Quantum Coupling and Strassen Theorem

Li Zhou[*]   Shenggang Ying[†]   Nengkun Yu[‡]

Mingsheng Ying[§]

September 18, 2018

Abstract

We introduce a quantum generalisation of the notion of coupling in probability theory. Several interesting examples and basic properties of quantum couplings are presented. In particular, we prove a quantum extension of Strassen theorem for probabilistic couplings, a fundamental theorem in probability theory that can be used to bound the probability of an event in a distribution by the probability of an event in another distribution coupled with the first.

1   Introduction

Coupling is a powerful technique in probability theory, with which random variables can be linked to or compared with each other. It has been widely used in the studies of random walks and Markov chains, interacting particle systems and diffusions, just name a few, in order to establish limit theorems about them, to develop approximations for them, or to derive correlation inequalities between them [7].

Recently, a very successful application of coupling in computer science was discovered by Barthe et al. [4] that it can serve as a solid mathematical foundation for defining the semantics of probabilistic relational Hoare logic. This discovery enables them to develop a series of powerful proof techniques for reasoning about relational properties of probabilistic computations, in particular, for verification of cryptographic protocols and differential privacy [2, 3, 1, 6].

[*]Department of Computer Science and Technology, Tsinghua University, Beijing, China; and Centre for Quantum Software and Information, School of Software, Faculty of Engineering and Information Technology, University of Technology Sydney, NSW, Australia

[†]Centre for Quantum Software and Information, School of Software, Faculty of Engineering and Information Technology, University of Technology Sydney, NSW, Australia

[‡]Centre for Quantum Software and Information, School of Software, Faculty of Engineering and Information Technology, University of Technology Sydney, NSW, Australia

[§]Centre for Quantum Software and Information, School of Software, Faculty of Engineering and Information Technology, University of Technology Sydney, NSW, Australia; and Institute of Software, Chinese Academy of Sciences, Beijing 100190, China; and Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

There is a simple and natural correspondence between probability theory and quantum theory: probability distributions/density operators (mixed quantum states), marginal distributions/partial traces, and more. This correspondence suggests us to explore the possibility of generalising the coupling techniques for reasoning about quantum systems. We expect that these techniques can help us to extend quantum Hoare logic [10] for proving relational properties between quantum programs and further for verifying quantum cryptographic protocols and differential privacy in quantum computation [11]. But in this paper, we focus on studying quantum couplings themselves.

Strassen theorem [9] is a fundamental theorem in probability theory that can be used to bound the probability of an event in a distribution by the probability of an event in another distribution coupled with the first. The main technical contribution of this paper is proving an elegant (in our opinion) quantum generalisation of Strassen theorem.

2 Background and Basic Definitions

2.1 Probabilistic Coupling

For convenience of the reader, we first briefly recall the basics of probabilistic coupling, following [6]. Let $\mathcal{A}$ be a finite or countably infinite set. A sub-distribution over $\mathcal{A}$ is a mapping $\mu : \mathcal{A} \to [0,1]$ such that $\sum_{a \in \mathcal{A}} \mu(a) \leq 1$. In paricular, if $\sum_{a \in \mathcal{A}} \mu(a) = 1$, then μ is called a distribution over $\mathcal{A}$. For a sub-distribution μ over $\mathcal{A}$, we define:

1. The weight of μ is $|\mu| = \sum_{a \in \mathcal{A}} \mu(a)$;

2. The support of μ is $\text{supp}(\mu) = \{a \in \mathcal{A} : \mu(a) > 0\}$;

3. The probability of an event $S \subseteq \mathcal{A}$ is $\mu(S) = \sum_{a \in S} \mu(a)$.

Moreover, let μ be a joint sub-distribution, i.e. a sub-distribution over Cartesian product $\mathcal{A}_1 \times \mathcal{A}_2$. Then its marginals $\pi_1(\mu), \pi_2(\mu)$ over $\mathcal{A}_1$ and $\mathcal{A}_2$ are, respectively, defined by

$$
\pi_1(\mu)(a_1) = \sum_{a_2 \in \mathcal{A}_2} \mu(a_1, a_2) \text{ for every } a_1 \in \mathcal{A}_1,
$$
$$
\pi_2(\mu)(a_2) = \sum_{a_1 \in \mathcal{A}_1} \mu(a_1, a_2) \text{ for every } a_2 \in \mathcal{A}_2.
$$

Now we can define the notion of coupling.

Definition 1 (Probabilistic Coupling). *Let μ_1, μ_2 be sub-distributions over $\mathcal{A}_1, \mathcal{A}_2$, respectively. Then a sub-distribution μ over $\mathcal{A}_1 \times \mathcal{A}_2$ is called a coupling for (μ_1, μ_2) if $\pi_1(\mu) = \mu_1$ and $\pi_2(\mu) = \mu_2$.*

Here are some simple examples of coupling taken from [6].

Example 1. *Let* **Flip** *be the uniform distribution over booleans, i.e.* $\textbf{Flip}(0) = \textbf{Flip}(1) = \frac{1}{2}$. *Then the following are two couplings for* $(\textbf{Flip}, \textbf{Flip})$:

1. *Identity coupling:* $\mu_{\mathrm{id}}(a_1, a_2) = \begin{cases} \frac{1}{2} & \text{if } a_1 = a_2, \\ 0 & \text{otherwise.} \end{cases}$

2. *Negation coupling:* $\mu_{\neg}(a_1, a_2) = \begin{cases} \frac{1}{2} & \text{if } \neg a_1 = a_2, \\ 0 & \text{otherwise.} \end{cases}$

More generally, let $\textbf{Unif}_{\mathcal{A}}$ *be the uniform distribution over a finite nonempty set* $\mathcal{A}$, *i.e.* $\textbf{Unif}_{\mathcal{A}}(a) = \frac{1}{|\mathcal{A}|}$ *for every* $a \in \mathcal{A}$. *Then each bijection* $f : \mathcal{A} \to \mathcal{A}$ *yields a coupling* μ_f *for* $(\textbf{Unif}_{\mathcal{A}}, \textbf{Unif}_{\mathcal{A}})$:

$$\mu_f(a_1, a_2) = \begin{cases} \frac{1}{|\mathcal{A}|} & \text{if } f(a_1) = a_2, \\ 0 & \text{otherwise.} \end{cases}$$

Example 2. *For any sub-distribution* μ *over* $\mathcal{A}$, *the identity coupling for* (μ, μ) *is:* $\mu_{\mathrm{id}}(a_1, a_2) = \begin{cases} \mu(a) & \text{if } a_1 = a_2 = a, \\ 0 & \text{otherwise.} \end{cases}$

Example 3. *For any distributions* μ_1, μ_2 *over* $\mathcal{A}_1, \mathcal{A}_2$, *respectively, the independent or trivial coupling is:* $\mu_{\times}(a_1, a_2) = \mu_1(a_1) \cdot \mu_2(a_2)$.

Obviously, coupling for a pair of distributions is not unique. Then the notion of lifting can be introduced to choose a desirable coupling.

Definition 2 (Probabilistic Lifting). *Let* μ_1, μ_2 *be sub-distributions over* $\mathcal{A}_1, \mathcal{A}_2$, *respectively, and let* $\mathcal{R} \subseteq \mathcal{A}_1 \times \mathcal{A}_2$ *be a relation. Then a sub-distribution* μ *over* $\mathcal{A}_1 \times \mathcal{A}_2$ *is called a witness for the* $\mathcal{R}$-*lifting of* (μ_1, μ_2) *if:*

1. μ *is a coupling for* (μ_1, μ_2);

2. $\mathrm{supp}(\mu) \subseteq \mathcal{R}$.

Whenever a witness exists, we say that μ_1 *and* μ_2 *are related by the* $\mathcal{R}$-*lifting and write* $\mu_1 \mathcal{R}^{\#} \mu_2$.

Example 4. 1. *Coupling* μ_f *in Example 1 is a witness for the lifting* $\textbf{Unif}_{\mathcal{A}}$ $\{(a_1, a_2) | f(a_1) = a_2\}^{\#} \textbf{Unif}_{\mathcal{A}}$.

2. *Coupling* μ_{id} *in Example 2 is a witness for the lifting* $\mu =^{\#} \mu$.

3. *Coupling* $\mu_{\times}$ *in Example 3 is a witness for the lifting* $\mu_1 T^{\#} \mu_2$, *where* $T = \mathcal{A}_1 \times \mathcal{A}_2$.

Proposition 1. 1. *Let* μ_1, μ_2 *be sub-distributions over* $\mathcal{A}_1, \mathcal{A}_2$, *respectively. If there exists a coupling for* (μ_1, μ_2), *then* $|\mu_1| = |\mu_2|$.

2. *Let* μ_1, μ_2 *be sub-distributions over the same* $\mathcal{A}$. *Then* $\mu_1 = \mu_2$ *if and only if* $\mu_1 =^{\#} \mu_2$.

2.2 Quantum Coupling

With the correspondence of probability distributions/density operators (mixed quantum states) and marginal distributions/partial traces mentioned in the Introduction, we can introduce the notion of quantum coupling. To this end, let us first recall several basic notions from quantum theory; for details, we refer to [8].

Suppose that $\mathcal{H}$ is a finite-dimensional Hilbert space. Let $\text{Herm}(\mathcal{H})$ be the set of Hermitian matrices in $\mathcal{H}$. Let $\text{Pos}(\mathcal{H})$ be the set of positive (semidefinite) matrices in $\mathcal{H}$, and $\mathcal{D}(\mathcal{H}) \subset \text{Pos}(\mathcal{H})$ is the set of partial density operators, *i.e.*, positive (semidefinite) matrices with trace one. A positive operator ρ in $\mathcal{H}$ is called a partial density operator if its trace $tr(\rho) = \sum_i \langle i|\rho|i \rangle \leq 1$, where $\{|i\rangle\}$ is an orthonormal basis of $\mathcal{H}$.

We define its support:

$$\text{supp}(\rho) = \text{span}\{\text{eigenvectors of } \rho \text{ with nonzero eigenvalues}\}$$
$$= \text{span}\{|\psi\rangle \mid tr(\rho|\psi\rangle\langle\psi|) = 0\}^{\perp}.$$

If A is an observable, i.e. Hermitian operator, in $\mathcal{H}$, then its expectation in state ρ is $\langle A \rangle_\rho = tr(A\rho)$. Furthermore, let $\mathcal{H}_1, \mathcal{H}_2$ be two Hilbert space. Then partial trace over $\mathcal{H}_1$ is a mapping $tr_1(\cdot)$ from operators in $\mathcal{H}_1 \otimes \mathcal{H}_2$ to operators in $\mathcal{H}_2$ defined by

$$tr_1(|\varphi_1\rangle\langle\psi_1| \otimes |\varphi_2\rangle\langle\psi_2|) = \langle\psi_1|\varphi_1\rangle \cdot |\varphi_2\rangle\langle\psi_2|$$

for all $|\varphi_1\rangle, |\psi_1\rangle \in \mathcal{H}_1$ and $|\varphi_2\rangle, |\psi_2\rangle \in \mathcal{H}_2$ together with linearity. The partial trace $tr_2(\cdot)$ over $\mathcal{H}_2$ can be defined dually.

Now we are ready to define the concept of coupling.

Definition 3 (Quantum Coupling). *Let* $\rho_1 \in \mathcal{D}(\mathcal{H})$ *and* $\rho_2 \in \mathcal{D}(\mathcal{H}_2)$. *Then* $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ *is called a coupling for* (ρ_1, ρ_2) *if* $tr_1(\rho) = \rho_2$ *and* $tr_2(\rho) = \rho_1$.

This is actually a very special case of the famous quantum marginal problem, see [12, 13, 14, 15] as a very incompleted list for recent development.

Example 5. *Let* $\mathcal{H}$ *be a Hilbert space and* $\mathcal{B} = \{|i\rangle\}$ *an orthonormal basis of* $\mathcal{H}$. *Then the uniform density operator on* $\mathcal{H}$ *is*

$$\mathbf{Unif}_\mathcal{H} = \frac{1}{d} \sum_i |i\rangle\langle i|$$

where $d = \dim \mathcal{H}$ *is the dimension of* $\mathcal{H}$. *Indeed, the uniform density operator on* $\mathcal{H}$ *is unique and independent with the choice of orthonormal basis. For each unitary operator* U *in* $\mathcal{H}$, *we write* $U(\mathcal{B}) = \{U|i\rangle\}$, *which is also an orthonormal basis of* $\mathcal{H}$. *Then*

$$\rho_U = \frac{1}{d} \sum_i (|i\rangle U|i\rangle)(\langle i|\langle i|U^\dagger)$$

is a coupling for $(\mathbf{Unif}_\mathcal{H}, \mathbf{Unif}_\mathcal{H})$. *In general, for different* U *and* U', $\rho_U \neq \rho_{U'}$, *though they are both the couplings for* $(\mathbf{Unif}_\mathcal{H}, \mathbf{Unif}_\mathcal{H})$.

Example 6. *Let ρ be a partial density operator in $\mathcal{H}$. Then by the spectral decomposition theorem, ρ can be written as $\rho = \sum_i p_i |i\rangle\langle i|$ for some orthonormal basis $\mathcal{B} = \{|i\rangle\}$ and $p_i \geq 0$ with $\sum_i p_i \leq 1$. We define:*

$$\rho_{\mathrm{id}(\mathcal{B})} = \sum_i p_i |ii\rangle\langle ii|.$$

Then it is to see that $\rho_{\mathrm{id}(\mathcal{B})}$ is a coupling for (ρ, ρ). A difference between this example and Example 2 is that ρ can be decomposed with other orthonormal bases, say $\mathcal{D} = \{|j\rangle\}$: $\rho = \sum_j q_j |j\rangle\langle j|$. In general, $\rho_{\mathrm{id}(\mathcal{B})} \neq \rho_{\mathrm{id}(\mathcal{D})}$, and we can define a different coupling:

$$\rho_{\mathrm{id}(\mathcal{D})} = \sum_j q_j |jj\rangle\langle jj|$$

for (ρ, ρ).

Example 7. *Let $\rho_1 \in \mathcal{D}(\mathcal{H}_1)$ and $\rho_2 \in \mathcal{D}(\mathcal{H}_2)$ be density operators. Then tensor product $\rho_\otimes = \rho_1 \otimes \rho_2$ is a coupling for (ρ_1, ρ_2).*

The notion of lifting can also be easily generalised into the quantum setting.

Definition 4 (Quantum Lifting). *Let $\rho_1 \in \mathcal{D}(\mathcal{H}_1)$ and $\rho_2 \in \mathcal{D}(\mathcal{H}_2)$, and let $\mathcal{X}$ be a subspace of $\mathcal{H}_1 \otimes \mathcal{H}_2$. Then $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is called a witness of the lifting $\rho_1 \mathcal{X}^{\#} \rho$ if:*

1. *ρ is a coupling for (ρ_1, ρ_2);*

2. *$supp(\rho) \subseteq \mathcal{X}$.*

Example 8.   1. *The coupling ρ_U in Example 5 is a witness for the lifting:*

$$\mathbf{Unif}_{\mathcal{H}} \mathcal{X}(\mathcal{B}, U)^{\#} \mathbf{Unif}_{\mathcal{H}}$$

where $\mathcal{X}(\mathcal{B}, U) = span\{|i\rangle U |i\rangle\}$ is a subspace of $\mathcal{H} \otimes \mathcal{H}$.

2. *The coupling $\rho_{\mathrm{id}(\mathcal{B})}$ in Example 6 is a witness of the lifting $\rho =_{\mathcal{B}}^{\#} \rho$, where $=_{\mathcal{B}} = span\{|ii\rangle\}$ defined by the orthonormal basis $\mathcal{B} = \{|i\rangle\}$ is a subspace of $\mathcal{H} \otimes \mathcal{H}$. It is interesting to note that the maximal entangled state $|\Psi\rangle = \frac{1}{\sqrt{d}} \sum_i |ii\rangle$ is in $=_{\mathcal{B}}$.*

3. *The coupling $\rho_\otimes$ in Example 7 is a witness of the lifting $\rho_1 (\mathcal{H}_1 \otimes \mathcal{H}_2)^{\#} \rho_2$.*

As a quantum generalisation of Proposition 1, we have:

Proposition 2.   1. *Let $\rho_1 \in \mathcal{D}(\mathcal{H}_1)$ and $\rho_2 \in \mathcal{D}(\mathcal{H}_2)$. If there exists a coupling for (ρ_1, ρ_2), then $tr(\rho_1) = tr(\rho_2)$.*

2. *Let $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$. Then $\rho_1 = \rho_2$ if and only if $\exists$ orthonormal basis $\mathcal{B}$ s.t. $\rho_1 =_{\mathcal{B}}^{\#} \rho_2$.*

Proof. Part 1 and Part 2 ($\Rightarrow$) are obvious. Here, we prove Part 2 ($\Leftarrow$). If $\rho_1 =^{\#}_{\mathcal{B}} \rho_2$, then there exists a coupling ρ for (ρ_1, ρ_2) such that $supp(\rho) \subseteq span\{|ii\rangle\}$, where $\mathcal{B} = \{|i\rangle\}$. Then we have: $\rho = \sum_j p_j |\Psi_j\rangle\langle\Psi_j|$ for some $|\Psi_j\rangle \in span\{|ii\rangle\}$ and p_j. Furthermore, for each j, we can write: $|\Psi_j\rangle = \sum_i \alpha_{ji}|ii\rangle$. Then it is routine to show that $tr_1(|\Psi_j\rangle\langle\Psi_j|) = tr_2(|\Psi_j\rangle\langle\Psi_j|) = \sum_i |\alpha_{ji}|^2 |i\rangle\langle i|$. Therefore, it holds that $\rho_1 = tr_2(\rho) = \sum_j p_j tr_2(|\Psi_j\rangle\langle\Psi_j|) = \sum_j p_j tr_1(|\Psi_j\rangle\langle\Psi_j|) = tr_1(\rho) = \rho_2$. $\square$

3   Quantum Strassen Theorem

As mentioned in the Introduction, a fundamental theorem for probabilistic coupling is the following:

Theorem 1 (Strassen Theorem). *Let μ_1, μ_2 be sub-distributions over $\mathcal{A}_1, \mathcal{A}_2$, respectively. Then*

$$\mu_1 \mathcal{R}^{\#} \mu_2 \Rightarrow \forall S \subseteq \mathcal{A}_1.\ \mu_1(S) \leq \mu_2(\mathcal{R}(S)) \tag{1}$$

where $\mathcal{R}(S)$ is the image of S under $\mathcal{R}$: $\mathcal{R}(S) = \{a_2 \in \mathcal{A}_2 | \exists a_1 \in S \text{ s.t. } (a_1, a_2) \in \mathcal{R}\}$. The converse of (1) holds if $|\mu_1| = |\mu_2|$.

In this section, we prove a quantum generalisation of the above Strassen Theorem. For this purpose, for any subspace $\mathcal{X}$ of $\mathcal{H}_1 \otimes \mathcal{H}_2$, we use $P_{\mathcal{X}}$ and $P_{\mathcal{X}^{\perp}}$ to denote the projections on $\mathcal{X}$ and $\mathcal{X}^{\perp}$ (the ortho-complement of $\mathcal{X}$), respectively. We use I_1, I_2, I_{12} to denote the identity matrix of $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_{12}$, respectively. $\langle \cdot, \cdot \rangle$ is employed to denote the inner product of matrices living in the same space,

$$\langle A, B \rangle = \text{tr}(A^{\dagger}B)$$

Then a quantum Strassen theorem can be stated as follows:

Theorem 2 (Quantum Strassen Theorem). *For any two partial density operators ρ_1 in $\mathcal{H}_1$ and ρ_2 in $\mathcal{H}_2$ with $\text{tr}(\rho_1) = \text{tr}(\rho_2)$, and for any subspace $\mathcal{X}$ of $\mathcal{H}_1 \otimes \mathcal{H}_2$, the following three statements are equivalent:*

1. *$\rho_1 \mathcal{X}^{\#} \rho_2$;*

2. *For all observables (Hermitian operators) Y_1 in $\mathcal{H}_1$ and Y_2 in $\mathcal{H}_2$ satisfying $P_{\mathcal{X}^{\perp}} \geq Y_1 \otimes I_2 - I_1 \otimes Y_2$, it holds that*

$$\text{tr}(\rho_1 Y_1) \leq \text{tr}(\rho_2 Y_2). \tag{2}$$

3. *For all positive observables Y_1 in $\mathcal{H}_1$ and Y_2 in $\mathcal{H}_2$ satisfying $P_{\mathcal{X}^{\perp}} \geq Y_1 \otimes I_2 - I_1 \otimes Y_2$, it holds that $\text{tr}(\rho_1 Y_1) \leq \text{tr}(\rho_2 Y_2)$.*

Proof. ($1 \Rightarrow 2$) Suppose ρ is a witness of the lifting $\rho_1 \mathcal{X}^{\#} \rho_2$. Then for all observables (Hermition operators) Y_1 in $\mathcal{H}_1$ and Y_2 in $\mathcal{H}_2$, if $P_{\mathcal{X}^{\perp}} \geq Y_1 \otimes I_2 -$

$I_1 \otimes Y_2$, then we have:

$$\text{tr}(\rho_1 Y_1) = \text{tr}(\rho(Y_1 \otimes I_2)) \tag{3}$$
$$\leq \text{tr}(\rho(P_{\mathcal{X}^\perp} + I_1 \otimes Y_2)) \tag{4}$$
$$= \text{tr}(\rho(I_1 \otimes Y_2)) \tag{5}$$
$$= \text{tr}(\rho_2 Y_2). \tag{6}$$

Equalities (3) and (6) are derived from the condition that ρ is a coupling for (ρ_1, ρ_2); that is, $\text{tr}_2(\rho) = \rho_1$ and $\text{tr}_1(\rho) = \rho_2$, (4) is due to the assumption for Y_1 and Y_2, and (5) is trivial as $\text{supp}(\rho) \subseteq \mathcal{X}$, so $\text{tr}(\rho P_{\mathcal{X}^\perp}) = 0$.

$(2 \Rightarrow 1)$ Let us first define the semidefinite program (Φ, A, B):

Primal problem		Dual problem	
maximize:	$\langle A, X \rangle$	minimize:	$\langle B, Y \rangle$
subject to:	$\Phi(X) = B,$	subject to:	$\Phi^*(Y) \geq A,$
	$X \in \text{Pos}(\mathcal{H}_1 \otimes \mathcal{H}_2).$		$Y \in \text{Herm}(\mathcal{H}_1 \oplus \mathcal{H}_2).$

where:

$$A = P_{\mathcal{X}}, \quad B = \begin{bmatrix} \rho_1 & \\ & \rho_2 \end{bmatrix},$$

$$\Phi(X) = \begin{bmatrix} \text{tr}_2(X) & \\ & \text{tr}_1(X) \end{bmatrix},$$

$$\Phi^*(Y) = \Phi^* \begin{bmatrix} Y_1 & \cdot \\ \cdot & Y_2 \end{bmatrix} = Y_1 \otimes I_2 + I_1 \otimes Y_2.$$

To show that the above problems are actually primal and dual, respectively, we only need to check the following equality:

$$\forall\ M,\ N,\ \langle \Phi(M), N \rangle = \text{tr}(\text{tr}_2(M)N_1 + \text{tr}_1(M)N_2)$$
$$= \text{tr}(M(N_1 \otimes I_2) + M(I_1 \otimes N_2))$$
$$= \langle M, \Phi^*(N) \rangle.$$

Moreover, the strong duality holds for this semidefinite program as we can check that the primal feasible set are not empty and there exists a Hermitian operator Y for which $\Phi^*(Y) > A$:

Primal feasible set $\mathcal{A} = \{X \in \text{Pos}(\mathcal{H}_1 \otimes \mathcal{H}_2) : \Phi(X) = B\} \ni \dfrac{1}{\text{tr}(\rho_1)}\rho_1 \otimes \rho_2$

Choose $Y = I_1 \oplus I_2 \in \text{Herm}(\mathcal{H}_1 \oplus \mathcal{H}_2),\ \Phi^*(Y) = 2I_{12} > P_{\mathcal{X}}.$

So, $\max\langle P_R, X \rangle = \min\langle B, Y \rangle = \min\{\langle \rho_1, Y_1 \rangle + \langle \rho_2, Y_2 \rangle\}$. Now, let us consider the following condition:

(A): For all observable (Hermitian operators) Y_1 in $\mathcal{H}_1$ and Y_2 in $\mathcal{H}_2$ satisfy $Y_1 \otimes I_2 + I_1 \otimes Y_2 \geq P_{\mathcal{X}}$, then

$$\langle B, Y \rangle = \text{tr}(\rho_1 Y_1 + \rho_2 Y_2) \geq \text{tr}\rho_1.$$

If condition (A) holds, then $\min\langle B, Y \rangle \geq \text{tr}\rho_1$. Still remember that $\max\langle P_{\mathcal{X}}, X \rangle \leq \text{tr}X = \text{tr}\rho_1$. Due to the strong duality, we have $\max\langle P_{\mathcal{X}}, X \rangle = \text{tr}\rho_1$. So, X_{max} which maximizes $\langle P_{\mathcal{X}}, X \rangle$ must satisfy $\langle P_{\mathcal{X}}, X_{max} \rangle = \text{tr}\rho_1 = \text{tr}X_{max}$. Consequently, $\text{supp}X_{max} \subseteq \mathcal{X}$; in other words, X_{max} is a witness of $\rho_1 \mathcal{X}^{\#} \rho_2$. Therefore, condition (A) $\Rightarrow \rho_1 R^{\#} \rho_2$. On the other hand, condition (A) is equivalent to statement 2 of the theorem. Indeed, this is not difficult to prove as if we replace $Y_1' = I_1 - Y_1$ in condition (A), then

$$Y_1 \in \text{Herm}(\mathcal{H}_1) \Longleftrightarrow Y_1' \in \text{Herm}(\mathcal{H}_1)$$
$$Y_1 \otimes I_2 + I_1 \otimes Y_2 \geq P_{\mathcal{X}} \Longleftrightarrow I_1 \otimes I_2 - P_{\mathcal{X}} \geq Y_1' \otimes I_2 - I_1 \otimes Y_2$$
$$\Longleftrightarrow P_{\mathcal{X}}^{\perp} \geq Y_1' \otimes I_2 - I_1 \otimes Y_2$$
$$\text{tr}(\rho_1 Y_1 + \rho_2 Y_2) \geq \text{tr}\rho_1 \Longleftrightarrow \text{tr}(\rho_2 Y_2) \geq \text{tr}(\rho_1 I_1) - \text{tr}(\rho_1(I_1 - Y_1'))$$
$$\Longleftrightarrow \text{tr}(\rho_2 Y_2) \geq \text{tr}(\rho_1 Y_1').$$

From the above, we can directly derive statement 2. In summary, we have: statement $2 \Leftrightarrow$ condition (A) $\Rightarrow \rho_1 R^{\#} \rho_2$.

$(2 \Rightarrow 3)$ Obvious.

$(3 \Rightarrow 2)$ We only need to show that, for any two observables Y_1 in $\mathcal{H}_1$ and Y_2 in $\mathcal{H}_2$ satisfy $P_{\mathcal{X}}^{\perp} \geq Y_1 \otimes I_2 - I_1 \otimes Y_2$, there exist two positive observables Y_1' in $\mathcal{H}_1$ and Y_2' in $\mathcal{H}_2$ such that $P_{\mathcal{X}}^{\perp} \geq Y_1' \otimes I_2 - I_1 \otimes Y_2'$ and

$$\text{tr}(\rho_1 Y_1) \leq \text{tr}(\rho_2 Y_2) \Longleftrightarrow \text{tr}(\rho_1 Y_1') \leq \text{tr}(\rho_2 Y_2').$$

Note that Y_1 and Y_2 are Hermitian, so their eigenvalues are real, and we can define $\lambda = \min\{\text{eigenvalues of } Y_1 \text{ and } Y_2\}$. Choose $Y_1' = Y_1 - \lambda I_1$ and $Y_2' = Y_2 - \lambda I_2$. Obviously, Y_1' and Y_2' are positive observables, and satisfy

$$P_{\mathcal{X}}^{\perp} \geq Y_1 \otimes I_2 - I_1 \otimes Y_2$$
$$= Y_1 \otimes I_2 - \lambda I_1 \otimes I_2 + \lambda I_1 \otimes I_2 - I_1 \otimes Y_2$$
$$= Y_1' \otimes I_2 - I_1 \otimes Y_2'.$$

Moreover, as $\text{tr}(\rho_1) = \text{tr}(\rho_2)$, we have

$$\text{tr}(\rho_1 Y_1) \leq \text{tr}(\rho_2 Y_2) \Longleftrightarrow \text{tr}(\rho_1 Y_1) - \text{tr}(\rho_1 \lambda I_1) \leq \text{tr}(\rho_2 Y_2) - \text{tr}(\rho_2 \lambda I_2)$$
$$\Longleftrightarrow \text{tr}(\rho_1 Y_1') \leq \text{tr}(\rho_2 Y_2').$$

$\square$

Remark: In the above proof, it is indeed naturally to employing the methods of semidefinite programming. In [6], Hsu deliberately constructs a flow network,

and then using the max-flow min-cut theorem to prove the Strassen theorem in the finite case. Essentially, the max-flow min-cut theorem is a special case of the duality theorem for linear programs (LP). Considering the fact that quantum states, quantum operations and so on are all described by matrices, similar to LP, semi-definite programming (SDP) is a powerful and widely used method of convex optimization in quantum theory. Indeed, when all matrices appeared in a SDP are diagonal, then the SDP reduces to LP. In the following section, we will see that in the degenerate case, quantum Strassen theorem also reduces to the classical Strassen theorem.

4   Classical Reduction of Quantum Strassen Theorem

At the first glance, Theorem 1 (Strassen Theorem for Probabilistic Coupling) and Theorem 2 (Quantum Strassen Theorem) are very different. In this section, we show that Theorem 2 is indeed a quantum generalisation of Theorem 1.

To this end, let μ_1 be a sub-distribution over $[m]$ ($[m] = \{i \in \mathbb{N} \mid 1 \le i \le m\}$) and μ_2 over $[n]$. And the corresponding degenerate partial density operators (quantum states) are:

$$
\rho_1 = \begin{bmatrix} \mu_1(1) & & & \\ & \mu_1(2) & & \\ & & \ddots & \\ & & & \mu_1(m) \end{bmatrix}, \quad \rho_2 = \begin{bmatrix} \mu_2(1) & & & \\ & \mu_2(2) & & \\ & & \ddots & \\ & & & \mu_2(n) \end{bmatrix}
$$

in $\mathcal{H}_1 = \mathrm{span}\{|i\rangle : i \in [m]\}$ and $\mathcal{H}_2 = \mathrm{span}\{|j\rangle : j \in [n]\}$, respectively. Furthermore, let $R \subseteq \{(i,j) \big| i \in [m],\ j \in [n]\}$ be a classical relation from $[m]$ to $[n]$. Then the corresponding (quantum relation) subspace of $\mathcal{H}_1 \otimes \mathcal{H}_2$ is defined as

$$
\mathcal{X}_R = \mathrm{span}\{|i\rangle|j\rangle \,\big|\, (i,j) \in R\}.
$$

Based on the above definition of the degenerate case, in the rest part of this section, Proposition 3 shows that the left hand side of Eqn.(1) in Theorem 1 is equivalent to the statement *1* in Theorem 2, while Proposition 4 states the equivalence of the right hand side of Eqn.(1) in Theorem 1 and the statement *3* in Theorem 2, concluding that Theorem 1 (Strassen Theorem) is indeed a reduction of Theorem 2 (Quantum Strassen Theorem).

The following proposition indicates that probabilistic lifting is a special case of quantum lifting.

Proposition 3. $\mu_1 R^\# \mu_2 \iff \rho_1 \mathcal{X}_R^\# \rho_2$.

Proof. ($\Rightarrow$) Suppose that there is a witness μ of the lifting $\mu_1 R^\# \mu_2$. We define the partial density operator:

$$
\rho : \langle i|\langle j|\rho|i'\rangle|j'\rangle = \begin{cases} \mu(i,j) & i = i',\ j = j' \\ 0 & i \ne i' \text{ or } j \ne j' \end{cases}.
$$

9

It is easy to check:

$$\langle i|\mathrm{tr}_2(\rho)|i'\rangle = \sum_{j=1}^n \langle i|\langle j|\rho|i'\rangle|j\rangle = \begin{cases} \sum_{j=1}^n \mu(i,j) = \mu_1(i) & i = i' \\ 0 & i \neq i' \end{cases},$$

$$\langle j|\mathrm{tr}_2(\rho)|j'\rangle = \sum_{i=1}^n \langle i|\langle j|\rho|i\rangle|j'\rangle = \begin{cases} \sum_{i=1}^m \mu(i,j) = \mu_2(j) & j = j' \\ 0 & j \neq j' \end{cases}.$$

So, $\mathrm{tr}_2(\rho) = \rho_1$ and $\mathrm{tr}_1(\rho) = \rho_2$; that is, ρ is a coupling for (ρ_1, ρ_2). Furthermore, we have:

$$\begin{aligned}
\mathrm{tr}(\rho P_{\mathcal{X}_R}) &= \sum_{(i,j)\in R} \langle i|\langle j|\rho|i\rangle|j\rangle \\
&= \sum_{(i,j)\in R} \mu(i,j) \\
&= \sum_{(i,j)\in R} \mu(i,j) + \sum_{(i,j)\notin R} \mu(i,j) \\
&= \mathrm{tr}(\rho)
\end{aligned}$$

Thus, $\mathrm{supp}(\rho) \subseteq \mathcal{X}_R$, and ρ is a witness of the quantum lifting $\rho_1 \mathcal{X}_R^{\#} \rho_2$.

($\Leftarrow$) Suppose there is a witness ρ of the quantum lifting $\rho_1 \mathcal{X}_R^{\#} \rho_2$. Let us construct the joint sub-distribution μ:

$$\mu(i,j) = \langle i|\langle j|\rho|i\rangle|j\rangle \text{ for all } i,j.$$

It is easy to check:

$$\sum_{j=1}^n \mu(i,j) = \sum_{j=1}^n \langle i|\langle j|\rho|i\rangle|j\rangle = \langle i|\rho_1|i\rangle = \mu_1(i),$$

$$\sum_{i=1}^m \mu(i,j) = \sum_{i=1}^m \langle i|\langle j|\rho|i\rangle|j\rangle = \langle j|\rho_2|j\rangle = \mu_2(j).$$

Also, if $(i,j) \notin R$, then $|i\rangle|j\rangle \perp \mathcal{X}_R$, then

$$\mu(i,j) = \langle i|\langle j|\rho|i\rangle|j\rangle = \mathrm{tr}(\rho|i\rangle|j\rangle\langle i|\langle j|) = 0$$

as $\mathrm{supp}(\rho) \subseteq \mathcal{X}_R$. Thus, $\mathrm{supp}(\mu) \subseteq R$, and μ is a witness of the lifting $\mu_1 R^{\#} \mu_2$. $\qquad\square$

The following proposition further shows that in the degenerate case, inequality (2) to (1). Surprisingly, such a reduction can be realized even without the condition of lifting.

10

Proposition 4. *Two statements are equivalent:*

1. *For any $S \subseteq [m]$, $\mu_1(S) \leq \mu_2(R(S))$;*

2. *For all positive observables Y_1 in $\mathcal{H}_1$ and Y_2 in $\mathcal{H}_2$ satisfy $P_{\mathcal{X}_R}^{\perp} \geq Y_1 \otimes I_2 - I_1 \otimes Y_2$, then*

$$\mathrm{tr}(\rho_1 Y_1) \leq \mathrm{tr}(\rho_2 Y_2)$$

Proof. As ρ_1, ρ_2 and $P_{\mathcal{X}_R}$ are diagonal density operators, so we only need to consider those Y_1 and Y_2 which are also diagonal. We use the notation $Y_{1,i} = (Y_1)_{ii}$ and $Y_{2,j} = (Y_2)_{jj}$ for simplicity. Then it holds that

$$P_{\mathcal{X}_R}^{\perp} \geq Y_1 \otimes I_2 - I_1 \otimes Y_2 \iff \forall\, i,j \left\{ \begin{array}{ll} Y_{2,j} \geq Y_{1,i} & (i,j) \in R \\ Y_{2,j} \geq Y_{1,i} - 1 & (i,j) \notin R \end{array} \right.$$

Now we need a technical lemma:

Lemma 1. *The following two statements are equivalent:*

1'. *If $Z_{1,i} \in \{0,1\}$, $Z_{2,j} \in \{0,1\}$, then*

$$\forall\, i,j \left\{ \begin{array}{ll} Z_{2,j} \geq Z_{1,i} & (i,j) \in R \\ Z_{2,j} \geq Z_{1,i} - 1 & (i,j) \notin R \end{array} \right. \Rightarrow \sum_{i=1}^{m} \mu_1(i) Z_{1,i} \leq \sum_{j=1}^{n} \mu_2(j) Z_{2,j}$$

2'. *If $Y_{1,i} \geq 0$, $Y_{2,j} \geq 0$, then*

$$\forall\, i,j \left\{ \begin{array}{ll} Y_{2,j} \geq Y_{1,i} & (i,j) \in R \\ Y_{2,j} \geq Y_{1,i} - 1 & (i,j) \notin R \end{array} \right. \Rightarrow \sum_{i=1}^{m} \mu_1(i) Y_{1,i} \leq \sum_{j=1}^{n} \mu_2(j) Y_{2,j}$$

where Z_1, Z_2 are also diagonal matrices, and $Z_{1,i} = (Z_1)_{ii}$, $Z_{2,j} = (Z_2)_{jj}$.

For readability, let us first use this lemma to finish the proof of the proposition, but postpone the proof of the lemma itself to the end of this section. As

$$\mathrm{tr}(\rho_1 Y_1) = \sum_{i=1}^{m} (\rho_1)_{ii} (Y_1)_{ii} = \sum_{i=1}^{m} \mu_1(i) Y_{1,i},$$

$$\mathrm{tr}(\rho_2 Y_2) = \sum_{j=1}^{m} (\rho_2)_{jj} (Y_2)_{jj} = \sum_{j=1}^{n} \mu_2(j) Y_{2,j},$$

it is direct to see that statement 2 of the proposition is equivalent to statement 2' of the above lemma. For the statement 1' of the above, we can define the set $S = \{i \in [m] \mid Z_{1,i} = 1\}$ and $T = \{j \in [n] \mid Z_{2,j} = 1\}$, then it is equivalent to:

$$\forall\, S \subseteq [m],\ T \subseteq [n],\ R(S) \subseteq T \ \Rightarrow\ \mu_1(S) \leq \mu_2(T),$$

which is exactly the statement 1 of the proposition. Therefore, using the above lemma, we see that statements 1 and 2 in the proposition are equivalent. $\square$

Combining Propositions 3 and 4, we see that Theorem 1 (Strassen Theorem for probabilistic coupling) is a reduction of Theorem 2 (Quantum Strassen Theorem).

To conclude this section, let us present the following:

Proof of Lemma 1. $(2' \Rightarrow 1')$ This is trivial as statement $1'$ is a special case of statement $2'$.

$(1' \Rightarrow 2')$ For any Y_1, we can construct a decreasing sequence $Z_{11} > \cdots > Z_{1k} > Z_{1(k+1)} > \cdots$ such that:

$$Y_1 = \sum_k \lambda_k Z_{1k}, \quad \lambda_k \geq 0.$$

We further define $S_k = \{i \in [m] \mid Z_{1k,i} = 1\}$ and the corresponding $T_k = R(S_k)$. Then, $\{S_k\}$ is a strictly decreasing sequence; that is, $S_k \supset S_{k+1}$ for all k, and $\{T_k\}$ is a non-increasing sequence; that is, $T_k \supseteq T_{k+1}$ for all k. Let us also define Z_{2k}:

$$Z_{2k,j} = \left\{ \begin{array}{ll} 1 & j \in T_k \\ 0 & j \notin T_k \end{array} \right.$$

and a new operator Y_{2min}:

$$Y_{2min} = \sum_k \lambda_k Z_{2k}.$$

Note that any pair of Z_{1k} and Z_{2k} satisfy statement $1'$. Then

$$
\begin{aligned}
\sum_{i=1}^m \mu_1(i) Y_{1,i} = \sum_{i=1}^m \mu_1(i) \sum_k \lambda_k Z_{1k,i} &= \sum_k \lambda_k \sum_{i=1}^m \mu_1(i) Z_{1k,i} \\
&\leq \sum_k \lambda_k \sum_{j=1}^n \mu_2(j) Z_{2k,j} = \sum_{j=1}^n \mu_2(j) \sum_k \lambda_k Z_{2k,j} \\
&= \sum_{j=1}^n \mu_2(j) Y_{2min,j}
\end{aligned}
$$

Now it suffices to prove that for any Y_2 satisfying the condition in statement $2'$, we have $Y_2 \geq Y_{2min}$. To show this, let us use $\mathbb{I}(\cdot)$ to represent the indication function, and consider the following two cases:

- Case 1: $t \notin T_1$. Then of course, $\forall \, k : \, t \notin T_k$, so,

$$Y_{2min,t} = \sum_k \lambda_k Z_{2k,t} = \sum_k \lambda_k \mathbb{I}(t \in T_k) = 0 \leq Y_{2,t}$$

- Case 2: $\exists k$ such that $t \in T_k$. Suppose $k_t = \max\{k : t \in T_k\}$. Then we have the following two facts: (1) for $k \leq k_t$, $t \in T_k$ and for $k > k_t$. $t \notin T_k$; (2) $\exists s \in S_{k_t}$ such that $(s,t) \in R$, and for $k \leq k_t$, $s \in S_k$. Combining these two facts, wel have:

$$
\begin{aligned}
Y_{2,t} \geq Y_{1,s} &= \sum_k \lambda_k Z_{1k,s} = \sum_k \lambda_k \mathbb{I}(s \in S_k) \\
&\geq \sum_{k=1}^{k_t} \lambda_k \mathbb{I}(s \in S_k) = \sum_{k=1}^{k_t} \lambda_k \mathbb{I}(t \in T_k) \\
&= \sum_k \lambda_k \mathbb{I}(t \in T_k) = \sum_k \lambda_k Z_{2k,t} \\
&= Y_{2min,t}
\end{aligned}
$$

So, for any Y_2 satisfies the condition in statement $\mathscr{2}'$, we have:

$$
\sum_{j=1}^n \mu_2(j) Y_{2,j} \geq \sum_{j=1}^n \mu_2(j) Y_{2min,j} \geq \sum_{i=1}^m \mu_1(i) Y_{1,i}.
$$

$\square$

5 Conclusion

In this paper, we defined the notion of quantum coupling and proved a quantum generalisation of Strassen theorem for probabilistic coupling. It is well-known that Strassen theorem is true in both the finite-dimensional and infinite-dimensional cases. However, Theorem 2 (quantum Strassen theorem) was proved only in the finite-dimensional case. So, an open problem is: whether is quantum Strassen theorem still valid in the infinite-dimensional case? Another interesting topic for further study is to use the coupling techniques to study the behaviours of quantum random walks and quantum Markov chains. As pointed out in the Introduction, in the future studies, we hope to apply quantum coupling to develop quantum relational Hoare logic and then use it in formal verification of quantum cryptographic protocols and differential privacy.

6 Acknowledgment

References

[1] G. Barthe, T. Espitau, J. Hsu, T. Sato and P. -Y. Strub, *-liftings for differential privacy, In: *Proceedings of International Colloquium on Automata, Languages and Programming (ICALP)*, 2017, pp. 102:1-12.

[2] G. Barthe, N. Fong, M. Gaboardi, B. Grégoire, J. Hsu and P. -Y. Strub, Advanced probabilistic couplings for differential privacy, In: *Proceedings of ACM SIGSAC Conference on Computer and CommunicationsSecurity (CCS)*, 2016, pp. 55-67.

[3] G. Barthe, M. Gaboardi, B. Grégoire, J. Hsu and P. -Y. Strub, Proving differential privacy via probabilistic couplings, In: *Proceedings of IEEE Symposium on Logic in Computer Science (LICS)*, 2016, pp. 749-758.

[4] G. Barthe, B. Grégoire and S. Zanella-Béguelin, Formal certification of code-based cryptographic proofs, In: *Proceedings of ACM Symposium on Principles of Programming Languages (POPL)*, 2009, pp. 90-101.

[5] E. D'Hondt and P. Panangaden, Quantum weakest preconditions, *Mathematical Structures in Computer Science*, 16(2006)429-451.

[6] J. Hsu, *Probabilistic Coupling for Probabilistic Reasoning*, PhD Thesis, University of Pennsylvania, 2017.

[7] T. Lindvall, *Lectures on the Coupling Method*, Courier Corporation, 2002.

[8] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

[9] V. Strassen, The existence of probability measures with given marginals, *The Annals of Mathematical Statistics* 36(1965)423-439.

[10] M. S. Ying, Floyd-Hoare logic for quantum programs, *ACM Transactions on Programming Languages and Systems*, 33(2011) art no. 19, pp. 1-49.

[11] L. Zhou and M. S. Ying, Differential privacy in quantum computation, In: *Proceedings of IEEE 30th Computer Security Foundations Symposium (CSF)*, 2017, pp. 249-262.

[12] S. Bravyi, Requirements for compatibility between local and multipartite quantum states, *Quant. Inf. Comp.*, **4**, 12-26, 2004.

[13] Jianxin Chen, Zhengfeng Ji, David Kribs, Norbert Lutkenhaus, and Bei Zeng, Symmetric extension of two-qubit states, *Phys. Rev. A*, **87**, 032318, (2014).

[14] Eric A. Carlen, Joel L. Lebowitz and Elliott H. Lieb, On an extension problem for density matrices, *Journal of Mathematical Physics*, **54** , 062103 (2013).

[15] Nengkun Yu, Li Zhou, Shenggang Ying, Mingsheng Ying, A New Class of Criteria for Tripartite Marginal Problem, *arXiv:1803.02673*, (2018).