on P_d:

$$P_d \leq \frac{\binom{rd}{\lfloor bd \rfloor} \cdot q^{s\lfloor bd \rfloor} \cdot q^{\lfloor bd \rfloor (rd - \lfloor bd \rfloor)}}{q^{srd}}. \tag{5.22}$$

When $1 \leq d \leq t/2$, we further derive that

$$\begin{aligned} \begin{bmatrix} t \\ d \end{bmatrix}_q P_d &\leq \frac{\binom{rd}{\lfloor bd \rfloor} \cdot q^{s\lfloor bd \rfloor} \cdot q^{\lfloor bd \rfloor (rd - \lfloor bd \rfloor)} \cdot \begin{bmatrix} t \\ d \end{bmatrix}_q}{q^{srd}} \\ &\leq \frac{q^{rd} \cdot q^{sbd} \cdot q^{bd(rd-bd)} \cdot q^{td}}{q^{srd}} \\ &\leq \frac{1}{q^{(sr-sb-t-r)d - b(r-b)d^2}}, \end{aligned} \tag{5.23}$$

where in the second inequality, we use the fact that $\binom{rd}{\lfloor bd \rfloor} \leq 2^{rd} \leq q^{rd}$, $\begin{bmatrix} t \\ d \end{bmatrix}_q \leq q^{td}$ and $\lfloor bd \rfloor (rd - \lfloor bd \rfloor) \leq bd(rd - bd)$ since $r \geq 4b$.

Let $f(d) = (sr - sb - t - r)d - b(r-b)d^2$. It is easy to see that $f(d)$ achieves minimum at $d = 1$ or $d = t/2$ when $1 \leq d \leq \frac{t}{2}$. We compute that $f(1) = (br - b^2 - 1)t + b^2 - r - br$ and $f(\frac{t}{2}) = (\frac{1}{4}br - \frac{1}{4}b^2 - \frac{1}{2})t^2 - \frac{1}{2}rt$. Since $r \geq 4b$ and $b \geq 1$, $br - b^2 - 1 \geq 3b^2 - 1 > 0$ and $\frac{1}{4}br - \frac{1}{4}b^2 - \frac{1}{2} \geq \frac{3}{4}b^2 - \frac{1}{2} > 0$, these two lower bounds then yield that

$$\begin{bmatrix} t \\ d \end{bmatrix}_q P_d \leq \frac{1}{q^{\Omega(t)}} \ \forall \ 1 \leq d \leq t/2. \tag{5.24}$$

When $t/2 \leq d \leq t - 3$, we replace $\begin{bmatrix} t \\ d \end{bmatrix}_q$ by $\begin{bmatrix} t \\ t-d \end{bmatrix}_q$ in inequality (5.23) and obtain

$$\begin{aligned} \begin{bmatrix} t \\ d \end{bmatrix}_q P_d &\leq \frac{\binom{rd}{\lfloor bd \rfloor} \cdot q^{s\lfloor bd \rfloor} \cdot q^{\lfloor bd \rfloor (rd - \lfloor bd \rfloor)} \cdot \begin{bmatrix} t \\ t-d \end{bmatrix}_q}{q^{srd}} \leq \frac{q^{rd} \cdot q^{sbd} \cdot q^{bd(rd-bd)} \cdot q^{t(t-d)}}{q^{srd}} \\ &\leq \frac{1}{q^{(sr-sb+t-r)d - b(r-b)d^2 - t^2}}. \end{aligned}$$

It can be seen that the function $g(d) = (sr - sb + t - r)d - b(r-b)d^2 - t^2$ achieves minimum at $d = t/2$ or $d = t-3$ when $t/2 \le d \le t-3$. We know that $g(\frac{t}{2}) = f(\frac{t}{2}) = (\frac{1}{4}br - \frac{1}{4}b^2 - \frac{1}{2})t^2 - \frac{1}{2}rt$ and compute $g(t-3) = (3br - 3b^2 - r - 3)t + 3r + 9b^2 - 9br$. Since $r \ge 4b$ and $b \ge 1$, $\frac{3}{4}b^2 - \frac{1}{2} > 0$ and $9b^2 - 4b - 3 > 0$ when $b \ge 1$, these two lower bounds then yield that

$$\begin{bmatrix} t \\ d \end{bmatrix}_q P_d \le \frac{1}{q^{\Omega(t)}}, \ \forall\ t/2 \le d \le t-3. \tag{5.25}$$

When $d = t-2$, recall that $P_{t-2} = \Pr[rk(A) \le b(t-2) | A \in M(s \times r(t-2), q)]$. Since $t \ge 16$ (i.e. $s \ge 16b$), $r(t-2) \ge 4b(t-2) \ge \lceil \frac{7}{2}s \rceil$. Also note that $b(t-2) < bt = s$. Therefore

$$P_{t-2} \le \Pr[rk(A) < s | A \in M(s \times \lceil \frac{7}{2}s \rceil, q)] \le s/q^{\frac{5}{2}s+1} \tag{5.26}$$

by Lemma 5.4 2. Then

$$\begin{bmatrix} t \\ t-2 \end{bmatrix}_q P_{t-2} \le \frac{sq^{2t}}{q^{\frac{5}{2}s+1}} = \frac{bt}{q^{(\frac{5}{2}b-2)t+1}} \le \frac{1}{q^{\Omega(t)}} \tag{5.27}$$

The case when $d = t-1$ is similar, and we can obtain

$$\begin{bmatrix} t \\ t-1 \end{bmatrix}_q P_{t-1} \le \frac{1}{q^{\Omega(t)}}. \tag{5.28}$$

We concludes the proof by combining inequalities (5.24), (5.25), (5.27), (5.28). □

5.4.3 Algorithm Analysis

We now present a detailed description and analysis of the main algorithm as Algorithm 2, and prove Theorem 5.1.

As described in Section 5.2, we first discuss the individualization step. Recall that an r-individualization is a direct sum decomposition of $\mathbb{F}_q^n = L \oplus R$, where L and R are subspaces of $\mathbb{F}_q$ and L is equipped with an ordered basis $(v_1, \ldots, v_r)$. In our main algorithm, we need to enumerate all r-individualizations, and the following proposition realizes this.

Proposition 5.24. *There is a deterministic algorithm that lists all r-individualizations in $\mathbb{F}_q^n$ in time $q^{O(rn)}$. Each individualization $\mathbb{F}_q^n = L \oplus R$ is represented as an invertible*

matrix $\begin{bmatrix} v_1 & \cdots & v_r & u_1 & \cdots & u_{n-r} \end{bmatrix} \in \mathrm{GL}(n,q)$*, where* $(v_1,\ldots,v_r)$ *is the chosen ordered basis of* L *and* $\{u_1,\ldots,u_{n-r}\}$ *forms a linear basis of* R*.*

Proof. We present the algorithm formally as Algorithm 1. Let us first outline what the

Algorithm 1 Algorithm for listing r-individualizations in $\mathbb{F}_q^n$

Input: A positive integer r.
Output: a list $\mathcal{T}$ of invertible matrices of size n.
1: $\mathcal{T} \leftarrow \emptyset$
2: Listing all r-tuples of linearly independent vectors.
3: **for all** dimension-r $L \leq \mathbb{F}_q^n$ with an ordered basis $(v_1,\ldots,v_r)$ **do**
4: $\quad(u_1,\ldots,u_{n-r}) \leftarrow$ An ordered basis of a complement of L
5: $\quad$**for all** $(n-r)$-tuples of vectors $(w_1,\ldots,w_{n-r})$ from L **do**
6: $\quad\quad\mathcal{T} \leftarrow [v_1,\ldots,v_r,u_1+w_1,\ldots,u_{n-r}+w_{n-r}]$
7: $\quad$**end for**
8: **end for**

algorithm does. Line 2 lists all r-tuples of linearly independent vectors, which can be viewed as enumerating all dimension-r subspaces L equipped with an ordered basis. This steps can be done easily in time $q^{rn}\cdot\mathrm{poly}(n,\log q)$ as there are $(q^n-1)(q^n-q)\cdots(q^n-q^{r-1}) \leq q^{rn}$ such r-tuples. From Line 3 to 8 the algorithm computes all complements of L, and represent each complement R by an $(n-r)$-tuple of vectors which span R. Collect these two tuples to form an $n\times n$ invertible matrix. To compute the complements, we first compute one linear basis of an arbitrary complement of L, by solving the linear equations defined by v_i's, e.g. $v_i^{\mathrm{t}}x = 0$ where x is a vector of variables for $i=1,\ldots,r$. Denote the solution by the $(n-r)$-tuple of (linear independent) vectors $(u_1,\ldots,u_{n-r})$. The linear spans of the $(n-r)$-tuples $(u_1+w_1,\ldots,u_{n-r}+w_{n-r})$ go over all complements of L when $(w_1,\ldots,w_{n-r})$ go over all $(n-r)$-tuples of vectors from L, which can be done in time $q^{r(n-r)}$. The total cost is $q^{2rn-r^2} \in q^{O(rn)}$. Other steps can be achieved via linear algebra computations. This concludes the proof. □

Algorithm 1 produces a list $\mathcal{T}$ of invertible matrices in $\mathrm{GL}(n,q)$. More precisely, matrices in $\mathcal{T}$ admits a structural property.

Proposition 5.25. *For every* $A = \begin{bmatrix} v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \end{bmatrix} \in \mathrm{GL}(n, q)$, *there uniquely exists* $A_1 = \begin{bmatrix} v_1 & \cdots & v_r & u_1 & \cdots & u_{n-r} \end{bmatrix} \in \mathcal{T}$, *where* $\langle v_{r+1}, \ldots, v_n \rangle = \langle u_1, \ldots, u_{n-r} \rangle$, *and* $A_0 \in \mathrm{GL}(n-r, q)$, *such that* $A = A_1 \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & A_0 \end{bmatrix}$.

Proof. Notes that every invertible matrix $A = \begin{bmatrix} v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \end{bmatrix}$ can be viewed as a change-of-basis matrix, mapping e_i to v_i for $i \in [n]$ where $V = \{v_1, \ldots, v_n\}$ is a linear basis of $\mathbb{F}_q^n$. Clearly, we can find $A_1 = \begin{bmatrix} v'_1 & \cdots & v'_r & u'_1 & \cdots & u'_{n-r} \end{bmatrix} \in \mathcal{T}$, where $\langle v_{r+1}, \ldots, v_n \rangle = \langle u_1, \ldots, u_{n-r} \rangle$. This is because matrices in $\mathcal{T}$ coincide with r-individualizations of $\mathbb{F}_q^n$, and A_1 corresponding to the one where $\mathbb{F}_q^n = \langle (v'_1, \ldots, v'_r) \rangle \oplus \langle u'_1, \ldots, u'_{n-r} \rangle$. We simply pick A_1 such that $v'_i = v_i$ for $i \in [r]$ and $\langle u'_1, \ldots, u'_{n-r} \rangle = \langle v_{r+1}, \ldots, v_n \rangle$. Note that A_1 is responsible to map e_i to v_i for $i \in [r]$, while the mapping from e_i to v_i for $i \in [n] \setminus [r]$ cannot be achieved by A_1, as it is only responsible to map the subspace $\langle e_1, \ldots, e_r \rangle$ to the subspace $\langle u'_1, \ldots, u'_{n-r} \rangle$. To further realize the basis change, we need another invertible A_0 and the decomposition $A = A_1 \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & A_0 \end{bmatrix}$ follows. The uniqueness can be derived as A_1 need to be unique. □

Now we formally present our average-case algorithm as Algorithm 2, with some implementation details.

Line 7. $\mathbb{B}_{\mathbb{G}}$ is constructed by taking the upper-right $r \times (n-r)$ corners of G_i's to get an m-matrix tuple $\mathbb{B}'_{\mathbb{G}} \in M(r \times (n-r), q)^m$, and flipping $\mathbb{B}'_{\mathbb{G}}$ to obtain an r-matrix tuple $\mathbb{B}_{\mathbb{G}} \in M((n-r) \times m, q)^r$. See also Figure 5.2 and 5.3.

Line 14. $\mathbb{C}_{\mathbb{H}}$ is constructed as follows. For a given invertible matrix $A_1 \in \mathcal{T}$. Let $\mathbb{H}_1 = A_1^{\mathrm{t}} \mathbb{H} A_1$. Then perform the same procedure as in Line 7 for $\mathbb{H}_1$.

Line 9, 16. π_1 denotes the projection of $M(n-r, q) \oplus M(m, q)$ to $M(n-r, q)$ along $M(m, q)$.

Line 22. To test whether $A_0 = A_2 A_1^{-1}$ is an isometry between $\mathcal{G}$ and $\mathcal{H}$, we test whether $A_0^{\mathrm{t}} \mathbb{G} A_0$ and $\mathbb{H}$ span the same alternating space.

Algorithm 2 Average-case algorithm for AltMatSpIso

Input: Two m-alternating tuples $\mathbb{G} = (G_1, \ldots, G_m)$ and $\mathbb{H} = (H_1, \ldots, H_m)$ in $\Lambda(n,q)^m$ representing m-alternating spaces $\mathcal{G}, \mathcal{H} \leq \Lambda(n,q)$, respectively. $m = cn$ for some constant c, and n is large enough (larger than some fixed function of c).

Output: Either certify that $\mathcal{G}$ does not satisfy $F(n,m,q,r)$, or a set S consisting of all isometries between $\mathcal{G}$ and $\mathcal{H}$. (If $S = \emptyset$ then $\mathcal{G}$ and $\mathcal{H}$ are not isometric.)

$S \leftarrow \emptyset$.
if $n - r \geq m$ **then**
  Set $r \in \mathbb{N}$ such that $r \geq 4 \cdot \frac{n-r}{m}$
else
  Set $r \in \mathbb{N}$ such that$r \geq 4 \cdot \frac{m}{n-r}$
end if
Construct $\mathbb{B}_{\mathbb{G}} \in M((n-r) \times m, q)^r$ as described before definition 5.21.
Compute a linear basis of $\mathrm{Adj}(\mathbb{B}_{\mathbb{G}})$.
if $\dim(\pi_1(\mathrm{Adj}(\mathbb{B}_{\mathbb{G}}))) > n - r$ **then**
  Return "$\mathcal{G}$ does not satisfy $F(n,m,q,r)$."
else
  Produce $\mathcal{T}$ which lists all r-individualizations in $\mathbb{F}_q^n$ as invertible matrices of size n by Algorithm 1.
end if
for all $A_1 \in \mathcal{T}$ **do**
  Construct $\mathbb{C}_{\mathbb{H}} \in M((n-r) \times m, q)^r$ with respect to the r-individualization represented by A_1.
  Compute a linear basis of $\mathrm{Adj}(\mathbb{B}_{\mathbb{G}}, \mathbb{C}_{\mathbb{H}})$.
  if $\dim(\pi_1(\mathrm{Adj}(\mathbb{B}_{\mathbb{G}}, \mathbb{C}_{\mathbb{H}}))) > n - r$ **then**
    Go to the next r-individualization.
  else
    for all $A_0 \in \pi_1(\mathrm{Adj}(\mathbb{B}_{\mathbb{G}}, \mathbb{C}_{\mathbb{H}}))$ **do**
      if A_0 is invertible **then**
        $A_2 \leftarrow \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & A_0 \end{bmatrix}$
        if $A = A_2 A_1^{-1}$ is an isometry between $\mathcal{G}$ and $\mathcal{H}$ **then**
          $S \leftarrow A$
        else
          Go to the next $A_0 \in \pi_1(\mathrm{Adj}(\mathbb{B}_{\mathbb{G}}, \mathbb{C}_{\mathbb{H}}))$.
        end if
      else
        Go to the next $A_0 \in \pi_1(\mathrm{Adj}(\mathbb{B}_{\mathbb{G}}, \mathbb{C}_{\mathbb{H}}))$.
      end if
    end for
  end if
end for

Complexity analysis. It is straightforward to verify that the algorithm runs in time $q^{O(n)}$: the multiplicative cost of enumerating r-individualization is at most q^{2rn-r^2}, and the multiplicative cost of enumerating $\pi_1(\mathrm{Adj}(\mathbb{B}_{\mathbb{G}}, \mathbb{C}_{\mathbb{H}}))$ starting from line 14 is at most q^{n-r}. All the other steps are basic tasks in linear algebra so can be carried out efficiently. When $m = cn$ and n larger than a fixed function of c, all but at most $1/q^{\Omega(n)}$ fraction of $\mathcal{G} \leq \Lambda(n, q)$ satisfy $F(n, m, q, r)$ by propositions 5.23, 5.22, and 5.14. Note that $\Omega(n)$ hides a constant depending on c.

To see the correctness, first note that by the test step in line 23, only isometries will be added to S. So we only need to verify that, when $\mathcal{G}$ is in $F(n, m, q, r)$, then every isometry $A \in \mathrm{Iso}(\mathcal{G}, \mathcal{H})$ will be added to S. Recall that $A \in \mathrm{GL}(n, q)$ is an isometry from $\mathcal{G}$ to $\mathcal{H}$ if and only if there exists $D \in \mathrm{GL}(m, q)$ such that ${A_0}^{\mathrm{t}}\mathbb{G}A_0 = \mathbb{H}^D$, which is equivalent to $\mathbb{G} = (A^{-1})^{\mathrm{t}}(\mathbb{H}^D)A^{-1}$. By Proposition 5.25, $A^{-1} \in \mathrm{GL}(n, q)$ can be written uniquely as $A^{-1} = A_1 A_2^{-1}$ where A_1 represents an r-individualization, and A_2 admits the block-diagonal form $\begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & A_0 \end{bmatrix}$ where $A_0 \in \mathrm{GL}(n - r, q)$. As we enumerate all the r-individualization, the invertible matrix A_1 will be encountered by Algorithm 1. The invertible matrix A_0 will be encountered when enumerating the elements of $\pi_1(\mathrm{Adj}(\mathbb{B}_{\mathbb{G}}, \mathbb{C}_{\mathbb{H}}))$ and forms A_2 as shown in line 22. Since $A_2^{\mathrm{t}}\mathbb{G}A_2 = A_1^{\mathrm{t}}(\mathbb{H}^D)A_1 = (A_1^{\mathrm{t}}\mathbb{H}A_1)^D$, we derive that $A \in \mathrm{Iso}(\mathcal{G}, \mathcal{H})$ and added to S.

5.5 Dynamic Programming

In this section, we utilize Luks' dynamic programming technique [Luk99] for GraphIso to derive a deterministic algorithm for AltMatSpIso. Note that, for a matrix group $G \leq \mathrm{GL}(n, q)$ (specified by a list of generators), we can view G as a permutation group on the domain $\mathbb{F}_q^n$. So basic tasks like *membership testing* and *pointwise transporter* can be implemented in time $q^{O(n)}$ by *permutation group algorithms*. Furthermore, we can obtain a *generating set* of a matrix group G of size $q^{O(n)}$ in time $q^{O(n)}$ by *Sims' method* [Sim78]. These algorithms are classical and can be found in [Luk90, Ser03]. Also we emphasize that, if we are required to compute a coset of a subgroup $H \leq G$, the coset will be specified by a list of generators of H and a coset representative $g \in G$.

As mentioned in Section 5.1, for GraphIso, the dynamic programming technique improves the (worst-case) brute-force $n! \cdot \mathrm{poly}(n)$ time bound to the $2^{O(n)}$ time bound, which can be understood as replacing the number of permutations $n!$ with the number of subsets 2^n. In the AltMatSpIso setting, the dynamic programming technique is more transparent when working with the *subset transporter problem*: Given a permutation group $P \leq S_n$ and two subsets $S_1, S_2 \subseteq [n]$ of size k, this technique gives a $2^k \cdot \mathrm{poly}(n)$-time algorithm to compute $P(S_1 \to S_2) := \{\sigma \in P : \sigma(S_1) = S_2\}$ (see also [BQ12] for a brief introduction). To illustrate the idea in the matrix group setting, we start with the *subspace transporter problem.*

Problem 5.3 (Subspace transporter problem)**.** *Let $G \leq \mathrm{GL}(n, q)$ be given by a set of generators, and let V, W be two subspaces of $\mathbb{F}_q^n$ of dimension d. The subspace transporter problem asks to compute the coset $G(V \to W) = \{g \in G : g(V) = W\}$.*

The subspace transporter problem admits the following brute-force algorithm. Fix a basis $(v_1, \ldots, v_d)$ of V, and enumerate all ordered basis of W at the multiplicative cost of q^{d^2}. For each ordered basis $(w_1, \ldots, w_d)$ of W, compute the coset $\{g \in G : \forall i \in [d], g(v_i) = w_i\}$ by adaptively using the pointwise transporter algorithm. This gives an algorithm running in time $q^{d^2+O(n)}$. Analogous to the permutation group setting, we aim to replace $O(q^{d^2})$, the number of ordered basis of $\mathbb{F}_q^d$, with $q^{\frac{1}{4}d^2+O(k)}$, the number of subspaces in $\mathbb{F}_q^d$, via the dynamic programming technique. For this, we first show how to enumerate all these subspaces.

Remark 5.26. The above algorithm can be also use to enumerate the autometry group of $\mathcal{G}$ by replacing $\mathcal{H}$ by $\mathcal{G}$ when $\mathcal{G}$ satisfy property $F(n, m, q, r)$. And corollary 5.2 follows immediately.

Lemma 5.27. *There exists a deterministic algorithm that enumerates all subspaces of $\mathbb{F}_q^n$, and for each subspace computes an ordered basis, in time $q^{\frac{1}{4}n^2+O(n)}$.*

Proof. Algorithm 3 enumerates all subspaces of $\mathbb{F}_q^n$, where a dimension-k subspace V is identified by a k-tuple of linear independent vectors $(v_1, \ldots, v_k)$ in $\mathbb{F}_q^n$ which span V. To analyze the complexity of Algorithm 3, note that the multiplicative cost which dominate

Algorithm 3 Enumerates all subspaces in $\mathbb{F}_q^n$

Input: $\mathbb{F}_q^n$.
Output: Lists S_k of k-tuples of (linear independent) vectors which represents dimension-k subspaces of $\mathbb{F}_q$ for $k = 0, \ldots, n$.

```
1: S_k ← ∅ for k = [n] and S_0 ← {0}.
2: for all i ∈ [n] do
3:    for all (v_1, ..., v_{i-1}) ∈ S_{i-1} do
4:       for all u ∉ ⟨v_1, ..., v_{i-1}⟩ do
5:          if (v_1, ..., v_{i-1}, u) ∉ S_i then
6:             S_i ← (v_1, ..., v_{i-1}, u)
7:          else
8:             Go to the next u ∉ ⟨v_1, ..., v_{i-1}⟩
9:          end if
10:      end for
11:   end for
12: end for
```

the time complexity is to go over all subspaces of $\mathbb{F}_q^n$. The total number equals

$$\begin{bmatrix} n \\ 0 \end{bmatrix}_q + \begin{bmatrix} n \\ 1 \end{bmatrix}_q + \cdots + \begin{bmatrix} n \\ n \end{bmatrix}_q \leq (n+1) \begin{bmatrix} n \\ \lceil \frac{n}{2} \rceil \end{bmatrix}_q \leq (n+1) q^{\frac{1}{4}n^2} \in q^{\frac{1}{4}n^2 + O(\log_q n)}. \tag{5.29}$$

Enumerate all vectors which are not in a given dimension-k subspaces (line 4) add a multiplicative cost $q^n - q^k \in q^{O(n)}$. All the other steps can be achieved by linear algebra computation. Thus, the total complexity of Algorithm 3 is $q^{\frac{1}{4}n^2 + O(n)}$. □

Theorem 5.28. *There exists a deterministic algorithm that solves the subspace transporter problem in time* $q^{\frac{1}{4}d^2 + O(n)}$.

Proof. Algorithm 4 can be used to solve the subspace transporter problem. We first fix an ordered basis $(v_1, \ldots, v_d)$ of V. For $k \in [d]$, let $V_k = \langle v_1, \ldots, v_k \rangle$. For $g \in G(V \to W)$, g maps V_k to some dimension-k subspace U of W. As the brute-force algorithm enumerate all possible d-tuples of basis vectors of W, we are motivate to replace the enumeration of basis by enumerating all possible subspaces. That is, we maintain a dynamic programming table consists of cells indexed by all subspaces of W. For a dimension-k subspace U, specified by a tuple of basis vectors, the corresponding cell will store the coset $G(V_k \to U) = \{g \in G : g(V_k) = U\}$. A list of subspaces can be obtained by Algorithm 3 in time $q^{\frac{1}{4}d^2 + O(d)}$ (Lemma 5.27). $G(V \to W)$ can be obtained by read-out the cell indexed by $U = W$.

Algorithm 4 Subspace transporter

Input: A matrix group $G \leq \mathrm{GL}(n, q)$. Two dimension-d subspaces $V, W \leq \mathbb{F}_q^n$.
Output: $G(V \to W)$.

1: Specify an ordered basis of V by $(v_1, \ldots, v_d)$. Let $V_k = \langle v_1, \ldots, v_k \rangle$ for $k \in [d]$.
2: Compute a list of all subspaces of W by Algorithm 3.
3: $G(V_l \to U) \leftarrow \emptyset$ for all dimension-l subspaces $U \leq W$, where $l \in [d]$.
4: $G(V_0 \to \{0\}) \leftarrow G$.
5: **for all** $k \in [d]$ **do**
6: **for all** $U \leq W$, $\dim(U) = k$ **do**
7: **for all** $U_0 \leq U$, $\dim(U_0) = k - 1$ **do**
8: **for all** $u \in U \setminus U_0$ **do**
9: Compute $[G(V_{k-1} \to U_0)](v_k \to u)$
10: $G(V_k \to U) \leftarrow [G(V_{k-1} \to U_0)](v_k \to u)$
11: **end for**
12: **end for**
13: Obtain a generating set of $G(V_k \to U)$ of size $q^{O(n)}$
14: **end for**
15: **end for**
16: $G(V \to W) \leftarrow G(V_d \to W)$

We fill in the dynamic programming table adaptively in an increasing order according to k. For $k = 0$, we set $G(\{0\} \to \{0\}) = G$. Assume for $k \in [d]$, we have computed $G(V_l \to U')$ for all subspaces $U' \leq W$ of dimension l where $0 \leq l \leq k - 1$. To compute $G(V_k \to U)$ for a fixed $U \leq W$ of dimension k, note that any $g \in G(V_k \to U)$ has to map V_{k-1} to some dimension-$(k-1)$ subspace $U_0 \leq U$, and v_k to some vector $u \in U \setminus U_0$. This gives that

$$G(V_k \to U) = \bigcup_{\substack{U_0 \leq U, \\ \dim(U_0) = k-1}} \bigcup_{u \in U \setminus U_0} [G(V_{k-1} \to U_0)](v_k \to u). \tag{5.30}$$

For a fixed dimension-$(k-1)$ subspace $U_0 \leq U$ and $u \in U \setminus U_0$, we compute $[G(V_{k-1} \to U_0)](v_k \to u)$ (line 9) as follows. We can read $G(V_{k-1} \to U_0)$ from the table, then compute $[G(V_{k-1} \to U_0)](v_k \to u)$ using the pointwise transporter algorithm (e.g. see proposition (3.9) in Luks' report [Luk93]). The number of u in $U \setminus U_0$ is no more than q^k, and the number of dimension-$(k-1)$ subspaces of U is also no more than q^k. After taking these two unions, apply Sims' method [Sim78] to get a generating set of size $q^{O(n)}$ (line 13). Therefore we can compute each cell in time at most $q^{2d} \cdot q^{O(n)} = q^{O(n)}$. Therefore the whole dynamic programming table can be filled in time $q^{\frac{1}{4}d^2 + O(d)} \cdot q^{O(n)} = q^{\frac{1}{4}d^2 + O(n)}$. $\square$

Now, we intend to derive an algorithm for AltMatSpIso, using the dynamic programming technique. We settle the following problem first.

Problem 5.4 (Alternating matrix transporter problem)**.** *Let $H \leq \mathrm{GL}(n,q)$ be given by a set of generators, and let $A, B \in \Lambda(n,q)$ be two alternating matrices. The alternating matrix transporter problem asks to compute the coset $H(A \to B) = \{g \in H : g^{\mathrm{t}}Ag = B\}$.*

The alternating matrix transporter problem admits a brute-force algorithm which runs in time $q^{\Theta(n^2)+O(n)}$, by simply enumerate all elements in H and test whether they transform A to B. Note that the alternating matrix transporter problem can not be converted into an instance of pointwise transporter problem over domain $F_q^{\frac{1}{2}n(n-1)}$, as we cannot transform the action of $g \in G$ on $B \in \Lambda(n,q)$ into the permutation group setting.

Theorem 5.29. *There exists a deterministic algorithm that solves the alternating matrix transporter problem in time $q^{\frac{1}{4}n^2+O(n)}$.*

Proof. Algorithm 5 solves the alternating matrix transporter problem.

Algorithm 5 Alternating matrix transporter

Input: A matrix group $H \leq \mathrm{GL}(n,q)$. Two alternating matrix $A, B \in \Lambda(n,q)$.
Output: $H(A \to B)$.

1: Specify an ordered basis of $\mathbb{F}_q^n = \langle e_1, \ldots, e_n \rangle$ represent A. Let $E_d = \langle e_1, \ldots, e_d \rangle$ for $d \in [n]$.
2: $H(A|_{E_d} \to B|_U) \leftarrow \emptyset$ for all dimension-d subspaces $U \leq \mathbb{F}_q^n$ where $d \in [n]$.
3: $H(A|_{\{0\}} \to B|_{\{0\}}) \leftarrow H$.
4: **for all** $d \in [n]$ **do**
5: **for all** dimension-d $U \leq \mathbb{F}_q^n$ **do**
6: **for all** dimension-$(d-1)$ $U_0 \leq U$ **do**
7: **for all** $u \in U \setminus U_0$ **do**
8: Compute $\big[[H(A|_{E_{d-1}} \to B|_{U_0})](e_d \to u)\big](A|_{E_{d-1}\times e_d} \to B|_{U_0 \times u})$
9: $H(A|_{E_d} \to B|_U) \leftarrow \big[[H(A|_{E_{d-1}} \to B|_{U_0})](e_d \to u)\big](A|_{E_{d-1}\times e_d} \to B|_{U_0 \times u})$
10: **end for**
11: **end for**
12: Obtain a generating set of $H(A|_{E_d} \to B|_U)$ of size $q^{O(n)}$
13: **end for**
14: **end for**
15: $H(A \to B) \leftarrow H(A|_{E_n} \to B|_{\mathbb{F}_q^n})$

Let $(e_1, \ldots, e_n)$ be the standard basis vectors of $\mathbb{F}_q^n$ which represents A. Let $E_d = \langle e_1, \ldots, e_d \rangle$ for $d \in [n]$. For an alternating matrix B, and an ordered basis $(u_1, \ldots, u_d)$ of

a dimension-d $U \leq \mathbb{F}_q^n$, we define

$$B|_U := \begin{bmatrix} u_1 & \cdots & u_d \end{bmatrix}^{\mathrm{t}} B \begin{bmatrix} u_1 & \cdots & u_d \end{bmatrix} \in \Lambda(d, q). \tag{5.31}$$

$B|_U$ stands for the restriction of B to the subspace U. For a vector v and U with the ordered basis as above, we define

$$B_{U \times v} = \begin{bmatrix} u_1 & \cdots & u_d \end{bmatrix}^{\mathrm{t}} Bv \in \mathbb{F}_q^d. \tag{5.32}$$

Algorithm 5 output a dynamic programming table, which is a list indexed by all subspaces of $\mathbb{F}_q^n$. These subspaces will be specified by an arbitrary tuple of ordered basis, which can be obtained in time $q^{\frac{1}{4}n^2+O(n)}$ utilizing Algorithm 3. For any $U = \langle u_1, \ldots, u_d \rangle \leq \mathbb{F}_q^n$ (of dimension d), its corresponding cell stores the coset

$$H(A|_{E_d} \to B|_U) = \{g \in H : \ g(E_d) = U, \ g^{\mathrm{t}}(A|_{E_d})g = B|_U\}. \tag{5.33}$$

We will fill in the list adaptively in the increasing order of the dimension d. The base case $d = 0$ is trivial. Assume we have already compute the coset $H(A|_{E_l} \to B|_{U'})$ for all dimension-l subspace $U' \leq \mathbb{F}_q^n$ and $0 \leq l \leq d-1$. To compute $H(A|_{E_d} \to B|_U)$ for some $U \leq \mathbb{F}_q^n$ of dimension d, we point out that any $g \in H(A|_{E_d} \to B|_U)$ satisfies the following: On the one hand, g must map E_{d-1} to some dimension-$(d-1)$ subspace $U_0 \leq U$, and $A|_{E_{d-1}} \in \Lambda(d-1, q)$ to $B|_{U_0} \in \Lambda(d-1, q)$. On the other hand, g sends e_d to some $u \in U \setminus U'$, and $A|_{E_{d-1} \times e_d} \in \mathbb{F}_q^{d-1}$ to $B|_{U' \times u} \in \mathbb{F}_q^{d-1}$. We construct $H(A|_{E_d} \to B|_U)$ by

$$H(A|_{E_d} \to B|_U) = \bigcup_{\substack{U_0 \leq U, \\ \dim(U_0) = d-1}} \bigcup_{u \in U \setminus U_0} \big[[H(A|_{E_{d-1}} \to B|_{U_0})](e_d \to u)\big](A|_{E_{d-1} \times e_d} \to B|_{U_0 \times u}). \tag{5.34}$$

For fixed dimension-$(d-1)$ subspace U_0 and $u \in U \setminus U_0$, we compute $\big[[H(A|_{E_{d-1}} \to B|_{U_0})](e_d \to u)\big](A|_{E_{d-1} \times e_d} \to B|_{U_0 \times u})$ (line 8) as follows. We read $H(A|_{E_{d-1}} \to B|_{U_0})$ from the table, compute $[H(A|_{E_{d-1}} \to B|_{U_0})](e_d \to u)$ using the pointwise transporter algorithm. As $[H(A|_{E_{d-1}} \to B|_{U_0})](e_d \to u)$ induces an action on $\mathbb{F}_q^{d-1}$ corresponding to the last column of $A|_{E_d}$ with the last entry (which is 0) being removed, $\big[[H(A|_{E_{d-1}} \to B|_{U_0})](e_d \to u)\big](A|_{E_{d-1} \times e_d} \to B|_{U_0 \times u})$ can be computed by another pointwise transporter

algorithm. We go over the two unions and apply Sims' method to obtain a generating set of size $q^{O(n)}$. The multiplicative cost for filling in each cell is at most $q^{2d} \cdot q^{O(n)} \in q^{O(n)}$, and the total time complexity is then $q^{\frac{1}{4}n^2+O(n)}$. □

We are now ready to prove Theorem 5.3.

Theorem 5.3, restated. Given $\mathbb{G} = (G_1, \ldots, G_m)$ and $\mathbb{H} = (H_1, \ldots, H_m)$ in $\Lambda(n,q)^m$ representing two m-alternating spaces $\mathcal{G}, \mathcal{H} \leq \Lambda(n,q)$, there exists a deterministic algorithm for AltMatSpIso in time $q^{\frac{1}{4}(m^2+n^2)+O(m+n)}$.

Proof. We claim that Algorithm 6 can be used to AltMatSpIso.

Algorithm 6 AltMatSpIso

Input: Two m-alternating tuples $\mathbb{G} = (G_1, \ldots, G_m)$ and $\mathbb{H} = (H_1, \ldots, H_m)$ in $\Lambda(n,q)^m$ representing m-alternating spaces $\mathcal{G}, \mathcal{H} \leq \Lambda(n,q)$.
Output: $\mathrm{Iso}(\mathcal{G}, \mathcal{H})$.

Specify an ordered basis of $\mathbb{F}_q^m = \langle e_1, \ldots, e_m \rangle$. Let $E_k = \langle e_1, \ldots, e_k \rangle$ for $k = [m]$.
$\mathrm{Iso}\mathcal{G}, \mathcal{H} \leftarrow \emptyset$. $\mathrm{Iso}(\mathbb{G}^{E_k}, \mathbb{H}^V) \leftarrow \emptyset$ for all dimension-k subspaces $V \leq \mathbb{F}_q^m$ and $k \in [m]$.
$\mathrm{Iso}(\mathbb{G}^{\{0\}}, \mathbb{H}^{\{0\}}) \leftarrow \mathrm{GL}(n,q) \times \mathrm{GL}(m,q)$.
for all $k \in [m]$ **do**
  for all dimension-k $V \leq \mathbb{F}_q^m$ **do**
    for all dimension-$(k-1)$ $V_0 \leq V$ **do**
      for all $v \in V \setminus V_0$ **do**
        Compute $\big[[\mathrm{Iso}(\mathbb{G}^{E_{k-1}}, \mathbb{H}^{V_0})](e_k \to v)\big](\mathbb{G}^{e_k} \to \mathbb{H}^v)$
        $\mathrm{Iso}(\mathbb{G}^{E_k}, \mathbb{H}^V) \to B|_U) \leftarrow \big[[\mathrm{Iso}(\mathbb{G}^{E_{k-1}}, \mathbb{H}^{V_0})](e_k \to v)\big](\mathbb{G}^{e_k} \to \mathbb{H}^v)$
      end for
    end for
    Obtain a generating set of $\mathrm{Iso}(\mathbb{G}^{E_k}, \mathbb{H}^V)$ of size $q^{O(n)}$
  end for
end for
$\mathrm{Iso}(\mathcal{G}, \mathcal{H}) \leftarrow \pi_1(\mathrm{Iso}(\mathbb{G}^{E_m}, \mathbb{H}^{\mathbb{F}_q^m}))$

Let $(e_1, \ldots, e_m)$ be the standard basis of $\mathbb{F}_q^m$ and let $E_k = \langle e_1, \ldots, e_k \rangle$ for $k \in [m]$. Let $\mathbb{G}^{E_k} = (G_1, \ldots, G_k) \in \Lambda(n,q)^k$. For a vector $v = \begin{bmatrix} a_1, \cdots, a_m \end{bmatrix}^{\mathrm{t}} \in \mathbb{F}_q^m$, define $\mathbb{H}^v := \sum_{i \in [m]} a_i H_i \in \Lambda(n,q)$. For a dimension-$k$ subspace $V \leq \mathbb{F}_q^m$ with an ordered basis $(v_1, \ldots, v_k)$, $\mathbb{H}^V := (\mathbb{H}^{v_1}, \ldots, \mathbb{H}^{v_k}) \in \Lambda(n,q)^k$.

The algorithm outputs a dynamic programming table, which is a list indexed by subspaces of $\mathbb{F}_q^m$. By Lemma 5.27, we can obtain these subspaces specified by an ordered basis in

time $q^{\frac{1}{4}m^2+O(m)}$ using Algorithm 3. The cell corresponding to a dimension-k subspace $V \leq \mathbb{F}_q^m$ stores the coset

$$\mathrm{Iso}(\mathbb{G}^{E_k}, \mathbb{H}^V) = \{(g,h) \in \mathrm{GL}(n,q) \times \mathrm{GL}(k,q) : g^{\mathrm{t}}(\mathbb{G}^{E_k})g = (\mathbb{H}^V)^h\}, \tag{5.35}$$

where $\mathbb{H}^h = (\sum_{j\in[m]} h_{1,j}H_j, \sum_{j\in[m]} h_{2,j}H_j, \ldots, \sum_{j\in[m]} h_{m,j}H_j)$, where $[h_{i,j}]_{i,j\in[m]}$ forms the invertible matrix representing $h \in \mathrm{GL}(n,q)$.

We will fill in the dynamic programming table adaptively in the increasing order of the dimension k. The base case $k = 0$ is trivial. Now assume we have computed $\mathrm{Iso}(\mathbb{G}^{E_l}, \mathbb{H}^V)$ for all $V \leq \mathbb{F}_q^n$ of dimension l and $0 \leq l \leq k-1$. To compute $\mathrm{Iso}(\mathbb{G}^{E_k}, \mathbb{H}^V)$ for a given dimension-k subspace $V \leq \mathbb{F}_q^n$, note that any $(g,h) \in \mathrm{Iso}(\mathbb{G}^{E_k}, \mathbb{H}^V)$ satisfies the following. On the one hand, h sends E_{d-1} to some dimension-$(k-1)$ subspace $V_0 \leq V$, and $(g,h) \in \mathrm{Iso}(\mathbb{G}^{E_{k-1}}, \mathbb{H}^{V_0})$. On the other hand, h sends e_k to some $v \in V \setminus V_0$, and g sends $\mathbb{G}^{e_k}$ to $\mathbb{H}^v$. We construct $\mathrm{Iso}(\mathbb{G}^{E_d}, \mathbb{H}^V)$ by

$$\mathrm{Iso}(\mathbb{G}^{E_k}, \mathbb{H}^V) = \bigcup_{\substack{V_0 \leq V, \\ \dim(V_0)=k-1}} \bigcup_{v \in V\setminus V_0} \big[[\mathrm{Iso}(\mathbb{G}^{E_{k-1}}, \mathbb{H}^{V_0})](e_k \to v)\big](\mathbb{G}^{e_k} \to \mathbb{H}^v). \tag{5.36}$$

Now, for fixed V_0 and v, we compute $\big[[\mathrm{Iso}(\mathbb{G}^{E_{k-1}}, \mathbb{H}^{V_0})](e_k \to v)\big](\mathbb{G}^{e_k} \to \mathbb{H}^v)$ (line 8) as follows. $\mathrm{Iso}(\mathbb{G}^{E_{k-1}}, \mathbb{H}^{V_0})$ can be read from the table. $[\mathrm{Iso}(\mathbb{G}^{E_{k-1}}, \mathbb{H}^{V_0})](e_k \to v)$ is an instance of the pointwise transporter problem of $\mathrm{GL}(n,q) \times \mathrm{GL}(k,q)$ acting on $\mathbb{F}_q^m$, which can be solved in time $q^{O(m)}$. $\big[[\mathrm{Iso}(\mathbb{G}^{E_{k-1}}, \mathbb{H}^{V_0})](e_k \to v)\big](\mathbb{G}^{e_k} \to \mathbb{H}^v)$ is an instance of the alternating matrix transporter problem, which can be solved, by Algorithm 5, in time $q^{\frac{1}{4}n^2+O(n)}$. Going over the two unions adds a multiplicative factor of q^{2k} for $k \in [m]$, and then we apply Sims' method to reduce the size of the generating set to $q^{O(n)}$ (line 12). Therefore each cell can be computed in $q^{2k} \cdot q^{\frac{1}{4}n^2+O(n+m)} \in q^{\frac{1}{4}n^2+O(m+n)}$. The whole dynamic programming table can be filled in time $q^{\frac{1}{4}m^2+O(m)} \cdot q^{\frac{1}{4}n^2+O(n+m)} \in q^{\frac{1}{4}(n^2+m^2)+O(n+m)}$. To obtain $\mathrm{Iso}(\mathcal{G}, \mathcal{H})$, we simply perform the projection π_1 of $\mathrm{GL}(n,q) \times \mathrm{GL}(m,q)$ to $\mathrm{GL}(n,q)$ along $\mathrm{GL}(m,q)$ (line 15). □

5.6 Summary and Discussion

In this chapter, we have exhibited an average-case efficient algorithm (Algorithm 2) which tests isometry for most alternating matrix spaces in the linear algebraic Erdős-Rényi model $LinER(n, m, q)$ with any other m-alternating spaces in $\Lambda(n, q)$ in time $q^{O(n)}$. We have also devised a deterministic algorithm (Algorithm 6) for AltMatSpIso in $q^{\frac{1}{4}}(n^2 + m^2) + O(n + m)$. Our average-case algorithm is inspired by the seminal work of Babai, Erdős and Selkow [BES80], which introduced the first average-case efficient algorithm for testing isomorphism between most graphs in the Erdős-Rényi model and any other graphs in linear time. To derive Algorithm 2, we have viewed and studied AltMatSpIso as a linear algebraic analog of GraphIso, and developed a linear algebraic analog of the individualization and refinement technique. To derive Algorithm 5, we have adapted Luks' dynamic programming technique for GraphIso to devise a $q^{\frac{1}{4}(n^2+m^2)+O(n+m)}$-time algorithm for AltMatSpIso, which is slightly better than the brute-force algorithm (when $m \in O(n)$) and Rosenbaum's algorithm [Ros13]. Algorithm 6 can be also served as a piece of evidence to support the "*linear algebraic*" viewpoint.

We would like to mention a bit more on GraphIso and AltMatSpIso. In the history of GraphIso, Two (families of) algorithmic ideas have been most responsible for the worst-case time complexity improvements. The first idea, which refers to the combinatorial idea, is to use certain combinatorial techniques including individualization, vertex or edge refinement, and more generally the Weisfeiler-Lehman refinement [WL68]. The second idea, which we call the group-theoretic idea, is to reduce GraphIso to certain problems in permutation group algorithms, and then settle those problems using group-theoretic techniques and structures. A major breakthrough utilizing the group-theoretic idea is the polynomial-time algorithm for graphs with bounded degree by Luks [Luk82]. Some combinatorial techniques have been implemented and used in practice [MP14], though the worst-case analysis usually does not favor such algorithms (see e.g. [CFI92]). On the other hand, while group-theoretic algorithms for GraphIso more than often come with a rigorous analysis, such algorithms usually only work with a restricted family of graphs (see e.g. [Luk82]). The major improvements on the worst-case time complexity of

GraphIso almost always rely on both ideas. The recent breakthrough, a quasipolynomial-time algorithm for GraphIso by Babai [Bab16a, Bab16b], is a clear evidence. Even the previous record, a $2^{\tilde{O}(\sqrt{n})}$-time algorithm by Babai and Luks [BL83], relies on both Luks' group-theoretic framework [Luk82] and Zemlyachenko's combinatorial partitioning lemma [ZKT85].

Let us return to AltMatSpIso. It is clear that AltMatSpIso can be studied in the context of matrix groups over finite fields. Despite that matrix groups algorithms are normally harder than permutation group algorithms. If a $q^{O(n+m)}$-time algorithm for AltMatSpIso is the main concern, then we can view $\mathrm{GL}(n, q)$ acting on the domain $\mathbb{F}_q^n$ of size q^n, so group-theoretic tasks are not a bottleneck. In addition, a group-theoretic framework for matrix groups in the vein of the corresponding permutation group results in [Luk82] has also been developed by Luks [Luk92]. Therefore, if we aim at a $q^{O(n+m)}$-time algorithm for AltMatSpIso, the group-theoretic aspect is relatively developed. On the other hand, the other major idea, namely the combinatorial refinement idea, seemed missing in the context of AltMatSpIso. We hope that, the ingredients presented in this chapter may open the door to systematically examine and adapt such combinatorial refinement techniques for GraphIso to improve the worst-case time complexity of AltMatSpIso.

Chapter 6

Conclusion

In this thesis, we apply the theory of matrix spaces to investigate fundamental problems in quantum information and computational complexity. We have contributed in the areas of PPT-distinguishability of orthogonal bipartite states, the parallel distinguishability of quantum channels, the tripartite-to-bipartite SLOCC entanglement transformation, and the alternating matrix isometry. We conclude that the theory of matrix spaces can be utilized in the following three aspects.

The first application of matrix spaces is to formulate problems in quantum information with respect to matrix spaces. We have shown how to convert the PPT-distinguishability of orthogonal bipartite states and the parallel distinguishability of quantum channels into the formalism of extendibility problems with respect to matrix spaces. Essentially, they are equivalent to deciding whether a given matrix space is strongly PPT-unextendible or strongly positive-unextendible. Note that similar problems have been widely studied with respect to vector spaces. Extendibility problems with respect to matrix spaces can be viewed as a natural generalizations, and admit a much more complicated structure. Another significant formulation appears in the study of tripartite-to-bipartite SLOCC entanglement transformation, where determining the (one-shot) SLOCC convertibility is equivalent to computing the maximal rank of matrix spaces. Such a formulation has led to the equivalence between asymptotic SLOCC convertibility of tripartite pure states to the bipartite maximally entangled state and the non-commutative symbolic determinant identity testing.

The second application of matrix spaces crucially relies on the first one. By formulating problems with respect to matrix spaces, We have access to powerful mathematical results in the theory of matrix spaces. For instance, we have proposed semidefinite programs to decide whether a given matrix space is PPT-unextendible or positive-unextendible. Specifically, we have capitalized on the theory of numerical range and the Farkas lemma of semidefinite programming to characterize the strong positive-unextendibility of two families of matrix spaces. We have resorted the shrunk subspaces concept, as well as matrix semi-invariants, to derive explicit formulas which compute the asymptotic SLOCC transformation rate for two families of tripartite states. These formulas are sufficient to derive a complete characterization of the asymptotic SLOCC convertibility to the bipartite maximally entangled state. In addition, to devise the average-case efficient algorithm for the alternating matrix space isometry problem, we have analyzed the property of the adjoint algebra of matrix spaces for the sake of average-case analysis, and the proof idea is inspired by the stable concept in geometric invariant theory.

The last application of matrix spaces, which might be the most interesting and important, is to provide a promising method to transform combinatorial ideas into the study of algebraic structures. More precisely, the "linear algebraic analog" viewpoint postulated that matrix spaces can be viewed and studied as a linear algebraic analog of bipartite graphs, and alternating matrix spaces can be viewed and studied as a linear algebraic analog of graphs. This viewpoint enables us to interpret concepts related to matrix spaces as linear algebraic generalizations from concepts in graphs. For instance, the shrunk subspace of matrix spaces can be viewed as a linear algebraic analog of the shrunk subset, originating from the well-known Hall's marriage theorem. On the other hand, such a viewpoint inspires us to adapt combinatorial techniques for graph-theoretic problems to study their linear algebraic counterparts. Important evidences include the linear algebraic analog of individualization and refinement technique, which can be utilized to devise an average-case efficient algorithm for testing isometry between alternating matrix space. Although the "linear algebraic" world can be completely different from the classical world, the "linear algebraic analog" viewpoint provides new insights and heuristics to tackle hard problems with respect to matrix spaces.

As the end of this thesis, we propose some open problems, which may deserve further investigations. The first one is to complete characterize the strongly PPT-unextendible and strongly positive-unextendible matrix spaces. These characterizations will lead to complete solutions for deciding the PPT-distinguishability of orthogonal bipartite states and the parallel distinguishability of quantum channels. The second open question is to compute the asymptotic maximal rank for shrinking matrix spaces. Resolving such a problem not only provides a complete characterization for the tripartite-to-bipartite SLOCC entanglement transformation in the asymptotic setting, but also brings new insight on derandomizing SDIT. The last but not least problem is to devise algorithms for the alternating matrix space isometry problem in time polynomial in the underlying vector space size. Such algorithms would remove perhaps the largest roadblock on the way to solving group isomorphism problem in time polynomial in the group order. Although there seems a long way to go, we believe that our viewpoint may lead to some non-trivial improvement over the brute-force algorithm.

Bibliography

[Ací01] A. Acín. Statistical distinguishability between unitary operations. *Phys. Rev. Lett.*, 87:177901, Oct 2001.

[ACMnT+07] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz Tapia, E. Bagan, Ll. Masanes, A. Acin, and F. Verstraete. Discriminating states: The quantum chernoff bound. *Phys. Rev. Lett.*, 98:160501, Apr 2007.

[ANS07] B. Adsul, S. Nayak, and K. V. Subrahmanyam. A geometric approach to the Kronecker problem II: rectangular shapes, invariants of matrices and the Artin–Procesi theorem. preprint, 2007.

[Ash90] R.B. Ash. *Information Theory*. Dover books on advanced mathematics. Dover Publications, 1990.

[Bab16a] László Babai. Graph isomorphism in quasipolynomial time. *arXiv preprint arXiv:1512.03547*, 2016.

[Bab16b] László Babai. Graph isomorphism in quasipolynomial time [extended abstract]. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, STOC '16, pages 684–697, New York, NY, USA, 2016. ACM.

[Bae38] Reinhold Baer. Groups with abelian central quotient group. *Transactions of the American Mathematical Society*, 44(3):357–386, 1938.

[Ban11] Somshubhro Bandyopadhyay. More nonlocality with less purity. *Phys. Rev. Lett.*, 106:210402, May 2011.

[BD06] Matthias Bürgin and Jan Draisma. The Hilbert null-cone on tuples of matrices and bilinear forms. *Mathematische Zeitschrift*, 254(4):785–809, Dec 2006.

[BDF+99] Charles H. Bennett, David P. DiVincenzo, Christopher A. Fuchs, Tal Mor, Eric Rains, Peter W. Shor, John A. Smolin, and William K. Wootters. Quantum nonlocality without entanglement. *Phys. Rev. A*, 59:1070–1091, Feb 1999.

[BDM+99] Charles H. Bennett, David P. DiVincenzo, Tal Mor, Peter W. Shor, John A. Smolin, and Barbara M. Terhal. Unextendible product bases and bound entanglement. *Phys. Rev. Lett.*, 82:5385–5388, Jun 1999.

[BES80] László Babai, Paul Erdős, and Stanley M. Selkow. Random graph isomorphism. *SIAM Journal on Computing*, 9(3):628–635, 1980.

[BK79] László Babai and Ludek Kučera. Canonical labelling of graphs in linear average time. In *20th Annual Symposium on Foundations of Computer Science (SFCS 1979)*, pages 39–46, Oct 1979.

[BL83] László Babai and Eugene M. Luks. Canonical labeling of graphs. In *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing*, STOC '83, pages 171–183, New York, NY, USA, 1983. ACM.

[BL08] Peter A. Brooksbank and Eugene M. Luks. Testing isomorphism of modules. *Journal of Algebra*, 320(11):4020 – 4029, 2008. Computational Algebra.

[BMW17] Peter A. Brooksbank, Joshua Maglione, and James B. Wilson. A fast isomorphism test for groups whose lie algebra has genus 2. *Journal of Algebra*, 473:545 – 590, 2017.

[BNS98] Howard Barnum, Michael A. Nielsen, and Benjamin Schumacher. Information transmission through a noisy quantum channel. *Phys. Rev. A*, 57:4153–4175, Jun 1998.

[BO08] Peter A. Brooksbank and E. A. O'Brien. Constructing the group preserving a system of forms. *International Journal of Algebra and Computation*, 18(02):227–241, 2008.

[Bol01] Béla Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2001.

[BQ12] László Babai and Youming Qiao. Polynomial-time isomorphism test for groups with abelian Sylow towers. In *29th International Symposium on Theoretical Aspects of Computer Science (STACS 2012)*, volume 14 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 453–464, Dagstuhl, Germany, 2012.

[BRS02] Somshubhro Bandyopadhyay, Vwani Roychowdhury, and Ujjwal Sen. Classification of nonasymptotic bipartite pure-state entanglement transformations. *Phys. Rev. A*, 65:052315, May 2002.

[BW09] Somshubhro Bandyopadhyay and Jonathan Walgate. Local distinguishability of any three quantum states. *Journal of Physics A: Mathematical and Theoretical*, 42(7):072002, 2009.

[BW12] Peter A. Brooksbank and James B. Wilson. Computing isometry groups of hermitian maps. *Transactions of the American Mathematical Society*, 364(4):1975–1996, 2012.

[CCD+10] Lin Chen, Eric Chitambar, Runyao Duan, Zhengfeng Ji, and Andreas Winter. Tensor rank and stochastic entanglement catalysis for multipartite pure states. *Phys. Rev. Lett.*, 105:200501, Nov 2010.

[CCH11] Toby S. Cubitt, Jianxin Chen, and Aram W. Harrow. Superactivation of the asymptotic zero-error classical capacity of a quantum channel. *IEEE Transactions on Information Theory*, 57(12):8114–8126, Dec 2011.

[CDS08] Eric Chitambar, Runyao Duan, and Yaoyun Shi. Tripartite entanglement transformations and tensor rank. *Phys. Rev. Lett.*, 101:140502, Oct 2008.

[CDS10] Eric Chitambar, Runyao Duan, and Yaoyun Shi. Multipartite-to-bipartite entanglement transformations and polynomial identity testing. *Phys. Rev. A*, 81:052310, May 2010.

[CFI92] Jin-Yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12(4):389–410, Dec 1992.

[Cho75] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285 – 290, 1975.

[CIK97] Alexander Chistov, Gábor Ivanyos, and Marek Karpinski. Polynomial time algorithms for modules over finite dimensional algebras. In *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation*, ISSAC '97, pages 68–74, New York, NY, USA, 1997. ACM.

[CKT+07] Anthony Chefles, Akira Kitagawa, Masahiro Takeoka, Masahide Sasaki, and Jason Twamley. Unambiguous discrimination among oracle operators. *Journal of Physics A: Mathematical and Theoretical*, 40(33):10183, 2007.

[CMW08] Toby Cubitt, Ashley Montanaro, and Andreas Winter. On the dimension of subspaces with bounded schmidt rank. *Journal of Mathematical Physics*, 49(2):022107, 2008.

[Coh75] Paul M. Cohn. The word problem for free fields: A correction and an addendum. *The Journal of Symbolic Logic*, 40(1):69–74, 1975.

[CY10] Jianxin Chen and Mingsheng Ying. Ancilla-assisted discrimination of quantum gates. *Quantum Info. Comput.*, 10(1):160–177, January 2010.

[DFJY05] Runyao Duan, Yuan Feng, Zhengfeng Ji, and Mingsheng Ying. Efficiency of deterministic entanglement transformation. *Phys. Rev. A*, 71:022305, Feb 2005.

[DFJY07] Runyao Duan, Yuan Feng, Zhengfeng Ji, and Mingsheng Ying. Distinguishing arbitrary multipartite basis unambiguously using local operations and classical communication. *Phys. Rev. Lett.*, 98:230502, Jun 2007.

[DFLY05a] Runyao Duan, Yuan Feng, Xin Li, and Mingsheng Ying. Multiple-copy entanglement transformation and entanglement catalysis. *Phys. Rev. A*, 71:042319, Apr 2005.

[DFLY05b] Runyao Duan, Yuan Feng, Xin Li, and Mingsheng Ying. Trade-off between multiple-copy transformation and entanglement catalysis. *Phys. Rev. A*, 71:062306, Jun 2005.

[DFM+99] David P. DiVincenzo, Christopher A. Fuchs, Hideo Mabuchi, John A. Smolin, Ashish Thapliyal, and Armin Uhlmann. *Entanglement of Assistance*, pages 247–257. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.

[DFY05] Runyao Duan, Yuan Feng, and Mingsheng Ying. Entanglement-assisted transformation is asymptotically equivalent to multiple-copy transformation. *Phys. Rev. A*, 72:024306, Aug 2005.

[DFY07] Runyao Duan, Yuan Feng, and Mingsheng Ying. Entanglement is not necessary for perfect discrimination between unitary operations. *Phys. Rev. Lett.*, 98:100503, Mar 2007.

[DFY09] Runyao Duan, Yuan Feng, and Mingsheng Ying. Perfect distinguishability of quantum operations. *Phys. Rev. Lett.*, 103:210501, Nov 2009.

[DGLL16] Runyao Duan, Cheng Guo, Chi-Kwong Li, and Yinan Li. Parallel distinguishability of quantum operations. In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 2259–2263, July 2016.

[DLPP01] G. Mauro D'Ariano, Paoloplacido Lo Presti, and Matteo G. A. Paris. Using entanglement improves the precision of quantum measurements. *Phys. Rev. Lett.*, 87:270404, Dec 2001.

[DM17] Harm Derksen and Visu Makam. Polynomial degree bounds for matrix semi-invariants. *Advances in Mathematics*, 310:44 – 63, 2017.

[DMS+03] David P. DiVincenzo, Tal Mor, Peter W. Shor, John A. Smolin, and Barbara M. Terhal. Unextendible product bases, uncompletable product

bases and bound entanglement. *Communications in Mathematical Physics*, 238(3):379–410, Jul 2003.

[DRA06] JAN DRAISMA. Small maximal spaces of non-invertible matrices. *Bulletin of the London Mathematical Society*, 38(5):764776, 2006.

[DSK05] G. Mauro D'Ariano, Massimiliano F. Sacchi, and Jonas Kahn. Minimax discrimination of two pauli channels. *Phys. Rev. A*, 72:052302, Nov 2005.

[DSW13] Runyao Duan, Simone Severini, and Andreas Winter. Zero-error communication via quantum channels, noncommutative graphs, and a quantum Lovász number. *IEEE Transactions on Information Theory*, 59(2):1164–1174, Feb 2013.

[Dua09] Runyao Duan. Super-activation of zero-error capacity of noisy quantum channels. *arXiv preprint arXiv:0906.2527*, 2009.

[DVC00] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A*, 62:062314, Nov 2000.

[DW00] Harm Derksen and Jerzy Weyman. Semi-invariants of quivers and saturation for Littlewood-Richardson coefficients. *Journal of the American Mathematical Society*, 13(3):467–479, 2000.

[DZ01] Mátyás Domokos and Alexandr N. Zubkov. Semi-invariants of quivers as determinants. *Transformation Groups*, 6(1):9–24, Mar 2001.

[Edm67] Jack Edmonds. Systems of distinct representatives and linear algebra. *J. Res. Nat. Bur. Standards Sect. B*, 71:241–245, 1967.

[EH88] David Eisenbud and Joe Harris. Vector spaces of matrices of low rank. *Advances in Mathematics*, 70(2):135 – 155, 1988.

[ELGO02] Bettina Eick, C. R. Leedham-Green, and E. A. O'Brien. Constructing automorphism groups of p-groups. *Communications in Algebra*, 30(5):2271–2295, 2002.

[ER59] P Erdős and A Rényi. On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

[ER63] P. Erdős and A. Rényi. Asymmetric graphs. *Acta Mathematica Academiae Scientiarum Hungarica*, 14(3):295–315, Sep 1963.

[Fan04] Heng Fan. Distinguishability and indistinguishability by local operations and classical communication. *Phys. Rev. Lett.*, 92:177905, Apr 2004.

[FDY05] Yuan Feng, Runyao Duan, and Mingsheng Ying. Catalyst-assisted probabilistic entanglement transformation. *IEEE Transactions on Information Theory*, 51(3):1090–1101, March 2005.

[FDY06] Yuan Feng, Runyao Duan, and Mingsheng Ying. Relation between catalyst-assisted transformation and multiple-copy transformation for bipartite pure states. *Phys. Rev. A*, 74:042312, Oct 2006.

[Fek23] Michael Fekete. Über die verteilung der wurzeln bei gewissen algebraischen gleichungen mit ganzzahligen koeffizienten. *Mathematische Zeitschrift*, 17(1):228–249, Dec 1923.

[FL07] Ben Fortescue and Hoi-Kwong Lo. Random bipartite entanglement from W and W-like states. *Phys. Rev. Lett.*, 98:260501, Jun 2007.

[FN70] V. Felsch and J. Neubüser. On a programme for the determination of the automorphism group of a finite group. In *Computational Problems in Abstract Algebra*, pages 59 – 60. Pergamon, 1970.

[FR04] Marc Fortin and Christophe Reutenauer. Commutative/noncommutative rank of linear matrices and subspaces of matrices of low rank. *Séminaire Lotharingien de Combinatoire*, 52:B52f, 2004.

[FS09] Yuan Feng and Yaoyun Shi. Characterizing locally indistinguishable orthogonal product states. *IEEE Transactions on Information Theory*, 55(6):2799–2806, June 2009.

[GB14] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, March 2014.

[GGOW16] Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. A deterministic polynomial time algorithm for non-commutative rational identity testing. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 109–117, Oct 2016.

[GKR+01] Sibasish Ghosh, Guruprasad Kar, Anirban Roy, Aditi Sen(De), and Ujjwal Sen. Distinguishability of bell states. *Phys. Rev. Lett.*, 87:277902, Dec 2001.

[GKRS04] Sibasish Ghosh, Guruprasad Kar, Anirban Roy, and Debasis Sarkar. Distinguishability of maximally entangled states. *Phys. Rev. A*, 70:022304, Aug 2004.

[GMS05] Gilad Gour, David A. Meyer, and Barry C. Sanders. Deterministic entanglement of assistance and monogamy constraints. *Phys. Rev. A*, 72:042329, Oct 2005.

[GMW86] O. Goldreich, S. Micali, and A. Wigderson. Proofs that yield nothing but their validity and a methodology of cryptographic protocol design. In *27th Annual Symposium on Foundations of Computer Science (SFCS 1986)*, pages 174–187, Oct 1986.

[Gou06] Gilad Gour. Entanglement of collaboration. *Phys. Rev. A*, 74:052307, Nov 2006.

[GQ17a] Joshua A. Grochow and Youming Qiao. Algorithms for group isomorphism via group extensions and cohomology. *SIAM Journal on Computing*, 46(4):1153–1216, 2017.

[GQ17b] Joshua A. Grochow and Youming Qiao. Isomorphism problems in linear algebra. In preparation, 2017.

[GR16] François Le Gall and David J. Rosenbaum. On the group and color isomorphism problems. *arXiv preprint arXiv:1609.08253*, 2016.

[GS06] Gilad Gour and Robert W. Spekkens. Entanglement of assistance is not a bipartite measure nor a tripartite monotone. *Phys. Rev. A*, 73:062331, Jun 2006.

[Gur03] Leonid Gurvits. Classical deterministic complexity of Edmonds' problem and quantum entanglement. In *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing*, STOC '03, pages 10–19, New York, NY, USA, 2003. ACM.

[GW10] Gilad Gour and Nolan R. Wallach. All maximally entangled four-qubit states. *Journal of Mathematical Physics*, 51(11):112201, 2010.

[Hal35] Philip Hall. On representatives of subsets. *Journal of the London Mathematical Society*, s1-10(1):26–30, 1935.

[HHH96] Michałl Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223(1):1 – 8, 1996.

[HHLW10] Aram W. Harrow, Avinatan Hassidim, Debbie W. Leung, and John Watrous. Adaptive versus nonadaptive strategies for quantum channel discrimination. *Phys. Rev. A*, 81:032339, Mar 2010.

[HJ12] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, NY, USA, 2nd edition, 2012.

[HMM+06] M. Hayashi, D. Markham, M. Murao, M. Owari, and S. Virmani. Bounds on multipartite entangled orthogonal state discrimination using local operations and classical communication. *Phys. Rev. Lett.*, 96:040501, Feb 2006.

[HSSH03] Michał Horodecki, Aditi Sen(De), Ujjwal Sen, and Karol Horodecki. Local indistinguishability: More nonlocality with less entanglement. *Phys. Rev. Lett.*, 90:047902, Jan 2003.

[IKQS15] Gábor Ivanyos, Marek Karpinski, Youming Qiao, and Miklos Santha. Generalized wong sequences and their applications to Edmonds' problems. *J. Comput. Syst. Sci.*, 81(7):1373–1386, November 2015.

[IKS10] Gábor Ivanyos, Marek Karpinski, and Nitin Saxena. Deterministic polynomial time algorithms for matrix completion problems. *SIAM Journal on Computing*, 39(8):3736–3751, 2010.

[IQ18] Gabor Ivanyos and Youming Qiao. Algorithms based on *-algebras, and their applications to isomorphism of polynomials with one secret, group isomorphism, and polynomial identity testing. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2357–2376, 2018.

[IQS17a] Gábor Ivanyos, Youming Qiao, and K. V. Subrahmanyam. Non-commutative Edmonds' problem and matrix semi-invariants. *Comput. Complex.*, 26(3):717–763, Sep 2017.

[IQS17b] Gábor Ivanyos, Youming Qiao, and K Venkata Subrahmanyam. Constructive non-commutative rank computation is in deterministic polynomial time. In *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*, volume 67 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 55:1–55:19, Dagstuhl, Germany, 2017.

[Ish04] Satoshi Ishizaka. Bound entanglement provides convertibility of pure entangled states. *Phys. Rev. Lett.*, 93:190501, Nov 2004.

[JFDY06] Zhengfeng Ji, Yuan Feng, Runyao Duan, and Mingsheng Ying. Identification and distance measures of measurement apparatus. *Phys. Rev. Lett.*, 96:200401, May 2006.

[Joh13] Nathaniel Johnston. Non-positive-partial-transpose subspaces can be as large as any entangled subspace. *Phys. Rev. A*, 87:064302, Jun 2013.

[JP99] Daniel Jonathan and Martin B. Plenio. Entanglement-assisted local manipulation of pure quantum states. *Phys. Rev. Lett.*, 83:3566–3569, Oct 1999.

[Kar79] Richard M. Karp. Probabilistic analysis of a canonical numbering algorithm for graphs. In *Proceedings of the AMS Symposium in Pure Mathematics*, volume 34, pages 365–378, 1979.

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *Comput. Complex.*, 13(1/2):1–46, December 2004.

[KIN94] A. D. KING. Moduli of representations of finite dimensional algebras. *The Quarterly Journal of Mathematics*, 45(4):515–530, 1994.

[Kra83] Karl Kraus. *States, effects, and operations : fundamental notions of quantum theory : lectures in mathematical physics at the University of Texas at Austin*. Springer, 1983. Includes bibliographical references.

[KST93] Johannes Köbler, Uwe Schöning, and Jacobo Torán. *The graph isomorphism problem: its structural complexity*. Birkhauser Verlag, Basel, Switzerland, Switzerland, 1993.

[KT10] S. Kıntaş and S. Turgut. Transformations of W-type entangled states. *Journal of Mathematical Physics*, 51(9):092202, 2010.

[KV12] Dmitry Kerner and Victor Vinnikov. Determinantal representations of singular hypersurfaces in $\mathbb{P}^n$. *Advances in Mathematics*, 231(3):1619 – 1654, 2012.

[Li16] Ke Li. Discriminating quantum states: The multiple chernoff distance. *Ann. Statist.*, 44(4):1661–1679, 08 2016.

[Lip78] Richard J. Lipton. The beacon set approach to graph isomorphism. Yale University. Department of Computer Science, 1978.

[LM15] Debbie Leung and William Matthews. On the power of PPT-preserving and non-signalling codes. *IEEE Transactions on Information Theory*, 61(8):4486–4499, Aug 2015.

[Lov79] László Lovász. On determinants, matchings, and random algorithms. In *FCT*, pages 565–574, 1979.

[Lov89] László Lovász. Singular spaces of matrices and their application in combinatorics. *Boletim da Sociedade Brasileira de Matemática - Bulletin/Brazilian Mathematical Society*, 20(1):87–99, Oct 1989.

[LQ17a] Yinan Li and Youming Qiao. Linear algebraic analogues of the graph isomorphism problem and the erdős-rènyi model. In *2017 IEEE 58th Annual*

Symposium on Foundations of Computer Science (FOCS), pages 463–474, Oct 2017.

[LQ17b] Yinan Li and Youming Qiao. On rank-critical matrix spaces. *Differential Geometry and its Applications*, 55:68 – 77, 2017. Geometry and complexity theory.

[LQ18] Yinan Li and Youming Qiao. On reducing isomorphism problems into hidden subgroup problems. In preparation, 2018.

[LQWD18] Yinan Li, Youming Qiao, Xin Wang, and Runyao Duan. Tripartite-to-bipartite entanglement transformation by stochastic local operations and classical communication and the structure of matrix spaces. *Communications in Mathematical Physics*, Jan 2018.

[Luk82] Eugene M. Luks. Isomorphism of graphs of bounded valence can be tested in polynomial time. *J. Comput. Syst. Sci.*, 25(1):42 – 65, 1982.

[Luk90] Eugene M. Luks. Lectures on polynomial-time computation in groups. *Lecture notes*, 1990.

[Luk92] Eugene M. Luks. Computing in solvable matrix groups. In *Proceedings of the 33rd Annual Symposium on Foundations of Computer Science*, pages 111–120, Oct 1992.

[Luk93] Eugene M. Luks. Permutation groups and polynomial-time computation. In *Groups and Computation: Workshop on Groups and Computation, October 7-10, 1991*, volume 11, page 139. American Mathematical Soc., 1993.

[Luk99] Eugene M. Luks. Hypergraph isomorphism and structural equivalence of boolean functions. In *Proceedings of the Thirty-first Annual ACM Symposium on Theory of Computing*, STOC '99, pages 652–658, New York, NY, USA, 1999. ACM.

[LW12] Mark L. Lewis and James B. Wilson. Isomorphism in expanding families of indistinguishable groups. *Groups - Complexity - Cryptology*, 4(1):73110, 2012.

[LWD17] Yinan Li, Xin Wang, and Runyao Duan. Indistinguishability of bipartite states by positive-partial-transpose operations in the many-copy scenario. *Phys. Rev. A*, 95:052346, May 2017.

[MFK94] David Mumford, John Fogarty, and Frances Kirwan. *Geometric invariant theory*. Springer-Verlag, 1994.

[Mil78] Gary L. Miller. On the $N \log N$ isomorphism technique (a preliminary report). In *Proceedings of the Tenth Annual ACM Symposium on Theory of Computing*, STOC '78, pages 51–58, New York, NY, USA, 1978. ACM.

[MP14] Brendan D. McKay and Adolfo Piperno. Practical graph isomorphism, ii. *Journal of Symbolic Computation*, 60:94 – 112, 2014.

[NC11] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, New York, NY, USA, 10th edition, 2011.

[Nie99] M. A. Nielsen. Conditions for a class of entanglement transformations. *Phys. Rev. Lett.*, 83:436–439, Jul 1999.

[O'B94] E. A. O'Brien. Isomorphism testing for p-groups. *Journal of Symbolic Computation*, 17(2):133 – 147, 1994.

[Par04] Kalyanapuram R. Parthasarathy. On the maximal dimension of a completely entangled subspace for finite level quantum systems. *Proceedings Mathematical Sciences*, 114(4):365–374, Nov 2004.

[Per96] Asher Peres. Separability criterion for density matrices. *Phys. Rev. Lett.*, 77:1413–1415, Aug 1996.

[PM07] Marco Piani and Caterina E. Mora. Class of positive-partial-transpose bound entangled states associated with almost any set of pure entangled states. *Phys. Rev. A*, 75:012305, Jan 2007.

[PVMDC05] M. Popp, F. Verstraete, M. A. Martín-Delgado, and J. I. Cirac. Localizable entanglement. *Phys. Rev. A*, 71:042306, Apr 2005.

[Rai01] Eric M. Rains. A semidefinite program for distillable entanglement. *IEEE Transactions on Information Theory*, 47(7):2921–2933, Nov 2001.

[Roc15] Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, 2015.

[Roo38] Thomas G. Room. *The Geometry of Determinantal Loci*. The Cambridge University Press, 1938.

[Ros13] David J. Rosenbaum. Bidirectional collision detection and faster deterministic isomorphism testing. *arXiv preprint arXiv:1304.3935*, 2013.

[Sch26] E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.*, 28:1049–1070, Dec 1926.

[Sch80] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, October 1980.

[SDY05] Xiaoming Sun, Runyao Duan, and Mingsheng Ying. The existence of quantum entanglement catalysts. *IEEE Transactions on Information Theory*, 51(1):75–80, Jan 2005.

[Ser03] Ákos Seress. *Permutation Group Algorithms*, volume 152. Cambridge University Press, 2003.

[Sim78] Charles C. Sims. *Some Group-theoretic Algorithms*, pages 108–124. Springer Berlin Heidelberg, Berlin, Heidelberg, 1978.

[SvdB01] Aidan Schofield and Michel van den Bergh. Semi-invariants of quivers for arbitrary dimension vectors. *Indagationes Mathematicae*, 12(1):125 – 138, 2001.

[SVW05] John A. Smolin, Frank Verstraete, and Andreas Winter. Entanglement of assistance and multipartite state distillation. *Phys. Rev. A*, 72:052317, Nov 2005.

[TGP10] S. Turgut, Y. Gül, and N. K. Pak. Deterministic transformations of multipartite entangled states with tensor rank 2. *Phys. Rev. A*, 81:012317, Jan 2010.

[Tut47] W. T. Tutte. The factorization of linear graphs. *Journal of the London Mathematical Society*, s1-22(2):107–111, 1947.

[VC15] Péter Vrana and Matthias Christandl. Asymptotic entanglement transformation between W and GHZ states. *Journal of Mathematical Physics*, 56(2):022204, 2015.

[VDDMV02] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde. Four qubits can be entangled in nine different ways. *Phys. Rev. A*, 65:052112, Apr 2002.

[Vid99] Guifré Vidal. Entanglement of pure states for a single copy. *Phys. Rev. Lett.*, 83:1046–1049, Aug 1999.

[VPC04] F. Verstraete, M. Popp, and J. I. Cirac. Entanglement versus correlations in spin systems. *Phys. Rev. Lett.*, 92:027901, Jan 2004.

[Wal02] Nolan R. Wallach. An unentangled Gleasons theorem. *Contemporary Mathematics*, 305:291, 2002.

[Wat05] John Watrous. Bipartite subspaces having no bases distinguishable by local operations and classical communication. *Phys. Rev. Lett.*, 95:080505, Aug 2005.

[Wat08] John Watrous. Distinguishing quantum operations having few kraus operators. *Quantum Info. Comput.*, 8(8):819–833, September 2008.

[Wat18] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[WH02] Jonathan Walgate and Lucien Hardy. Nonlocality, asymmetry, and distinguishing bipartite states. *Phys. Rev. Lett.*, 89:147901, Sep 2002.

[Wil09] James B. Wilson. Decomposing p-groups via Jordan algebras. *Journal of Algebra*, 322(8):2642 – 2679, 2009.

[Wil13] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013.

[Wil14] James B. Wilson. 2014 conference on *Groups, Computation, and Geometry* at Colorado State University, co-organized by P. Brooksbank, A. Hulpke, T. Penttila, J. Wilson, and W. Kantor. Personal communication, 2014.

[WL68] Boris Weisfeiler and Andrei A. Lehman. A reduction of a graph to a canonical form and an algebra arising during this reduction. *Nauchno-Technicheskaya Informatsia*, 2(9):12–16, 1968.

[WSHV00] Jonathan Walgate, Anthony J. Short, Lucien Hardy, and Vlatko Vedral. Local distinguishability of multipartite orthogonal quantum states. *Phys. Rev. Lett.*, 85:4972–4975, Dec 2000.

[WXD17] Xin Wang, Wei Xie, and Runyao Duan. Semidefinite programming strong converse bounds for classical capacity. *IEEE Transactions on Information Theory*, PP(99):1–1, 2017.

[XD07] Yu Xin and Runyao Duan. Conditions for entanglement transformation between a class of multipartite pure states with generalized schmidt decompositions. *Phys. Rev. A*, 76:044301, Oct 2007.

[YDY14] Nengkun Yu, Runyao Duan, and Mingsheng Ying. Distinguishability of quantum states by positive operator-valued measures with positive partial transpose. *IEEE Transactions on Information Theory*, 60(4):2069–2079, April 2014.

[YE09] Dong Yang and Jens Eisert. Entanglement combing. *Phys. Rev. Lett.*, 103:220501, Nov 2009.

[YGD14] Nengkun Yu, Cheng Guo, and Runyao Duan. Obtaining a W state from a Greenberger-Horne-Zeilinger state via stochastic local operations and classical communication with a rate approaching unity. *Phys. Rev. Lett.*, 112:160401, Apr 2014.

[ZKT85] V. N. Zemlyachenko, N. M. Korneenko, and R. I. Tyshkevich. Graph isomorphism problem. *Journal of Soviet Mathematics*, 29(4):1426–1481, May 1985.