The Sign Problem in Boson Sampling

Bachelor-Thesis

presented by
Dominik Michels

This present work was submitted to the Institute for Quantum Information at RWTH Aachen University

First examiner: Prof. Dr. Barbara M. Terhal
Second examiner: Prof. Dr. David P. DiVincenzo

Aachen, June 14, 2016

Eidesstattliche Versicherung

______________________________ ______________________________

Name, Vorname Matrikelnummer (freiwillige Angabe)

Ich versichere hiermit an Eides Statt, dass ich die vorliegende Arbeit/Bachelorarbeit/ Masterarbeit* mit dem Titel

____

____

____

selbständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Für den Fall, dass die Arbeit zusätzlich auf einem Datenträger eingereicht wird, erkläre ich, dass die schriftliche und die elektronische Form vollständig übereinstimmen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

______________________________ ______________________________

Ort, Datum Unterschrift

*Nichtzutreffendes bitte streichen

Belehrung:

§ 156 StGB: Falsche Versicherung an Eides Statt

Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

§ 161 StGB: Fahrlässiger Falscheid; fahrlässige falsche Versicherung an Eides Statt

(1) Wenn eine der in den §§ 154 bis 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.

(2) Straflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Abs. 2 und 3 gelten entsprechend.

Die vorstehende Belehrung habe ich zur Kenntnis genommen:

______________________________ ______________________________

Ort, Datum Unterschrift

Contents

1 Notation

1. We denote the probability space by (Ω, Σ, P). Ω is the sample space, Σ the set of possible outcomes and P the probability measure of an element $A \in \Sigma$.

2. Let $X : \Omega \to \chi$ be a random variable and $\mathbb{P}_X$ the probability distribution. The mean value of X is denoted by $\mathbb{E}[X]$. The variance of X is denoted by $\mathbb{V}[X]$.

3. Let Z be a complex random variable then we call $\mathbb{E}[Z^2]$ the the second moment and $\mathbb{E}[|Z|^2]$ the absolute second moment.

4. Let M be a set then the cardinality of M is denoted by $|M|$.

5. Let $A, B \in \mathbb{C}^{n\times n}$ then the Hadamard product between A and B is denoted by $A \circ B$.

6. Let $A \in \mathbb{C}^{n\times n}$ then $|A|$ is defined by $|A|_{ij} = |A_{ij}|$.

7. Let $A \in \mathbb{C}^{n\times n}$ then $\sqrt{A}$ is defined by $\sqrt{A}_{ij} = \sqrt{A_{ij}}$.

8. Let $\boldsymbol{x} \in \mathbb{R}^n$ then $\boldsymbol{x}$ is denoted by a bold letter.

9. $\underline{n} = \{i | i \in \mathbb{N}, i <= n\}$.

10. The symmetric group with n elements is denoted by $\mathcal{S}_n$.

11. The group of unitary matrices with dimension n is denoted by $\mathcal{U}(n)$.

12. The d'Alambert operator is denoted by $\square$.

2 Introduction

Since the emergence of computers in the 1950s it has been an important task to study the computational limits of computers. Early on Alonzo Church and Alan Turing claimed in their famous *Extended Church-Turing Thesis* that all computational problems, that are efficiently solvable by real devices, are efficiently solvable by a probabilistic Turing machine.

With the development of quantum physics this claim has been in danger because the calculation of quantum physical phenomena, which are computable by real devices, seem not to be efficiently computable by a probabilistic Turing machine. Nevertheless, for both science and industry, it would be a major step forward to efficiently simulate and predict quantum physics in order to develop new technologies like superconducters or understand fundamental properties of matter like the substructure of the proton in *Quantum Chromo Dynamics*. This urge let to the idea of a quantum simulator or quantum computer first introduced by Richard Feynman in 1982 [9]. Since quantum computers are still beyond the scope of what is technologically feasible one has to stick with algorithms on classical computers.

Today quantum physics is simulated on classical computers with the help of Quantum Monte Carlo algorithms. These algorithms usually are efficient for systems involving bosons and suffer from an exponential slowdown for systems involving fermions. This exponential slowdown is due to sign fluctuations in fermionic systems and thus called the *Sign Problem*.

In this bachelor thesis we explore a problem called *Boson Sampling* in quantum optics. *Boson Sampling* is explained in detail in section 3. In rough words it can be seen as simulating a linear optical network in quantum optics. From a computational complexity point of view *Boson Sampling* is not expected to be efficiently solvable on a classical computer but known to be efficiently solvable on a quantum computer. We show that *Boson Sampling* shows a *Sign Problem* in section 5. Furthermore *Boson Sampling* is strongly related to the calculation of permanents of complex valued matrices. Calculating permanents of general complex valued matrices is believed to be a computationally hard task. We show in section 5 that the calculation of permanents of general complex valued matrices also shows a *Sign Problem* and thus may be related to fermionic systems.

The goal of this bachelor thesis is to study both the *Sign Problem* in *Boson Sampling* and the *Sign Problem* in the calculation of permanents and explore the connection between the calculation of permanents and fermionic systems.

The bachelor thesis is structured as follows. We start by giving a general introduction into quantum and bosonic optics in section 3. Then we review classical methods to calculate the outcomes of linear bosonic networks in section 4. In section 5 we show that *Boson Sampling* and the calculation of permanents have a *Sign Problem* and in section 6 we explore the connection of the calculation of permanents to fermionic systems and study ideas to reduce the hardness of the *Sign Problem* with the use of noncommuting algebras.

3 Quantum Optics

3.1 General Quantum Optics

In this section we present a short introduction into the quantization of the electromagnetic field. We follow the presentation by Gerry and Knight in [11]. In classical physics light is understood as an excitation of the electromagnetic field. The propagation of light is given by Maxwell's equations. With the vector potential $\boldsymbol{A}$ and the Coulomb gauge condition

$$\nabla \cdot \boldsymbol{A} = 0$$

the vacuum Maxwell's equations are given by

$$\Box \boldsymbol{A} = 0. \tag{1}$$

The electric and magnetic field can be retrieved from the vector potential by

$$\boldsymbol{E}(\boldsymbol{r}, t) = -\frac{\partial \boldsymbol{A}(\boldsymbol{r}, t)}{\partial t}$$

and

$$\boldsymbol{B}(\boldsymbol{r}, t) = \nabla \times \boldsymbol{A}(\boldsymbol{r}, t).$$

Suppose the electromagnetic field is contained in some volume $V \subseteq \mathbb{R}^3$ then the energy H contained in the electromagnetic field is

$$H = \frac{1}{2} \int\limits_V \left(\epsilon_0 \boldsymbol{E} \cdot \boldsymbol{E} + \frac{1}{\mu_0} \boldsymbol{B} \cdot \boldsymbol{B} \right) dV.$$

The solution of Equation 1 can be written as a superposition of plane waves with wavevector $\boldsymbol{k}$ and two different polarizations s.

$$\boldsymbol{A}(\boldsymbol{r}, t) = \sum_{\boldsymbol{k}, s} \boldsymbol{e}_{\boldsymbol{k}, s} \left(A_{\boldsymbol{k}s} e^{i\boldsymbol{k} \cdot \boldsymbol{r}} + A^*_{\boldsymbol{k}s} e^{-i\boldsymbol{k} \cdot \boldsymbol{r}} \right)$$

After some calculation we omit here (see pp. 21-22 [11]) the energy of the electromagnetic field can be rewritten as

$$H = \frac{1}{2} \sum_{\boldsymbol{k}, s} \left(p^2_{\boldsymbol{k}s} + \omega^2_{\boldsymbol{k}} q^2_{\boldsymbol{k}s} \right). \tag{2}$$

This is particularly interesting because every summand in Equation 2 has the same the form as the energy of a simple harmonic oscillator with frequency $\omega_{\boldsymbol{k}}$ and unit mass. Hence we can treat the electromagnetic field as a collection of many independent harmonic oscillators. The variables $p_{\boldsymbol{k}s}$ and $q_{\boldsymbol{k}s}$ are treated as general canonical variables.
Now we proceed with quantizing the electromagnetic field energy. Since the energy is given by the sum of energies of independent harmonic oscillators we quantize the total energy by quantizing the energy of every harmonic oscillator in the sum. In standard canonical quantization the canonical variables become operators satisfying the commutation relations

$$[\hat{q}_{\boldsymbol{k}s}, \hat{q}_{\boldsymbol{k}'s'}] = [\hat{p}_{\boldsymbol{k}s}, \hat{p}_{\boldsymbol{k}'s'}] = 0$$
$$[\hat{q}_{\boldsymbol{k}s}, \hat{p}_{\boldsymbol{k}'s'}] = i\hbar \delta_{\boldsymbol{k}\boldsymbol{k}'} \delta_{ss'}.$$

With the definition of the creation and annihilation operator

$$\hat{a}_{\boldsymbol{k}s} = \frac{1}{\sqrt{2\hbar\omega_{\boldsymbol{k}}}} \left(\omega_{\boldsymbol{k}} \hat{q}_{\boldsymbol{k}s} + i\hat{p}_{\boldsymbol{k}s}\right)$$
$$\hat{a}^{\dagger}_{\boldsymbol{k}s} = \frac{1}{\sqrt{2\hbar\omega_{\boldsymbol{k}}}} \left(\omega_{\boldsymbol{k}} \hat{q}_{\boldsymbol{k}s} - i\hat{p}_{\boldsymbol{k}s}\right)$$

the Hamilton operator $\hat{H}$ can be calculated (see pp. 12-13 [11]) from Equation 2 to

$$\hat{H} = \sum_{\boldsymbol{k}s} \hbar\omega_{\boldsymbol{k}} \left(\hat{a}^{\dagger}_{\boldsymbol{k}s}\hat{a}_{\boldsymbol{k}s} + \frac{1}{2}\right) \tag{3}$$
$$= \sum_{j} \hbar\omega_{j} \left(\hat{n}_{j} + \frac{1}{2}\right) \tag{4}$$

where the indices $\boldsymbol{k}s$ have been relabelled by j. Each j is called a mode in the following.
The possible energies of the electromagnetic field are given by the eigenvalues of the Hamilton operator. They are calculated by summing over all eigenvalues of the independent harmonic oscillators. For one harmonic oscillator the eigenvalues are given by $E_{n_j} = \hbar\omega\left(n_j + \frac{1}{2}\right)$ with the corresponding eigenvector $|n_j\rangle$ and $n_j \in \mathbb{N}$. If the energy eigenvalue in the j-th mode is E_{n_j}, we say that there are n_j photons in the mode j. The eigenvectors of the harmonic oscillator form a basis of the harmonic oscillator Hilbert space $\mathcal{H} = \mathcal{L}^2(\mathbb{C})$. An eigenvector $|n_1, n_2, n_3, \ldots\rangle$ of the Hamilton operator of the electromagnetic field in Equation 3 is then given by the tensor product of the eigenvectors of all independent harmonic oscillators

$$|n_1, n_2, n_3, \ldots\rangle = |n_1\rangle \otimes |n_2\rangle \otimes |n_3\rangle \otimes \ldots.$$

All eigenvectors of the Hamilton operator of the electromagnetic field are part of a larger space called Fock space $\mathcal{F}$. An arbitrary quantized state of the electromagnetic field is given by an element of $\mathcal{F}$. The vacuum where no excitation of the electromagnetic field is present is called $|0\rangle = |0_1, 0_2, 0_3, \ldots\rangle$. The action of the creation and annihilation operators defined above on an arbitrary basis element $|n_1, n_2, n_3, \ldots\rangle$ is given by

$$\hat{a}_j |n_1, n_2, \ldots, n_j, \ldots\rangle = \sqrt{n_j}\, |n_1, n_2, \ldots, n_j - 1, \ldots\rangle$$
$$\hat{a}^{\dagger}_j |n_1, n_2, \ldots, n_j, \ldots\rangle = \sqrt{n_j + 1}\, |n_1, n_2, \ldots, n_j + 1, \ldots\rangle .$$

Hence they can be thought of as creating or annihilating one quantum of energy. The commutation relations are

$$[\hat{a}_i, a^{\dagger}_j] = \delta_{ij} \tag{5}$$
$$[\hat{a}_i, \hat{a}_j] = [\hat{a}^{\dagger}_i, \hat{a}^{\dagger}_j] = 0. \tag{6}$$

We can generate an arbitrary basis state out of the vacuum by

$$|n_1, n_2, \ldots\rangle = \prod_j \frac{\left(\hat{a}^{\dagger}_j\right)^n}{\sqrt{n_j!}} |0\rangle . \tag{7}$$

In the following we omit the hat over an operator, i.e. we write a and $a^{\dagger}$ instead of $\hat{a}$ and $\hat{a}^{\dagger}$.

3.2 Linear Bosonic Optics

In this section we give an introduction into linear bosonic optics and discuss how to compute outcomes of a linear bosonic network. In bosonic optics we restrict ourself to creation and annihilation operators satisfying the commutation relations in Equation 5. One quantum of energy in a mode is similarly to the previous chapter called a boson. Due to the commutativity of the creation and annihilation operators the state of the system does not change sign under the exchange of particles. Obviously our photons from the previous chapter satisfy bosonic commutation relations and therefore are called bosons.

We understand a bosonic system as a computer. We start with an input state $|x\rangle \in \mathcal{F}$. Then some network acts on the input state and produces the output state $|y\rangle \in \mathcal{F}$. We require that no bosons or no energy gets lost in our computation. The number of bosons k stays constant. Furthermore we require that the number of modes in the system is fixed. The number of modes in our system is called n. This implies that the state space can be reduced from the whole Fock space to a subspace with n modes. This space is given by $H^{\otimes n}$ and forms a Hilbert space. The action of the linear network is described by a unitary matrix U, i.e. $|y\rangle = U|x\rangle$.

Let $d = poly(n)$ and the unitary $U = U_d \dots U_1$ be a sequence of unitary operators acting on only two modes. We denote the modes a unitary out of the sequence acts on with i and j.

The network is called passive linear or non-interacting (Terhal, DiVincenzo [20]) if U can be written as $U = \exp(iH_g)$ with the Hamiltonian

$$H_g = b_{ii} a_i^\dagger a_i + b_{jj} a_j^\dagger a_j + b_{ij} a_i^\dagger a_j + b_{ij}^* a_j^\dagger a_j.$$

The coefficients $b_{\alpha\beta}$ form a 2×2 hermitian matrix. We will consider these coefficients to be part of an $n \times n$ matrix b, which is only non-zero for matrix elements involving modes i and j. The action of U onto a state $a_i^\dagger |0\rangle$ containing only one boson can be rewritten as

$$U a_i^\dagger |0\rangle = U a_i^\dagger U^\dagger U |0\rangle = U a_i^\dagger U^\dagger |0\rangle$$

because $U|0\rangle = |0\rangle$ due to boson number conservation. This conjugation of the creation operator of mode i can be thought of as creating bosons in many different modes depending on U. This is formalized in the following theorem.

Theorem 1 *There exists a matrix $V \in \mathbb{C}^{2\times 2}$ with $V = \exp(ib)$ such that*

$$U a_i^\dagger U^\dagger = \sum_{m=1}^{2} V_{im} a_m^\dagger$$

Proof We will consider b to be the 2×2 hermitian matrix in this proof. Without loss of generality we relabel modes i, j with 1 and 2. The spectral theorem tells us that we can find a matrix $W \in \mathcal{U}(2)$ such that $W^\dagger b W = \text{diag}(\lambda_1, \lambda_2) =: D$ where λ_1, λ_2 are eigenvalues of b. This implies

$$\begin{aligned} U = \exp(iH) &= \exp\Big(i \begin{pmatrix} a_1^\dagger a_2^\dagger \end{pmatrix} W D W^\dagger \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}\Big) \\ &= \exp\Big(i \boldsymbol{c}^\dagger D \boldsymbol{c}\Big) \end{aligned}$$

with $\boldsymbol{c}^\dagger = \begin{pmatrix} a_1^\dagger a_2^\dagger \end{pmatrix} W$ and $\boldsymbol{c} = W^\dagger \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$. We calculate

$$\begin{aligned} U a_i^\dagger U^\dagger &= \exp\Big(c_1^\dagger c_1 \lambda_1 + c_2^\dagger c_2 \lambda_2\Big) \left(\sum_{j=1}^{2} W_{ij}^\dagger c_i^\dagger\right) \exp\Big(-i(c_1^\dagger c_1 \lambda_1 + c_2^\dagger c_2 \lambda_2)\Big) \\ &= \sum_{j=1}^{2} \Big(\exp\Big(i c_j^\dagger c_j \lambda_j\Big) W_{ij}^\dagger c_j^\dagger \exp\Big(-i c_j^\dagger c_j \lambda_j\Big)\Big) \end{aligned}$$

We look further at

$$\begin{aligned} \exp\Big(i c_j^\dagger c_j \lambda_j\Big) c_j^\dagger &= \sum_{k=0}^{\infty} \frac{1}{k!} \Big(i c_j^\dagger c_j \lambda_j\Big)^k c_j^\dagger \\ &= c_j^\dagger \sum_{k=0}^{\infty} \frac{1}{k!} \Big(i c_j c_j^\dagger \lambda_j\Big)^k \\ &= c_j^\dagger \exp\Big(i c_j c_j^\dagger \lambda_j\Big) \\ &= c_j^\dagger \exp\Big(i c_j^\dagger c_j \lambda_j + i \mathbb{1} \lambda_j\Big) \end{aligned}$$

where the commutation relations of the bosonic operators has been used in the last equality. With that we get

$$\begin{aligned} U a_i^\dagger U^\dagger &= \sum_{j=1}^{2} W_{ij}^\dagger c_j^\dagger \exp(i \mathbb{1} \lambda_j) \\ &= \sum_{m=1}^{2} \sum_{j=1}^{2} W_{ij}^\dagger \exp(i\lambda_j) W_{jm} a_m^\dagger \\ &= \sum_{m=1}^{2} \left(\exp(ih)\right)_{im} a_m^\dagger \end{aligned}$$

where $(\exp(ih))_{im} = \big(\exp\big(iW^\dagger D W\big)\big)_{im} = (W^\dagger \exp(iD) W)_{im} = (W^\dagger \operatorname{diag}(\exp(i\lambda_1), \exp(i\lambda_2))\, W)_{im} = \Big(\sum_{j=1}^{2} W_{ij}^\dagger \exp(i\lambda_j) W_{jm}\Big)_{im}$ and the unitarity of W has been used in the last equality. □

Again we consider the V_i to be part of a larger $n \times n$ matrix which is only non-zero for matrix elements involving modes i and j. The V for the entire network $U = U_k \ldots U_1$ is retrieved by matrix multiplication

$$U a_i^\dagger U^\dagger = U_d \ldots U_1 a_i^\dagger U_1^\dagger \ldots U_d^\dagger = \sum_{m=1}^{n} (V_1 \cdot \ldots \cdot V_d)_{im} a_m^\dagger \tag{8}$$

We are interested in the probability of the outcomes of the linear optical network. Therefore we show how to evaluate the matrix element $\langle y|U|x\rangle$ for arbitrary basis states $|x\rangle, |y\rangle \in \mathcal{H}^{\otimes n}$. We start by giving the definition of the occupation number function.

Definition 1 *Let $|x\rangle$ be a basis state of $\mathcal{H}^{\otimes n}$. The function $\#_x : \mathbb{N} \to \mathbb{N}_0$ returns the number of bosons in each mode. This number is also called occupation number.*

We label the modes where a boson is present with the indices $i_1, \ldots, i_k$. An index may appear multiple times because many bosons can occupy the same mode. With the help of Equation 7, valid for general bosonic systems, we write

$$U |x\rangle = U \prod_{j=1}^{k} \frac{a_{i_j}^\dagger}{\sqrt{\#_x(i_j)!}} |0\rangle .$$

With the use of Equation 8 we write

$$U |x\rangle = \prod_{j=1}^{k} \left(\frac{1}{\sqrt{\#_x(i_j)!}} \right) \sum_{p_1,\ldots,p_k}^{k} V_{i_1 p_1} V_{i_2 p_2} \ldots V_{i_k p_k} a_{p_1}^\dagger a_{p_2}^\dagger \ldots a_{p_k}^\dagger |0\rangle$$

similarly to the input state we write the output state in the form $\langle y| = \langle 0| \prod_{j=1}^{n} \frac{a_{l_j}}{\sqrt{\#_y(l_j)!}}$ where $l_1, \ldots, l_k$ label the indices where a boson is present in the output state. We write

$$\langle y|U|x\rangle = \prod_{j=1}^{k} \left(\frac{1}{\sqrt{\#_x(i_j)!\,\#_y(l_j)!}} \right) \sum_{p_1,\ldots,p_k}^{k} V_{i_1 p_1} V_{i_2 p_2} \ldots V_{i_k p_k} \langle 0| a_{l_1} a_{l_2} \ldots a_{l_k} a_{p_1}^\dagger a_{p_2}^\dagger \ldots a_{p_k}^\dagger |0\rangle .$$

Contributions in the sum only arise when $l_1, \ldots, l_k$ is a permutation of $p_1, \ldots, p_k$. To see that, assume $l_1, \ldots, l_k$ is not a permutation of $p_1, \ldots, p_k$. Then there exists a mode $j \in \mathbb{N}$ such that the number of bosons in the mode j of the input mode is different to the output mode. All basis states of the Hilbert space are orthogonal. Hence then $\langle 0| a_{l_1} a_{l_2} \ldots a_{l_k} a_{p_1}^\dagger a_{p_2}^\dagger \ldots a_{p_k}^\dagger |0\rangle = 0$ With the use of the commutation relation of the creation and annihilation operators we get

$$\begin{aligned} \langle y|U|x\rangle &= \prod_{j=1}^{k} \left(\frac{1}{\sqrt{\#_x(i_j)!\,\#_y(l_j)!}} \right) \sum_{\pi \in S_k} V_{i_1 \pi(l_1)} V_{i_2 \pi(l_2)} \ldots V_{i_k \pi(l_k)} \prod_{j=1}^{k} \left(\sqrt{\#_y(l_j)!} \right) \\ &= \prod_{j=1}^{k} \left(\frac{1}{\sqrt{\#_x(i_j)!}} \right) \sum_{\pi \in S_k} V_{i_1 \pi(l_1)} V_{i_2 \pi(l_2)} \ldots V_{i_k \pi(l_k)}. \end{aligned}$$

We define $\widetilde{V}$ as the matrix where we have selected the rows $i_1, \ldots, i_k$ and columns $l_1, \ldots, l_k$ from V. Note that rows and columns in $\widetilde{V}$ may appear multiple times due to the multiple selection of modes in $i_1, \ldots, i_k$ and $l_1, \ldots, l_k$. With that definition we reduce the calculation of the probability amplitude $\langle y|U|x\rangle$ to the calculation of the permanent of $\widetilde{V}$.

$$\langle y|U|x\rangle = \prod_{j=1}^{k} \left(\frac{1}{\sqrt{\#_x(i_j)!}} \right) \operatorname{Perm} \widetilde{V}$$

3.2.1 Computational complexity

In the previous section we reduced the calculation of the probability amplitude of a linear bosonic network to the calculation of the permanent of a $k \times k$ complex valued matrix, where k is the number of bosons in the network. The calculation of the permanent of a matrix is believed to be a computationally hard task. Vailant [22] showed that the computation of the permanent of a general $k \times k$ 0,1 valued matrix is $\#P$-complete. A problem is $\#P$-complete if it can be reduced to counting accepting paths of a non-deterministic Turing machine in polynomial time. A problem is in NP if it can be solved by a non-deterministic Turing machine in polynomial time. Thus $\#P$ problems can be thought of as counting the number of solutions of NP problems. Besides $\#P$ and NP there is P containing all the problems solvable on a Turing machine in polynomial time and BQP containing all the problems solvable on a quantum computer in polynomial time. It is known that $P \subset BQP$ and $P \subset NP$. It is generally believed that the solution of $\#P$-complete problems is not possible in polynomial time on a classical computer; or more formally it is believed that $\#P \neq P$. Hence we cannot hope to efficiently compute the probability amplitude of a linear bosonic network on a classical computer. Furthermore even the approximate calculation of the permanent of a general matrix $A \in \mathbb{C}^{n \times n}$ within a constant factor is $\#P$-complete [1]. Nevertheless for specific classes of matrices it is possible to approximate the value of the permanent in polynomial time. Gurvits constructed an algorithm that approximates $\operatorname{Perm} A$ within an additive error $\pm\epsilon||A||$ in $O(n^2/\epsilon^2)$ time [2]. The matrix norm of the submatrices of V out of section 3.2 used in quantum optics is smaller than or equal to 1. Thus Gurvit's algorithm approximates the permanent within an additive error of ϵ. We are going to study approximate methods to compute the permanent in more detail later.
The problem known as *Boson Sampling*, called BS in the following, is drawing samples $|y\rangle$ from the probability distribution given by $|\langle y|U|x\rangle|^2$. In [1] Aaronson and Arkhipov claim that the *Boson Sampling* problem is computationally hard on a classical computer. In the following we are concerned with calculating the probability amplitude and more generally calculating permanents of general complex valued matrices. This problem is not the same as drawing samples from the distribution and such not the same as *Boson Sampling*, but strongly related to it. As stated above this problem is $\#P$-complete.
It is known that a quantum computer can efficiently perform *Boson Sampling*, hence $BS \subset BQP$. But BS itself may not be enough to perform universal quantum computation (Knill, Laflamme, Milburn [8]). It is expected that BS is in between classical and universal quantum computation and thus of theoretical interest (Aaronson, Arkhipov [1]). The relation between the complexity classes are shown graphically in Figure 1.

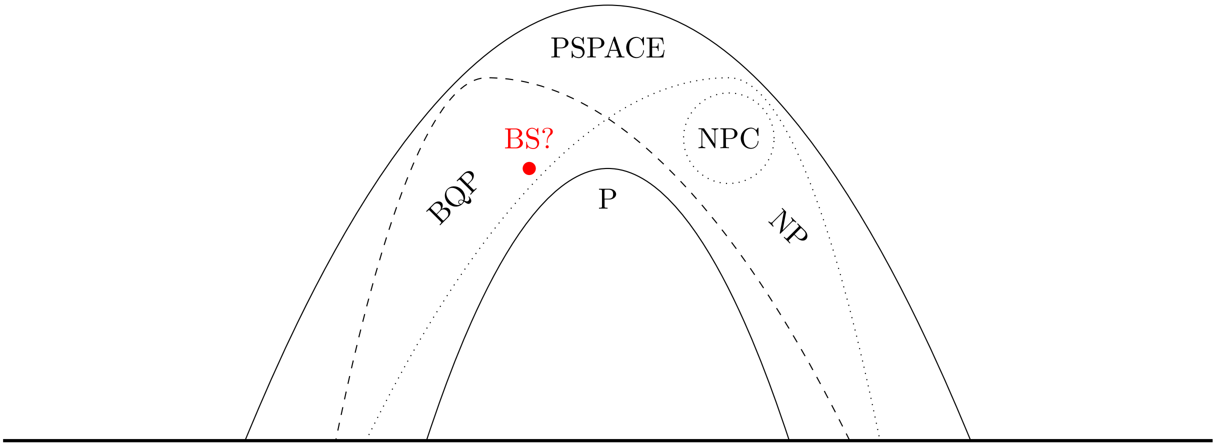

Figure 1: Relation between the complexity classes. Proven relations are drawn in solid and believed relations in dashed lines. BS is supposed to be contained in BQP but not in P. Graphic inspired by: `http://www.texample.net/tikz/examples/complexity-classes/`

3.3 Fermionic linear optics

In this section we give a short introduction into fermionic linear optics. The fermionic situation is similar to the bosonic one. The only difference is that we choose the creation and annihilation operators to be anti commuting.

$$\{a_i, a_j^\dagger\} = \delta_{ij} \tag{9}$$

$$\{a_i, a_j\} = \{a_i^\dagger, a_j^\dagger\} = 0 \tag{10}$$

This anti commutation enforces the Pauli principle because the anti commutativity implies that $\{a_i^\dagger, a_i^\dagger\} = 2a_i^\dagger a_i^\dagger = 0$ and thus the creation of multiple fermions in one mode is not possible. The derivation of the expression for the matrix element $\langle y|U|x\rangle$ for arbitrary fermionic input and output basis states is analogous to the bosonic case. Similar to Theorem 1 one obtains a transformation V representing the linear fermionic network. The complete derivation can be found in [20]. The result is

$$\langle y|U|x\rangle = \det \widetilde{V}$$

where again $\widetilde{V}$ is obtained by taking the rows and columns from V where a particle is present in the input and respectively in the output state. Multiple selections of rows is not possible because one mode can only be occupied by one fermion.

3.3.1 Computational complexity

Hence for fermionic optics the calculation of probability amplitudes reduces to the calculation of determinants. In contrast to permanents determinants can be calculated efficiently on a classical computer (Terhal, DiVincenzo [20]).

4 Approximate methods to calculate permanents

In this section we analyse several approximate methods to compute the permanent of a complex valued matrix. We start by looking at a connection between the permanent and determinant first discovered by Godsil and Gutman [12].

4.1 The generalized Godsil-Gutman estimator

Before we start introducing the Godsil-Gutman and related estimators we would like to introduce some notation. In the following we need a matrix filled with random variables. We call this matrix X and the random variables occupying the sites of the matrix X_{ij}. X can be seen as a random variable $X : \Omega \to \mathbb{R}^{n\times n}$ itself.

Theorem 2 *Let $A \in \mathbb{C}^{n\times n}$ and X a $n \times n$ Matrix of random variables with $\mathbb{E}[X_{ij}] = 0, \mathbb{V}[X_{ij}] = 1\ \forall i, j \in \underline{n} \times \underline{n}$ and $B = \sqrt{A} \circ X$ then*

$$\operatorname{Perm} A = \mathbb{E}[\det(B(X))^2] \tag{11}$$

PROOF

$$\begin{aligned}
\mathbb{E}[\det(B(X))^2] &= \mathbb{E}\left[\sum_{\pi\in\mathcal{S}_n}\sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma)\prod_{i=1}^n B_{i\pi(i)}B_{i\sigma(i)}\right] \\
&= \sum_{\pi\in\mathcal{S}_n}\sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma)\prod_{i=1}^n \sqrt{A_{i\pi(i)}A_{i\sigma(i)}}\,\mathbb{E}\left[X_{i\pi(i)}X_{i\sigma(i)}\right]] \\
&= \sum_{\pi\in\mathcal{S}_n}\prod_{i=1}^n A_{i\pi(i)} \\
&= \operatorname{Perm} A
\end{aligned}$$

where the linearity of the expectation value has been used in line 2 and the mean and variance condition in line 3. □

This estimator of the permanent of a matrix is an interesting connection between the permanent and the determinant because, as pointed out above, the permanent is believed to be hard and the determinant known to be easy to compute. The estimator is called Godsil-Gutman estimator if the random variables X_{ij} are chosen to be i.i.d. Bernoulli random variables. If the X_{ij} are chosen to be i.i.d. Gaussian random variables, the estimator is called Barvinok estimator [3]. We call the estimator in Theorem 2 the generalized Godsil-Gutman estimator.
Regardless of the explicit choice of the random variable it is possible to approximate the mean value by drawing and averaging samples of $Y(X) = \det(B(X))^2$. More formally we define the random variable

$$\overline{Y} = \frac{Y_1 + \cdots + Y_N}{N}$$

where $Y_1, \ldots, Y_N$ are N independent copies of Y. The mean value of $\overline{Y}$ is equal to the mean value of Y and thus equal to the permanent of the matrix. The variance of $\overline{Y}$ and so the variance of the averaging process will be large such that an efficient approximation of the permanent is not possible. In the following we are concerned with calculating the corresponding error of the approximation. We say that that the relative error of the approximation is ϵ, if

$\mathbb{P}(|Y - \mathbb{E}[Y]| \geq \epsilon E[Y]) < \delta$. In other words this means that the estimated value of the permanent is with probability $1-\delta$ within a factor of $1+\epsilon$ or $1-\epsilon$ equal to the permanent. We say that that the additive error of the approximation is ϵ, if $\mathbb{P}(|Y - \mathbb{E}[Y]| \geq \epsilon) < \delta$. We use Chebyshev's inequality:

Chebyshev inequality [10] *Let X be an arbitrary random variable and $c > 0$. Then it holds that*

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq c) \leq \frac{\mathbb{V}[X]}{c^2}.$$

We are interested in the deviation of $\overline{Y}$ from its mean. We bound the deviation with the help of Chebyshev's inequality.

$$\begin{aligned}
\mathbb{P}(|\overline{Y} - \mathbb{E}[\overline{Y}]| \geq c) &\leq \frac{\mathbb{V}[\overline{Y}]}{c^2} && \text{(Chebyshev inequality)} \\
&= \frac{1}{c^2}\left(\mathbb{E}\left[|\overline{Y} - \mathbb{E}[\overline{Y}]|^2\right]\right) \\
&= \frac{1}{c^2}\left(\mathbb{E}\left[|\overline{Y}|^2\right] - \left|\mathbb{E}[\overline{Y}]\right|^2\right) \\
&= \frac{1}{N^2c^2}\left(\mathbb{E}\left[\left|\sum_{i=1}^{N} Y_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{N} Y_i Y_j\right|\right] - N^2|\mathbb{E}[Y]|^2\right) \\
&\leq \frac{1}{N^2c^2}\left(N\mathbb{E}[|Y|^2] - N^2\,|\mathbb{E}[Y]|^2\right) && \text{(Covariances vanish)} \\
&\leq \frac{\mathbb{V}[Y]}{Nc^2} && (12)
\end{aligned}$$

If we set $c = \epsilon \cdot \mathbb{E}[\overline{Y}]$ with $\epsilon \in \mathbb{R}$ we get a bound for the relative error ϵ. Intuitively speaking this means that by drawing enough samples from $\det(B(X))^2$ and averaging them we approximate the permanent of A with the help of Equation 11 within an relative error ϵ of

$$\epsilon = \frac{1}{\delta\sqrt{N}} \frac{\sqrt{\mathbb{V}[\det(B(X))^2]}}{\mathbb{E}\left[\det(B(X))^2\right]}. \qquad (13)$$

Hence the error of the approximation decreases with N and is dependent on $\mathbb{V}[\det(B(X))^2]$. This means that for large enough N we get a good approximation of Perm A with high probability. How large N has to be is determined by $\mathbb{V}[\det(B(X))^2]$. We analyse the variance in the next sections in detail.

4.1.1 Error estimation without importance sampling

In the following we are concerned with analysing the variance of the Godsil-Gutman estimator in order to estimate its power. The Godsil-Gutman estimator has been proven to have an exponentially large error for general complex $n \times n$ matrices [15]. This is not surprising because calculating the permanent of a complex matrix is $\#P$-hard and even an efficient approximate solution of a general $\#P$-hard problem is not known today.
What turns out to be working well is the approximation of the permanent of matrices with nonnegative entries, i.e. $A \in \mathbb{R}_+^{n\times n}$. Although the approximation of nonnegative matrices is not

related to the *Boson Sampling* problem, approximating permanents is an interesting problem by itself. In 2004 Jerrum, Sinclair and Vigoda constructed a fully-polynomial randomized approximation scheme or short FPRAS for calculating the permanent of a $n \times n$ matrix with nonnegative entries [16]. The approximation in their algorithm is, with high probability, within an arbitrarily small specified relative error of the true value of the permanent. Their algorithm runs in $O(n^{10})$ operations.
We estimated the variance of the Godsil-Gutman estimator for nonnegative matrices numerically. For every dimension from $n = 2$ to $n = 14$ we chose 100 different nonnegative matrices A randomly from $(0, 100]^{n \times n}$ and evaluated Y 10000 times. The interval $(0, 100]$ and no larger interval was chosen because otherwise the value of permanents became to large for the computer. For every A we evaluated the variance and the mean of Y out of the 10000 samples. In order to evaluate the power of the estimator we plotted the relative standard deviation $\frac{\sqrt{\mathbb{V}[Y]}}{\mathbb{E}[Y]}$ against the dimension of the matrices. The result is displayed in Figure 2. The numerical data shows that the relative standard deviation of Y can be bounded by the dimension n of the matrix. An even lower bound of $\sqrt{n}$ could be possible. More data is needed to confirm or disprove that. With a standard deviation bounded by n the error bound in Equation 13 implies that the random variable $\overline{Y}$, retrieved by averaging N copies of Y, has a relative error of $\frac{n}{\delta\sqrt{N}}$. The calculation of the determinant of a matrix with Gaussian elimination takes $O(n^3)$ steps. Therefore we propose that the approximation of the permanent of the matrix A, retrieved with the Godsil-Gutman estimator, is, with high probability, within an arbitrarily small specified relative error of the true value of the permanent after $O(n^5)$ operations. If that claim holds the Godsil-Gutman estimator performs better than the algorithm by Jerrum, Sinclair and Vigoda for nonnegative matrices.

4.1.2 Error estimation with importance sampling

In this section we apply the technique of importance sampling to the Godsil-Gutman estimator. We prove that the relation in Equation 11 cannot be evaluated efficiently for arbitrary matrices with a standard Monte Carlo algorithm. This will be important in 5 where we classify the *Sign Problem* in *Boson Sampling*. We start by giving a short introduction into Monte Carlo algorithms and importance sampling.
Consider the discrete random variable $X : \Omega \to \chi$ with probability density function $P_X(\omega)$. Then the expected value of a function $g(X)$ is

$$\mathbb{E}[g(X)] = \sum_{\omega \in \chi} g(\omega) P_X(\omega).$$

By taking N independent samples of X, i.e. $x = (x_1, \ldots, x_N)$, we can evaluate the mean of g by evaluating the so called Monte Carlo estimate g_N for $\mathbb{E}[g(X)]$.

$$g_N(x) = \frac{1}{N} \sum_{i=1}^{N} g(x_i). \tag{14}$$

For large N we expect that the Monte Carlo estimate is close to $\mathbb{E}[g(X)]$ with high probability. The convergence of the Monte Carlo estimate to $\mathbb{E}[g(X)]$ and its variance is analogues to the previous section given by Chebychev's inequality. We get for the variance

$$\mathbb{V}[g_N(X)] = \frac{\mathbb{V}[g(X)]}{N}.$$

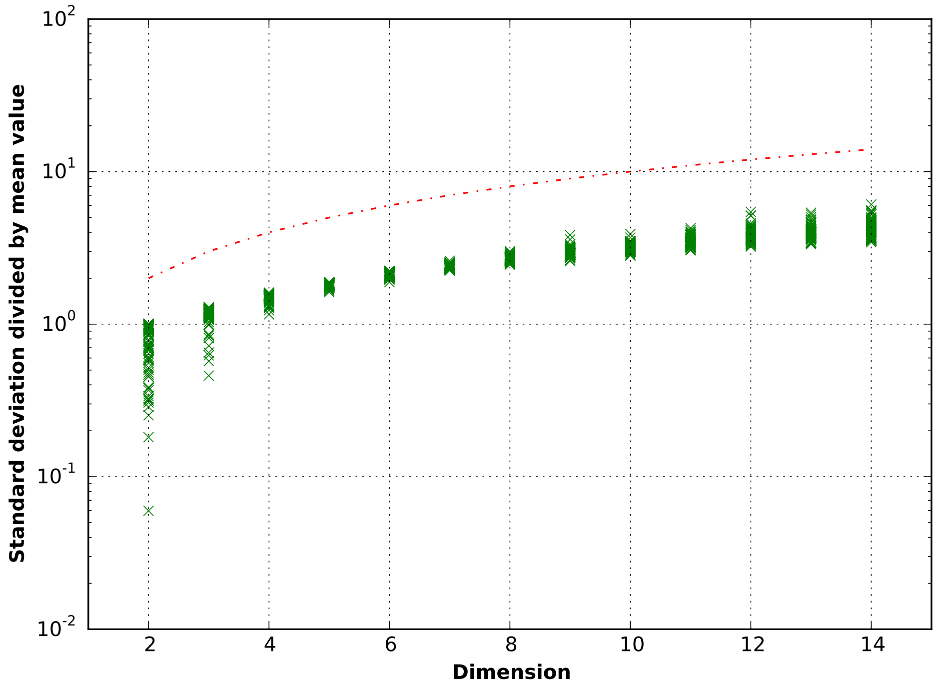

Figure 2: Results on the variance estimation of the Godsil-Gutman estimator. For every dimension we calculated the distribution of Y for 100 matrices with 10000 samples. The standard deviation divided by the mean of Y is plotted against the dimension of the matrix. The curve $n \mapsto n$ is plotted in red dashed lines for comparison.

Our goal, of course, is to make the variance of the Monte Carlo estimate as low as possible such that the convergence of the Monte Carlo estimate to $\mathbb{E}[g(X)]$ is fast. In order to achieve faster convergence we use the technique of *importance sampling.* Consider that we choose the samples $x = (x_1, \ldots, x_n)$ in Equation 14 with high probability at a position where the value of $g(\omega)$ is large and with low probability at a position where the value of $g(\omega)$ is small. Then we would expect that the Monte Carlo estimate converges faster to $\mathbb{E}[g(X)]$ and thus the variance of the Monte Carlo estimate decreases. In Equation 14 the x_i are drawn independently according to the probability density function $P_X(\omega)$. We would like to exchange the probability density function $P_X(\omega)$ without changing the mean value. The technique of changing the probability density function is called *importance sampling.* We name the new probability density function $h_X(\omega)$ and require that $h_X(\omega) \neq 0 \; \forall \omega \in \chi$. The exchange of the probability density function works as follows

$$\begin{aligned}\mathbb{E}[g(X)] &= \sum_{\omega \in \chi} g(\omega) P_X(\omega) \\ &= \sum_{\omega \in \chi} h_X(\omega) \frac{g(\omega) P_X(\omega)}{h_X(\omega)} \\ &= \mathbb{E}_h \left[\frac{g(X) \cdot P_X(X)}{h_X(X)} \right].\end{aligned}$$

Intuitively we would choose the new probability density function h_X to be proportional to $|g|$ because then naturally the probability of choosing the element ω is high if the value of $g(\omega)$ is large. It can be shown [19] that this intuition is valid and the variance of the new estimator E_h is minimal if $h(\omega) \propto |g(\omega)|$. Now we return from the general case to the permanent estimator and we choose h to be proportional to $|Y(\omega)| = |(\det B(\omega))^2|$, i.e.

$$h_X(\omega) = \frac{|Y(\omega)|}{\sum\limits_{\omega \in \chi} |Y(\omega)|}. \tag{15}$$

With the definition $Z : \chi \to \mathbb{C}, \omega \to \frac{Y(\omega) P_X(\omega)}{h_X(\omega)}$ we get

$$\operatorname{Perm} A = \mathbb{E}_h \left[\frac{Y(X) P_X(X)}{h_X(X)} \right] = \mathbb{E}_h[Z(X)] \tag{16}$$

This could be seen a slightly modified version of the generalized Godsil-Gutman estimator. The interesting question is how the variance has decreased. In order to calculate the variance we need to prove a help theorem first.

Lemma 1 *Under the same conditions as Theorem 2 it holds that*

$$\mathbb{E}[|\det(B(X))|^2] = \operatorname{Perm}|A|$$

Proof By calculation

$$
\begin{aligned}
\mathbb{E}[|Y(X)|] &= \mathbb{E}\left[(\det B(X))(\det B(X))^*\right] \\
&= \sum_{\sigma\in\mathcal{S}_n}\sum_{\pi\in\mathcal{S}_n} \operatorname{sgn}\sigma \operatorname{sgn}\pi \ \mathbb{E}\left[\prod_{i=1}^{n} X_{i\sigma(i)}X_{i\pi(i)}\sqrt{A_{i\sigma(i)}}\sqrt{A^*_{i\pi(i))}}\right] \\
&= \sum_{\sigma\in\mathcal{S}_n}\sum_{\pi\in\mathcal{S}_n} \operatorname{sgn}\sigma \operatorname{sgn}\pi \prod_{i=1}^{n} \mathbb{E}\left[X_{i\sigma(i)}X_{i\pi(i)}\right]\sqrt{A_{i\sigma(i)}}\sqrt{A^*_{i\pi(i))}} \\
&= \sum_{\sigma\in\mathcal{S}_n}\prod_{i=1}^{n}|A_{i\sigma(i)}| \\
&= \operatorname{Perm}|A|
\end{aligned}
$$

□

For the rest of the section we require that the X_{ij} are Bernoulli random variables. This implies that the probability density function of the random variable X is given by $P_X(\omega) = (\frac{1}{2})^{n^2}$. We calculate the variance of the complex valued random variable $Z(X)$.

$$
\begin{aligned}
\mathbb{V}[Z] &= \mathbb{E}\left[|Z - \mathbb{E}[Z]|^2\right] \\
&= \mathbb{E}\left[|Z|^2\right] + |\mathbb{E}\left[Z\right]|^2
\end{aligned}
$$

Of interest for the relative variance is the absolute second moment $\mathbb{E}\left[|Z^2|\right]$ of Z. We calculate the absolute second moment to

$$
\begin{aligned}
\mathbb{E}_h[|Z(X)|^2] &= \mathbb{E}_h\left[\left|\frac{Y(X)P_X(X)}{h_X(X)}\right|^2\right] \\
&= \mathbb{E}_h\left[\left(\frac{|Y(X)|P_X(X)}{|Y(X)|}\right)^2\left(\sum_{\omega\in\chi}|Y(\omega)|\right)^2\right] \\
&= \sum_{\omega\in\chi} h(\omega)\left(\left(\frac{1}{2}\right)^{n^2}\right)^2\left(\sum_{\omega\in\chi}|Y(\omega)|\right)^2 \\
&= \sum_{\omega\in\chi}\frac{|Y(\omega)|}{\sum_{\omega\in\chi}|Y(\omega)|}\left(\left(\frac{1}{2}\right)^{n^2}\right)^2\left(\sum_{\omega\in\chi}|Y(\omega)|\right)^2 \\
&= \sum_{\omega\in\chi}\left(\left(\frac{1}{2}\right)^{n^2}|Y(\omega)|\right)\sum_{\omega\in\chi}\left(\left(\frac{1}{2}\right)^{n^2}|Y(\omega)|\right) \\
&= \mathbb{E}[|Y|]^2 \\
&= (\operatorname{Perm}|A|)^2
\end{aligned} \tag{17}
$$

where Lemma 1 has been used in the last row. The permanent of the absolute matrix can be exponentially larger then the permanent of the matrix itself. To see that assume without loss of generality that n is even and consider the block diagonal matrix $D = \operatorname{diag}(C, \ldots, C)$ with $n/2$ blocks where each block C is given by

$$
C = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.
$$

The permanent of D is zero but the permanent of the absolute matrix $|C|$ is $2^{\frac{n}{2}}$. Hence even with *importance sampling* the permanent of a general complex $n \times n$ matrix cannot be evaluated efficiently with the Godsil-Gutman estimator.

4.2 Approximation theorems

As a conclusion of this section we formulate two approximation theorems for the permanent. The first one is general but has a large additive error. The second one is only for nonnegative matrices but provides a constant relative error bound. Both theorems require importance sampling.

Theorem 3 *Suppose we can sample from h_X defined in Equation 15 in $O(f(n))$ time. There exists a randomized classical algorithm that takes a matrix $A \in \mathbb{C}^{n\times n}$ as input runs in $O(\frac{f(n)n^3}{\epsilon^2})$ time and with probability $1-\delta$ approximates* $\operatorname{Perm} A$ *to within an additive error* $\pm\frac{\epsilon}{\delta}\operatorname{Perm}|A|$.

PROOF Choose $N \in \mathbb{N}, N = \frac{1}{\epsilon^2}$ and use the estimator in Equation 16. Calculating the determinant takes $O(n^3)$, generating the samples takes $O(f(n))$ and averaging N trials $O\left(\frac{1}{\epsilon^2}\right)$ time. The variance is with the help of Equation 17 given by

$$\mathbb{V}[Z(X)] = (\operatorname{Perm}|A|)^2 - |\operatorname{Perm} A|^2 \leq (\operatorname{Perm}|A|)^2.$$

By applying Equation 12 we get

$$\mathbb{P}\left(|\overline{Z} - \operatorname{Perm} A| \geq c\right) \leq \frac{(\operatorname{Perm}|A|)^2}{Nc^2}$$

and so an additive error bound of $c = \frac{\epsilon}{\delta}\operatorname{Perm}|A|$. □

Theorem 4 *Suppose we can sample from h_X defined in Equation 15 in $O(f(n))$ time. There exists a randomized classical algorithm that takes a matrix $A \in \mathbb{R}_+^{n\times n}$ as input runs in $O(\frac{f(n)n^3}{\epsilon^2})$ time and with probability $1-\delta$ approximates* $\operatorname{Perm} A$ *to within a relative error $\frac{\epsilon}{\delta}$.*

PROOF Choose $N \in \mathbb{N}, N = \frac{1}{\epsilon^2}$. Calculating the determinant takes $O(n^3)$, generating the samples takes $O(f(n))$ and averaging N trials $O\left(\frac{1}{\epsilon^2}\right)$ time. If all entries of A are positive $\operatorname{Perm} A = \operatorname{Perm}|A|$. Analogues to the proof of the previous theorem we get

$$\mathbb{P}\left(|\overline{Z} - \operatorname{Perm} A| \geq c\operatorname{Perm} A\right) \leq \frac{(\operatorname{Perm}|A|)^2}{Nc^2(\operatorname{Perm} A)^2} = \frac{1}{Nc^2}$$

and so a relative error bound of $c = \frac{\epsilon}{\delta}$. □

5 The numerical sign problem

In this section we give an introduction into the *Sign Problem* and show that the Godsil-Gutman estimator has a sign problem. Furthermore we define the *Sign Problem* in *Boson Sampling*. We write the permanent of the real Matrix $A \in \mathbb{R}^{n \times n}$ according to 4.1 as

$$\operatorname{Perm} A = \int\limits_{\chi} P(\omega) \left(\det B\right)^2 d\omega.$$

Note that if the probability distribution function is not continuous we can still write the expectation value as an integral by using delta distributions. The permanent of a real valued matrix has to be real. With the definitions $O(\omega) = \Re \left(\det B\right)^2$ we write the permanent as

$$\operatorname{Perm} A = \int\limits_{\chi} P(\omega) \operatorname{sign}\left(O(\omega)\right) |O(\omega)| d\omega.$$

Hence the permanent of the matrix A relates to averaging the observable O over χ. As we have already discussed in 4.1.1 the average can be efficiently computed if the sign vanishes. If the sign does not vanish the average can in general not be efficiently computed by standard Monte Carlo algorithms as shown in 4.1.2. Therefore the computational complexity in approximating the permanent of complex valued matrices is solely due to the negativity of the observable. These kind of problems are referred to as problems showing a *Sign Problem* [18] [21].
We continue by defining the *Sign Problem* in *Boson Sampling*. We consider arbitrary input basis states $|x\rangle$ and output states $|y\rangle$ from the Fock space. We write the transition amplitude $\mathbb{P}(y|x)$ according to 3.2 as

$$\mathbb{P}(y|x) = \prod_{j=1}^{k} \left(\frac{1}{\#_x(i_j)!} \right) |\operatorname{Perm} \widetilde{V}|^2.$$

For a moment consider that $\widetilde{V}$ is nonnegative. Then we can calculate $\mathbb{P}(y|x)$ with the algorihtm by Jerrum, Sinclair and Vigoda [16]. With the probability distribution function we can set up a Markov chain which produces samples. Unfortunately submatrices of unitary matrices and thus $\widetilde{V}$ are in general not nonnegative. This is what we call the *Sign Problem* in *Boson Sampling* and what prevents us from setting up a Markov chain to produce samples. It is however not clear if the complexity is due to the negativity of the matrix. It is left to show that the Markov chain converges after a polynomial number of steps to its stationary distribution.

6 Operator methods to calculate permanents

In this section we try to reduce the hardness of the *Sign Problem* in calculating the permanent of a complex valued matrix by formulating a new unbiased estimator which uses random variables with non commuting algebras as image space. This idea is inspired by the work of Chien, Rasmussen and Sinclair [6]. With the help of general Clifford Algebras of with more than three generators they reduced the variance of the generalized Godsil-Gutman estimator to a constant value. Of course this does not mean that the permanent of a matrix can be efficiently approximated. Elements of a Clifford algebra may anticommute. The determinant of a matrix with anticommuting elements can in general not be computed efficiently. It has been proven that calculating the determinant of a matrix over a non commutative algebra is $\#P$-hard [4]. Hence the complexity is shifted from the large variance to the computation of the determinants. But still this is interesting because Clifford Algebras relate to physical fermionic systems. So maybe one could set up an experiment with fermions to approximate the permanent of a matrix. We present techniques involving creation and annihilation operators as well as Majorana fermion operators.

6.1 The Jordan-Wigner-Transformation

We start by introducing the *Jordan-Wigner-Transformation* which maps $2n$ fermionic creation and annihilation operators to Pauli operators X, Y and Z acting on n qubits. The three Pauli operators are defined by

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Besides the creation and annihilation operators we define the Majorana fermion operators c_{2k} and c_{2k-1}. [5]

$$\begin{aligned} c_{2k} &= -i(a_k - a_k^\dagger) \\ c_{2k-1} &= a_k + a_k^\dagger \end{aligned} \tag{18}$$

They satisfy the commutation relations

$$\{c_i, c_j\} = 2\delta_{ij}\mathbb{1}$$

and therefore generate a Clifford Algebra. We define $\sigma_- = \frac{1}{2}(X - iY)$ and $\sigma_+ = \frac{1}{2}(X + iY)$. The *Jordan-Wigner-Transformation* for creation and annihilation operators is given by [14]

$$\begin{aligned} a_j &= -\left(\bigotimes_{k=1}^{j-1} Z\right) \otimes \sigma_{-j} \left(\bigotimes_{k=j+1}^{n} \mathbb{1}\right) \\ a_j^\dagger &= -\left(\bigotimes_{k=1}^{j-1} Z\right) \otimes \sigma_{+j} \left(\bigotimes_{k=j+1}^{n} \mathbb{1}\right) \end{aligned}$$

and for Majorana fermion operators by

$$c_{2k-1} = \left(\bigotimes_{k=1}^{j-1} Z \right) \otimes X_j \left(\bigotimes_{k=j+1}^{n} \mathbb{1} \right)$$

$$c_{2k} = \left(\bigotimes_{k=1}^{j-1} Z \right) \otimes Y_j \left(\bigotimes_{k=j+1}^{n} \mathbb{1} \right).$$

6.2 Annihilation and creation operators

We start by reviewing fermionic creation and annihilation operators. These operators anticommute as stated in Equation 9. We try to define an unbiased estimator of the permanent of a matrix $A \in \mathbb{C}^{n\times n}$ with the help of creation and annihilation operators. This is done similarly to the Godsil-Gutman estimator. We define the matrix B by $B_{ij} = \sqrt{A_{ij}} \cdot a_{ij}$ where a_{ij} is a fermion annihilation operator. All a_{ij} are chosen differently, i.e. are acting on different modes. In total we need n^2 modes. Note that the difference to the Godsil-Gutman estimator is, that the random variables X_{ij} have been replaced by a fermionic annihilation operator.

Definition 2 *The matrix B^* is defined by $B^*_{ij} = \sqrt{A^*_{ij}} \cdot a_{ij}$.*

In the following Lemma we explore properties of the number operator we need to formulate the unbiased estimator.

Lemma 2 *Let $\pi, \sigma \in \mathcal{S}_n$ and $M \subset \underline{n}$ an index set with $m = |M|$. It holds that $\operatorname{tr}\left[\prod_{i\in M} a^\dagger_{i\pi(i)} a_{i\sigma(i)}\right] = 2^{n^2-m}$ if $\forall i \in M : \pi(i) = \sigma(i)$ and $\operatorname{tr}\left[\prod_{i\in M} a^\dagger_{i\pi(i)} a_{i\sigma(i)}\right] = 0$ if $\exists i \in M : \pi(i) \neq \sigma(i)$.*

Proof We prove both cases.

Case 1: $\forall i \in M : \pi(i) = \sigma(i)$

We label the modes $(i, \pi(i)), i \in M$ as $l_1, \ldots, l_m$ with $l_i \in \underline{n^2}$. With the help of the Jordan-Wigner transformation and the observation that

$$\sigma_+\sigma_- = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} =: N$$

we get for the number operator of the l_i-th mode

$$N_{l_i} = \mathbb{1}_2^{\otimes^{l_i-1}} \otimes N \otimes \mathbb{1}_2^{\otimes^{n^2-l_i+1}}$$

We calculate

$$\begin{aligned}
\operatorname{tr}\left[\prod_{i\in M} a^\dagger_{i\pi(i)} a_{i\pi(i)}\right] &= \operatorname{tr}\left[\prod_{i\in M} N_{i\pi(i)}\right] \\
&= \operatorname{tr}\left[\prod_{i\in M} \mathbb{1}_2^{\otimes^{l_i-1}} \otimes N \otimes \mathbb{1}_2^{\otimes^{n^2-l_i+1}}\right] \\
&= \prod_{i\in \underline{m}} \operatorname{tr}[N] \prod_{i\in M}^{n^2-m} \operatorname{tr}[\mathbb{1}_2] \\
&= 2^{n^2-m}
\end{aligned}$$

where the property $\operatorname{tr}[A \otimes B] = \operatorname{tr}[A] \operatorname{tr}[B]$ of the trace operator has been used between line two and three.

Case 2: $\exists i \in M : \pi(i) \neq \sigma(i)$

Let k be the number of indices where $\pi(i) \neq \sigma(i)$. Then

$$\operatorname{tr}\left[\prod_{i \in M} a^{\dagger}_{i\pi(i)} a_{i\sigma(i)}\right] = (-1)^{g(m)} \operatorname{tr}\left[N_1 \dots N_{m-k} M_1 \dots M_k\right]$$

where $N_1 \dots N_{m-k}$ are number operators and $M_1 \dots M_k$ are the products which are not number operators. The sign in front may occur from rearranging the terms. We define $V = N_1 \dots N_{m-k} M_1 \dots M_{k-1}$. Then we get

$$\begin{aligned}
(-1)^{g(m)} \operatorname{tr}\left[N_1 \dots N_{m-k} M_1 \dots M_k\right] &= (-1)^{g(m)} \operatorname{tr}\left[V M_k\right] \\
&= (-1)^{g(m)} \operatorname{tr}\left[V a^{\dagger}_{j\pi(j)} a_{j\sigma(j)}\right] \\
&= -(-1)^{g(m)} \operatorname{tr}\left[a_{j\sigma(j)} V a^{\dagger}_{j\pi(j)}\right] \\
&= (-1)^{g(m)} \operatorname{tr}\left[a_{j\sigma(j)} V a^{\dagger}_{j\pi(j)}\right]
\end{aligned}$$

Where again the property $\operatorname{tr}[ABC] = \operatorname{tr}[CAB]$ of the trace has been used. This implies

$$\operatorname{tr}\left[\prod_{i \in M} a^{\dagger}_{i\pi(i)} a_{i\sigma(i)}\right] = 0 \qquad \square$$

Corollary 1 *Let $\pi, \sigma \in \mathcal{S}_n$. It holds that*

$$\operatorname{tr}\left[\prod_{i=1}^{n} a^{\dagger}_{i\pi(i)} a_{i\sigma(i)}\right] = 2^{n^2-n} \prod_{i=1}^{n} \operatorname{tr}\left[a^{\dagger}_{i\pi(i)} a_{i\sigma(i)}\right]$$

Theorem 5 *Let $\pi, \sigma, \gamma, \iota \in \mathcal{S}_n$. It holds that*

$$\operatorname{tr}\left[a^{\dagger}_{i\pi(i)} a_{i\sigma(i)} a^{\dagger}_{i\gamma(i)} a_{i\iota(i)}\right] = 2^{(n^2-2)} \left(\delta_{\pi(i)\sigma(i)} \delta_{\gamma(i)\iota(i)} + \delta_{\pi(i)\iota(i)} \delta_{\gamma(i)\sigma(i)}\right) \tag{19}$$

Proof Without loss of generality we assume that for every $i \in \underline{n}$ the permutations $\pi, \sigma, \gamma, \iota$ have to be equal in pairs, i.e. $(\pi(i) = \sigma(i) \wedge \gamma(i) = \iota(i))$ or $(\pi(i) = \gamma(i) \wedge \sigma(i) = \iota(i))$ or $(\pi(i) = \iota(i) \wedge \sigma(i) = \gamma(i))$. If not then there exists one permutation, without loss of generality $\pi(i)$, which is unequal to all other permutations, i.e. $\pi(i) \neq \sigma(i), \pi(i) \neq \gamma(i)$ and $\pi(i) \neq \iota(i)$. Let l_i be the mode associated with $(i, \pi(i))$. With the help of the Jordan-Wigner-Transformation we get a tensor decomposition

$$a^{\dagger}_{i\pi(i)} a_{i\sigma(i)} a^{\dagger}_{i\gamma(i)} a_{i\iota(i)} = A \otimes (\sigma_{+/-})_{l_i} (Z_{l_i})^c \otimes B \tag{20}$$

with operators $A \in (\mathbb{C}^{2\times 2})^{\otimes(l_i-1)}$, $B \in (\mathbb{C}^{2\times 2})^{\otimes(n^2-l_i+1)}$ and a constant $c \in \mathbb{N} \cup \{0\}$. The trace property $\operatorname{tr}[A \otimes B] = \operatorname{tr}[A] \operatorname{tr}[B]$ and the property $\operatorname{tr}\left[(Z_{l_i})^{c_1} \sigma_{+/-} (Z_{l_i})^{c_2}\right] = 0$ imply that the trace over Equation 20 is zero and Equation 19 holds. We proceed by case differentiation:

Case 1: $(\pi(i) = \sigma(i) \wedge \gamma(i) = \iota(i)) \wedge \pi(i) \neq \gamma(i)$

We get $\mathrm{tr}\left[a^\dagger_{i\pi(i)} a_{i\sigma(i)} a^\dagger_{i\gamma(i)} a_{i\iota(i)}\right] = \mathrm{tr}\left[a^\dagger_{i\pi(i)} a_{i\pi(i)} a^\dagger_{i\gamma(i)} a_{i\gamma(i)}\right]$ and Equation 19 holds by Lemma 2.

Case 2: $(\pi(i) = \gamma(i) \wedge \sigma(i) = \iota(i)) \wedge \pi(i) \neq \sigma(i)$

We get $\mathrm{tr}\left[a^\dagger_{i\pi(i)} a_{i\sigma(i)} a^\dagger_{i\gamma(i)} a_{i\iota(i)}\right] = \mathrm{tr}\left[a^\dagger_{i\pi(i)} a_{i\sigma(i)} a^\dagger_{i\pi(i)} a_{i\sigma(i)}\right] = -\,\mathrm{tr}\left[a^\dagger_{i\pi(i)} a^\dagger_{i\pi(i)} a_{i\sigma(i)} a_{i\sigma(i)}\right] = 0$ with the anti commutation of the creation and annihilation operators explained in 3.3 and Equation 19 holds.

Case 3: $(\pi(i) = \iota(i) \wedge \sigma(i) = \gamma(i)) \wedge \pi(i) \neq \gamma(i)$

We get $\mathrm{tr}\left[a^\dagger_{i\pi(i)} a_{i\sigma(i)} a^\dagger_{i\gamma(i)} a_{i\iota(i)}\right] = \mathrm{tr}\left[a^\dagger_{i\pi(i)} a_{i\sigma(i)} a^\dagger_{i\sigma(i)} a_{i\pi(i)}\right] = \mathrm{tr}\left[a^\dagger_{i\pi(i)} a_{i\pi(i)} a^\dagger_{i\gamma(i)} a_{i\gamma(i)}\right]$ and Equation 19 holds by Lemma 2.

Case 4: $(\pi(i) = \iota(i) = \sigma(i) = \gamma(i))$

A straightforward calculation with the Jordan-Wigner-Transformation implies that $\mathrm{tr}\left[a^\dagger_{i\pi(i)} a_{i\pi(i)} a^\dagger_{i\pi(i)} a_{i\pi(i)}\right] = 2^{n^2-1}$ and so Equation 19 holds. □

Theorem 6 *Let $\pi, \sigma, \gamma, \iota \in \mathcal{S}_n$. It holds that*

$$\mathrm{tr}\left[\prod_{i=1}^{n} a^\dagger_{i\pi(i)} a_{i\sigma(i)} a^\dagger_{i\gamma(i)} a_{i\iota(i)}\right] = 2^{n^2-n^3} \prod_{i=1}^{n} \mathrm{tr}\left[a^\dagger_{i\pi(i)} a_{i\sigma(i)} a^\dagger_{i\gamma(i)} a_{i\iota(i)}\right] \tag{21}$$

Proof Without loss of generality we assume that for every $i \in \underline{n}$ the permutations $\pi, \sigma, \gamma, \iota$ have to be equal in pairs, i.e. $(\pi(i) = \sigma(i) \wedge \gamma(i) = \iota(i))$ or $(\pi(i) = \gamma(i) \wedge \sigma(i) = \iota(i))$ or $(\pi(i) = \iota(i) \wedge \sigma(i) = \gamma(i))$. If not then the same argument as in the proof of Theorem 5 implies that the trace on both sides of Equation 21 is zero and so the equation holds. If there exists an $i \in \underline{n}$ such that $(\pi(i) = \gamma(i) \wedge \sigma(i) = \iota(i))$ and $\gamma(i) \neq \sigma(i)$ then the anti commutativity of the creation and annihilation operators explained in 3.3 implies that the trace on both sides is zero and the equation holds. So we are left with $\forall i \in \underline{n} : (\pi(i) = \sigma(i) \wedge \gamma(i) = \iota(i)) \vee (\pi(i) = \iota(i) \wedge \sigma(i) = \gamma(i))$. In both cases we get two number operator for each i. Therefore we can write

$$\mathrm{tr}\left[\prod_{i=1}^{n} a^\dagger_{i\pi(i)} a_{i\sigma(i)} a^\dagger_{i\gamma(i)} a_{i\iota(i)}\right] = \mathrm{tr}\left[\prod_{i=1}^{n} N_{i\pi'(i)} N_{i\sigma'(i)}\right]$$

for some permutations $\pi', \sigma' \in \mathcal{S}_n$. Now we define l to be number of i where $N_{i,\pi'(i)} = N_{i,\sigma'(i)}$, i.e. $l = \left|\left\{i \in \underline{n} | N_{i,\pi'(i)} = N_{i,\sigma'(i)}\right\}\right|$. The Jordan-Wigner transformation and the trace property $\mathrm{tr}[A \otimes B] = \mathrm{tr}[A]\,\mathrm{tr}[B]$ imply after a straightforward calculation that

$$\begin{aligned}\mathrm{tr}\left[\prod_{i=1}^{n} N_{i\pi'(i)} N_{i\sigma'(i)}\right] &= 2^{n^2-2\cdot(n-l)-l} \\ &= 2^{n^2-2n+l}\end{aligned}$$

and

$$\begin{aligned}\prod_{i=1}^{n} \mathrm{tr}\left[N_{i\pi'(i)} N_{i\sigma'(i)}\right] &= \left(2^{n^2-2}\right)^{(n-l)} \cdot \left(2^{n^2-1}\right)^{l} \\ &= 2^{n^3-2n+l}.\end{aligned}$$

We write

$$\begin{aligned}\operatorname{tr}\left[\prod_{i=1}^{n} N_{i\pi'(i)} N_{i\sigma'(i)}\right] &= 2^{n^2-2n+l} \\ &= 2^{n^2-n^3} 2^{n^3-2n+l} \\ &= 2^{n^2-n^3} \prod_{i=1}^{n} \operatorname{tr}\left[N_{i\pi'(i)} N_{i\sigma'(i)}\right]\end{aligned}$$

which completes the proof. □

With these properties the unbiased estimator of the permanent is given by

Theorem 7 *There exists a function $f : \mathbb{N} \to \mathbb{N}$ such that*

$$\operatorname{Perm} A = \frac{(-1)^{f(n)}}{2^{n^2-n}} \operatorname{tr}\left[(\det B^*)^\dagger (\det B)\right] \tag{22}$$

Proof

$$\begin{aligned}\operatorname{tr}\left[(\det B^*)^\dagger (\det B)\right] &= \sum_{\sigma\in\mathcal{S}_n,\pi\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi)\operatorname{tr}\left[\left(\prod_{i=1}^{n}\sqrt{A^*_{i\pi(i)}}a_{i\pi(i)}\right)^\dagger\left(\prod_{i=1}^{n}\sqrt{A_{i\sigma(i)}}a_{i\sigma(i)}\right)\right] \\ &= \sum_{\sigma\in\mathcal{S}_n,\pi\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi)\operatorname{tr}\left[\left(\prod_{i=n}^{1}\sqrt{A_{i\pi(i)}}a^\dagger_{i\pi(i)}\right)\left(\prod_{i=1}^{n}\sqrt{A_{i\sigma(i)}}a_{i\sigma(i)}\right)\right] \\ &= (-1)^{f(n)} \sum_{\sigma\in\mathcal{S}_n,\pi\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi)\operatorname{tr}\left[\left(\prod_{i=1}^{n}\sqrt{A_{i\pi(i)}}\sqrt{A_{i\sigma(i)}}a^\dagger_{i\pi(i)}a_{i\sigma(i)}\right)\right] \\ &= (-1)^{f(n)}\cdot 2^{n^2-n} \sum_{\sigma\in\mathcal{S}_n,\pi\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi)\left(\prod_{i=1}^{n}\operatorname{tr}\left[\sqrt{A_{i\pi(i)}}\sqrt{A_{i\sigma(i)}}a^\dagger_{i\pi(i)}a_{i\sigma(i)}\right]\right) \\ &= (-1)^{f(n)}\cdot 2^{n^2-n}\sum_{\pi\in\mathcal{S}_n}\prod_{i=1}^{n} A_{i\pi(i)} \\ &= (-1)^{f(n)}\cdot 2^{n^2-n}\cdot \operatorname{Perm} A\end{aligned}$$

Where Corollary 1 has been used between line three and four. □

This estimator looks like the evaluation of an expectation value in a quantum mechanical system. It is the trace of a non-linear combination of fermionic creation and annihilation operators. In order to make contact to a physical system, we introduce the the density operator $\rho = \frac{1}{2^{n^2}}\mathbb{1}_{2^{n^2}}$. This density operator is hermitian, positive definite, and has trace one. With the help of the density operator we rewrite Equation 27

$$\begin{aligned}\operatorname{Perm} A &= \frac{(-1)^{f(n)}2^{n^2}}{2^{n^2-n}} \operatorname{tr}\left[\rho(\det B^*)^\dagger (\det B)\right] \\ &= (-1)^{f(n)}\cdot 2^n \operatorname{tr}\left[\rho\widetilde{H}\right]\end{aligned}$$

where the definition $\widetilde{H} = (\det B^*)^\dagger(\det B)$ has been used. This can be seen as calculating the expectation value of the observable $\widetilde{H}$ for a system in a completely mixed state given by ρ.
$\widetilde{H}$ is not hermitian and therefore cannot relate to a physical observable. We define the hermitian operator H out of $\widetilde{H}$ by $H = \frac{\widetilde{H}+\widetilde{H}^\dagger}{2}$. The expectation value of H corresponds to the real part of the permanent.

$$\Re\,\mathrm{Perm}\,A = (-1)^{f(n)} \cdot 2^n \,\mathrm{tr}\left[\rho\frac{\widetilde{H}+\widetilde{H}^\dagger}{2}\right] \tag{23}$$

$$= (-1)^{f(n)} \cdot 2^n \,\mathrm{tr}\left[\rho H\right] \tag{24}$$

Analogously we receive $\Im\,\mathrm{Perm}\,A$ by calculating the expectation value of the hermitian operator $H' = \frac{\widetilde{H}-\widetilde{H}^\dagger}{2i}$. [1]

6.2.1 Error estimation

The variance and so the error of the estimator is determined by its second moment.

$$\begin{aligned}\mathrm{tr}\left[\rho\widetilde{H}^2\right] &= \frac{1}{2^{n^2}}\,\mathrm{tr}\left[(\det B^*)^\dagger(\det B)(\det B^*)^\dagger(\det B)\right]\\ &= \frac{1}{2^{n^2}}\sum_{\pi,\gamma,\sigma,\iota\in\mathcal{S}_n}\mathrm{sgn}(\pi\sigma\gamma\iota)\\ &\quad\mathrm{tr}\left[\left(\prod_{i=1}^n\sqrt{A^*_{i\pi(i)}}a_{i\pi(i)}\right)^\dagger\left(\prod_{i=1}^n\sqrt{A_{i\sigma(i)}}a_{i\sigma(i)}\right)\left(\prod_{i=1}^n\sqrt{A^*_{i\gamma(i)}}a_{i\gamma(i)}\right)^\dagger\left(\prod_{i=1}^n\sqrt{A_{i\iota(i)}}a_{i\iota(i)}\right)\right]\end{aligned}$$

We simplify the trace term

$$\begin{aligned}&\mathrm{tr}\left[\left(\prod_{i=1}^n\sqrt{A^*_{i\pi(i)}}a_{i\pi(i)}\right)^\dagger\left(\prod_{i=1}^n\sqrt{A_{i\sigma(i)}}a_{i\sigma(i)}\right)\left(\prod_{i=1}^n\sqrt{A^*_{i\gamma(i)}}a_{i\gamma(i)}\right)^\dagger\left(\prod_{i=1}^n\sqrt{A_{i\iota(i)}}a_{i\iota(i)}\right)\right]\\ &= \mathrm{tr}\left[\left(\prod_{i=n}^1\sqrt{A_{i\pi(i)}}a^\dagger_{i\pi(i)}\right)\left(\prod_{i=1}^n\sqrt{A_{i\sigma(i)}}a_{i\sigma(i)}\right)\left(\prod_{i=n}^1\sqrt{A_{i\gamma(i)}}a^\dagger_{i\gamma(i)}\right)\left(\prod_{i=1}^n\sqrt{A_{i\iota(i)}}a_{i\iota(i)}\right)\right]\\ &= \mathrm{tr}\left[\left(\prod_{i=1}^n\sqrt{A_{i\pi(i)}A_{i\sigma(i)}}a^\dagger_{i\pi(i)}a_{i\sigma(i)}\right)\left(\prod_{i=1}^n\sqrt{A_{i\gamma(i)}A_{i\iota(i)}}a^\dagger_{i\gamma(i)}a_{i\iota(i)}\right)\right]\\ &= \mathrm{tr}\left[\prod_{i=1}^n\sqrt{A_{i\pi(i)}A_{i\sigma(i)}A_{i\gamma(i)}A_{i\iota(i)}}a^\dagger_{i\pi(i)}a_{i\sigma(i)}a^\dagger_{i\gamma(i)}a_{i\iota(i)}\right]\\ &= 2^{n^2-n^3}\prod_{i=1}^n\mathrm{tr}\left[\sqrt{A_{i\pi(i)}A_{i\sigma(i)}A_{i\gamma(i)}A_{i\iota(i)}}a^\dagger_{i\pi(i)}a_{i\sigma(i)}a^\dagger_{i\gamma(i)}a_{i\iota(i)}\right] && \text{(Theorem 6)}\\ &= 2^{n^2-n^3}\prod_{i=1}^n A_{i\pi(i)}A_{i\gamma(i)}2^{(n^2-2)}\left(\delta_{\pi(i)\sigma(i)}\delta_{\gamma(i)\iota(i)}+\delta_{\pi(i)\iota(i)}\delta_{\gamma(i)\sigma(i)}\right) && \text{(Theorem 5)}\\ &= 2^{n^2-2n}\prod_{i=1}^n A_{i\pi(i)}A_{i\gamma(i)}\left(\delta_{\pi(i)\sigma(i)}\delta_{\gamma(i)\iota(i)}+\delta_{\pi(i)\iota(i)}\delta_{\gamma(i)\sigma(i)}\right)\end{aligned}$$

[1] If we restrict ourself to a real matrices A we get an hermitian operator H by defining $H = \det B \det \widetilde{B}$ with $B_{ij} = \sqrt{|A_{ij}|}a_{ij}$ and $\widetilde{B}_{ij} = (-1)^{f_{ij}}\sqrt{|A_{ij}|}a_{ij}$ where f_{ij} is equal to 1 if A_{ij} is negative and zero otherwise. The error estimation remains the same.

Hence the second moment is given by

$$\operatorname{tr}\left[\rho \widetilde{H}^2\right] = \frac{1}{2^{2n}} \sum_{\pi,\gamma} \prod_{i=1}^{n} A_{i\pi(i)} A_{i\gamma(i)} \sum_{\sigma,\iota} \left(\delta_{\pi(i)\sigma(i)}\delta_{\gamma(i)\iota(i)} + \delta_{\pi(i)\iota(i)}\delta_{\gamma(i)\sigma(i)}\right).$$

We compare the variances of the different estimators in 6.4.

6.3 Majorana fermion operators

We continue with constructing an unbiased estimator for the permanent with the help of Majorana fermion operators defined in Equation 18. We proceed similarly to the previous section. The estimator and the variance are going to be similar but differ in some detail. We define the matrix B by $B_{ij} = \sqrt{A_{ij}} \cdot c_{ij}$ where c_{ij} is a Majorana fermion operator. All c_{ij} are chosen differently.

Again we explore the properties of the trace of Majorana fermion operators before we formulate the estimator.

Lemma 3 *Let $\pi, \sigma \in \mathcal{S}_n$ and $M \subset \underline{n}$ an index set. It holds that* $\mathrm{tr}\left[\prod_{i \in M} c_{i\pi(i)} c_{i\sigma(i)}\right] = 2^{n^2}$ *if* $\forall i \in M : \pi(i) = \sigma(i)$ *and* $\mathrm{tr}\left[\prod_{i \in M} c_{i\pi(i)} c_{i\sigma(i)}\right] = 0$ *if* $\exists i \in M : \pi(i) \neq \sigma(i)$.

PROOF We prove both cases.

Case 1: $\forall i \in M : \pi(i) = \sigma(i)$

Calculation with the Jordan-Wigner-Transformation for Majorana fermion operators analogous to Lemma 2.

Case 2: $\exists i \in M : \pi(i) \neq \sigma(i)$

Let $i \in M$ be the index with $\pi(i) \neq \sigma(i)$. Let l_i be the mode associated with $(i, \pi(i))$. Without loss of generality let l_i be an even number. With the help of the Jordan-Wigner-Transformation we get a tensor decomposition

$$\prod_{i \in M} c_{i\pi(i)} c_{i\sigma(i)} = A \otimes (Z)^{c_1} Y (Z)^{c_2} \otimes B$$

with operators $A \in \left(\mathbb{C}^{2\times 2}\right)^{\otimes (l_i - 1)}$, $B \in \left(\mathbb{C}^{2\times 2}\right)^{\otimes (n^2 - l_i + 1)}$ and constants $c_1, c_2 \in \mathbb{N} \cup \{0\}$.

The trace property $\mathrm{tr}[A \otimes B] = \mathrm{tr}[A] \cdot \mathrm{tr}[B]$ and the property $\mathrm{tr}[(Z)^{c_1} Y (Z)^{c_2}] = 0$ of the Pauli operators imply that

$$\mathrm{tr}\left[\prod_{i \in M} c_{i\pi(i)} c_{i\sigma(i)}\right] = 0\,.$$

Corollary 2 *Let $\pi, \sigma \in \mathcal{S}_n$. It holds that*

$$\mathrm{tr}\left[\prod_{i=1}^{n} c_{i\pi(i)} c_{i\sigma(i)}\right] = 2^{n^2} \prod_{i=1}^{n} \mathrm{tr}\left[c_{i\pi(i)} c_{i\sigma(i)}\right]$$

Theorem 8 *Let $\pi, \sigma, \gamma, \iota \in \mathcal{S}_n$. It holds that*

$$\mathrm{tr}\left[c_{i\pi(i)} c_{i\sigma(i)} c_{i\gamma(i)} c_{i\iota(i)}\right] = 2^{n^2} \left(\delta_{\pi(i)\sigma(i)} \delta_{\gamma(i)\iota(i)} + \delta_{\pi(i)\iota(i)} \delta_{\gamma(i)\sigma(i)} - \delta_{\pi(i)\gamma(i)} \delta_{\sigma(i)\iota(i)}\right)$$

PROOF The proof is analogous to the proof of Theorem 5. The differences are that we use Lemma 3 instead of Lemma 2 and case 2 contributes with a minus sign in front. □

Theorem 9 *Let $\pi, \sigma, \gamma, \iota \in \mathcal{S}_n$. It holds that*

$$\mathrm{tr}\left[\prod_{i=1}^{n} c_{i\pi(i)} c_{i\sigma(i)} c_{i\gamma(i)} c_{i\iota(i)}\right] = 2^{n^2 - n^3} \prod_{i=1}^{n} \mathrm{tr}\left[c_{i\pi(i)} c_{i\sigma(i)} c_{i\gamma(i)} c_{i\iota(i)}\right] \tag{25}$$

Proof Without loss of generality we assume that for every $i \in \underline{n}$ the permutations $\pi, \sigma, \gamma, \iota$ have to be equal in pairs, i.e. $(\pi(i) = \sigma(i) \wedge \gamma(i) = \iota(i))$ or $(\pi(i) = \gamma(i) \wedge \sigma(i) = \iota(i))$ or $(\pi(i) = \iota(i) \wedge \sigma(i) = \gamma(i))$. If not then there exists a $i \in \underline{n}$ such that one permutation, without loss of generality $\pi(i)$, is unequal to all other permutations, i.e. $\pi(i) \neq \sigma(i), \pi(i) \neq \gamma(i)$ and $\pi(i) \neq \iota(i)$. Let l_i be the mode associated with $(i, \pi(i))$. Without loss of generality let l_i be an even number. With the help of the Jordan-Wigner-Transformation we get a tensor decomposition

$$\prod_{i=1}^{n} c_{i\pi(i)} c_{i\sigma(i)} c_{i\gamma(i)} c_{i\iota(i)} = A \otimes (Z)^{c_1} Y (Z)^{c_2} \otimes B \tag{26}$$

with operators $A \in (\mathbb{C}^{2\times 2})^{\otimes(l_i-1)}$, $B \in (\mathbb{C}^{2\times 2})^{\otimes(n^2-l_i+1)}$ and constants $c_1, c_2 \in \mathbb{N} \cup \{0\}$. The trace property $\operatorname{tr}[A \otimes B] = \operatorname{tr}[A] \operatorname{tr}[B]$ and the property $\operatorname{tr}[(Z)^{c_1} Y (Z)^{c_2}] = 0$ of the Pauli operators imply that the trace over Equation 26 is zero and Equation 25 holds.
The Jordan-Wigner-Transformation implies that $c_{ij}^2 = \mathbb{1}_{2^{n^2}}$. Therefore we write

$$\begin{aligned}
\operatorname{tr}\left[\prod_{i=1}^{n} c_{i\pi(i)} c_{i\sigma(i)} c_{i\gamma(i)} c_{i\iota(i)}\right] &= \operatorname{tr}\left[\prod_{i=1}^{n} \mathbb{1}_{2^{n^2}}\right] \\
&= 2^{n^2} = 2^{n^2-n^3} \prod_{i=1}^{n} 2^{n^2} \\
&= 2^{n^2-n^3} \prod_{i=1}^{n} \operatorname{tr}\left[\mathbb{1}_{2^{n^2}}\right] \\
&= 2^{n^2-n^3} \prod_{i=1}^{n} \operatorname{tr}\left[c_{i\pi(i)} c_{i\sigma(i)} c_{i\gamma(i)} c_{i\iota(i)}\right]
\end{aligned}$$

□

We formulate the unbiased estimator of the permanent.

Theorem 10 *There exists a function $f : \mathbb{N} \to \mathbb{N}$ such that*

$$\operatorname{Perm} A = \frac{(-1)^{f(n)}}{2^{n^2}} \operatorname{tr}\left[(\det B)^2\right] \tag{27}$$

Proof

$$\begin{aligned}
\operatorname{tr}\left[(\det B)(\det B)\right] &= \sum_{\sigma \in \mathcal{S}_n, \pi \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \operatorname{sgn}(\pi) \operatorname{tr}\left[\left(\prod_{i=1}^{n} \sqrt{A_{i\pi(i)}} c_{i\pi(i)}\right)\left(\prod_{i=1}^{n} \sqrt{A_{i\sigma(i)}} c_{i\sigma(i)}\right)\right] \\
&= (-1)^{f(n)} \sum_{\sigma \in \mathcal{S}_n, \pi \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \operatorname{sgn}(\pi) \operatorname{tr}\left[\left(\prod_{i=1}^{n} \sqrt{A_{i\pi(i)}} \sqrt{A_{i\sigma(i)}} c_{i\pi(i)} c_{i\sigma(i)}\right)\right] \\
&= (-1)^{f(n)} \cdot 2^{n^2} \sum_{\sigma \in \mathcal{S}_n, \pi \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \operatorname{sgn}(\pi) \left(\prod_{i=1}^{n} \operatorname{tr}\left[\sqrt{A_{i\pi(i)}} \sqrt{A_{i\sigma(i)}} c_{i\pi(i)} c_{i\sigma(i)}\right]\right) \\
&= (-1)^{f(n)} \cdot 2^{n^2} \sum_{\pi \in \mathcal{S}_n} \prod_{i=1}^{n} A_{i\pi(i)} \\
&= (-1)^{f(n)} \cdot 2^{n^2} \cdot \operatorname{Perm} A
\end{aligned}$$

Where Corollary 2 has been used between line two and three. □

Completely analogous to 6.2 we define the density operator $\rho = \frac{1}{2^{n^2}} \mathbb{1}_{n^2}$ and the matrix $\widetilde{H} = (\det B)^2$ to make contact to a physical system. We get

$$\begin{aligned}
\operatorname{Perm} A &= \frac{(-1)^{f(n)} 2^{n^2}}{2^{n^2}} \operatorname{tr}\left[\rho \widetilde{H}\right] \\
&= (-1)^{f(n)} \operatorname{tr}\left[\rho \widetilde{H}\right].
\end{aligned}$$

Analogues to 6.2 we define the hermitian operators $H = \frac{\widetilde{H}+\widetilde{H}^\dagger}{2}$ and $H' = \frac{\widetilde{H}-\widetilde{H}}{2i}$ to calculate the real and imaginary part of the permanent from the expectation value of a quantum mechanical hermitian operator.

6.3.1 Error Estimation

The variance is determined by the second moment.

$$\begin{aligned}
\operatorname{tr}\left[\rho \widetilde{H}^2\right] &= \frac{1}{2^{n^2}} \operatorname{tr}\left[(\det B)(\det B)(\det B)(\det B)\right] \\
&= \frac{1}{2^{n^2}} \sum_{\pi,\gamma,\sigma,\iota \in \mathcal{S}_n} \operatorname{sgn}(\pi\sigma\gamma\iota) \\
&\quad \operatorname{tr}\left[\left(\prod_{i=1}^n \sqrt{A_{i\pi(i)}} c_{i\pi(i)}\right)\left(\prod_{i=1}^n \sqrt{A_{i\sigma(i)}} c_{i\sigma(i)}\right)\left(\prod_{i=1}^n \sqrt{A_{i\gamma(i)}} c_{i\gamma(i)}\right)\left(\prod_{i=1}^n \sqrt{A_{i\iota(i)}} c_{i\iota(i)}\right)\right]
\end{aligned}$$

We simplify the trace term

$$\begin{aligned}
&\operatorname{tr}\left[\left(\prod_{i=1}^n \sqrt{A_{i\pi(i)}} c_{i\pi(i)}\right)\left(\prod_{i=1}^n \sqrt{A_{i\sigma(i)}} c_{i\sigma(i)}\right)\left(\prod_{i=1}^n \sqrt{A_{i\gamma(i)}} c_{i\gamma(i)}\right)\left(\prod_{i=1}^n \sqrt{A_{i\iota(i)}} c_{i\iota(i)}\right)\right] \\
&= \operatorname{tr}\left[\left(\prod_{i=1}^n \sqrt{A_{i\pi(i)} A_{i\sigma(i)}} c_{i\pi(i)} c_{i\sigma(i)}\right)\left(\prod_{i=1}^n \sqrt{A_{i\gamma(i)} A_{i\iota(i)}} c_{i\gamma(i)} c_{i\iota(i)}\right)\right] \\
&= \operatorname{tr}\left[\prod_{i=1}^n \sqrt{A_{i\pi(i)} A_{i\sigma(i)} A_{i\gamma(i)} A_{i\iota(i)}} c_{i\pi(i)} c_{i\sigma(i)} c_{i\gamma(i)} c_{i\iota(i)}\right] \\
&= 2^{n^2-n^3} \prod_{i=1}^n \operatorname{tr}\left[\sqrt{A_{i\pi(i)} A_{i\sigma(i)} A_{i\gamma(i)} A_{i\iota(i)}} c_{i\pi(i)} c_{i\sigma(i)} c_{i\gamma(i)} a_{i\iota(i)}\right] && \text{(Theorem 9)} \\
&= 2^{n^2-n^3} \prod_{i=1}^n A_{i\pi(i)} A_{i\gamma(i)} 2^{n^2} \left(\delta_{\pi(i)\sigma(i)} \delta_{\gamma(i)\iota(i)} + \delta_{\pi(i)\iota(i)} \delta_{\gamma(i)\sigma(i)} - \delta_{\pi(i)\gamma(i)} \delta_{\sigma(i)\iota(i)}\right) && \text{(Theorem 8)} \\
&= 2^{n^2} \prod_{i=1}^n A_{i\pi(i)} A_{i\gamma(i)} \left(\delta_{\pi(i)\sigma(i)} \delta_{\gamma(i)\iota(i)} + \delta_{\pi(i)\iota(i)} \delta_{\gamma(i)\sigma(i)} - \delta_{\pi(i)\gamma(i)} \delta_{\sigma(i)\iota(i)}\right)
\end{aligned}$$

Hence the second moment is given by

$$\begin{aligned}
&\operatorname{tr}\left[\rho \widetilde{H}^2\right] \\
&= \sum_{\pi,\sigma,\gamma,\iota \in \mathcal{S}_n} \prod_{i=1}^n \sqrt{A_{i\pi(i)} A_{i\sigma(i)} A_{i\gamma(i)} A_{i\iota(i)}} \left(\delta_{\pi(i)\sigma(i)} \delta_{\gamma(i)\iota(i)} + \delta_{\pi(i)\iota(i)} \delta_{\gamma(i)\sigma(i)} - \delta_{\pi(i)\gamma(i)} \delta_{\sigma(i)\iota(i)}\right).
\end{aligned}$$

6.4 Comparison

In this section we compare the variances of the Godsil-Gutman and Barvinok estimator and the two estimators using fermionic operators from the sections 6.2 and 6.3. We compare the variances by comparing the respective second moments. The second moment of the Barvinok estimator is given by $E[\det(B(X))^4]$ which evaluates with Isserlis-Theorem [23] to

$$E[\det(B(X))^4] = \sum_{\pi,\sigma,\gamma,\iota \in \mathcal{S}_n} \prod_{i=1}^{n} \sqrt{A_{i\pi(i)} A_{i\sigma(i)} A_{i\gamma(i)} A_{i\iota(i)}} \left(\delta_{\pi(i)\sigma(i)}\delta_{\gamma(i)\iota(i)} + \delta_{\pi(i)\iota(i)}\delta_{\gamma(i)\sigma(i)} + \delta_{\pi(i)\gamma(i)}\delta_{\sigma(i)\iota(i)}\right). \tag{28}$$

The second moment of the Godsil-Gutman estimator has no expression with Kronecker deltas. This is due to the use of Bernoulli and not Gaussian random variables. The expectation value $E[X^4]$ is equal to one for Bernoulli random variables X but equal to three for Gaussian random variables X. Therefore differences in the variance between the Barvinok and Godsil-Gutman estimator arise in parts of the sum in Equation 28 where $\pi(i) = \gamma(i) = \sigma(i) = \iota(i)$. The second moments of the annihilation operator estimator and Majorana fermion operator estimator are according to 6.2 and 6.3 given by

$$\mathrm{tr}\left[\rho \widetilde{H}^2\right] \propto \sum_{\pi,\sigma,\gamma,\iota \in \mathcal{S}_n} \prod_{i=1}^{n} A_{i\pi(i)} A_{i\gamma(i)} \left(\delta_{\pi(i)\sigma(i)}\delta_{\gamma(i)\iota(i)} + \delta_{\pi(i)\iota(i)}\delta_{\gamma(i)\sigma(i)}\right) \tag{29}$$

$$\mathrm{tr}\left[\rho \widetilde{H}^2\right] \propto \sum_{\pi,\sigma,\gamma,\iota \in \mathcal{S}_n} \prod_{i=1}^{n} \sqrt{A_{i\pi(i)} A_{i\sigma(i)} A_{i\gamma(i)} A_{i\iota(i)}} \left(\delta_{\pi(i)\sigma(i)}\delta_{\gamma(i)\iota(i)} + \delta_{\pi(i)\iota(i)}\delta_{\gamma(i)\sigma(i)} - \delta_{\pi(i)\gamma(i)}\delta_{\sigma(i)\iota(i)}\right) \tag{30}$$

where the first line contains the variance of the former and the second line the variance of the latter.

We see that the main difference of the three variances in Equation 28, Equation 29 and Equation 30 is given by the delta term at the end of the expressions. In the variance of the Barvinok estimator the term $\delta_{\pi(i)\gamma(i)}\delta_{\sigma(i)\iota(i)}$ contributes with a plus sign. In the variance of the annihilation and creation operator estimator it is gone and in the variance of the Majorana fermion operator estimator it contributes with a minus sign.

In order to compare the variances we look at the identity matrix $\mathbb{1}_n$. The second moment of the Barvinok estimator for the identity is $3^n = 3^n \cdot (\mathrm{Perm}\, A)^2$. The second moment of the annihilation operator estimator for the identity matrix is $2^n = 2^n \cdot (\mathrm{Perm}\, A)^2$. Thus in both cases the relative variance is exponentially large in the dimension of the matrix. So the annihilation and creation operator estimator does not have a small variance for a general complex valued matrix.

We expect the variance of the Majorana fermion operator estimator to be the smallest of the three variances because the third delta term contributes with a minus sign and thus cancellation of terms can happen. Unfortunately we have not been able to proof that so far. What we would like to show is that

$$\mathrm{tr}\left[\rho \widetilde{H}^2\right] = O(\mathrm{poly}(n))(\mathrm{Perm}\, A)^2$$

such that the relative variance is polynomial in n. For example this would be the case if the second delta term in Equation 29 vanishes or the minus term in front of the last delta expression

in Equation 30 cancels another delta term contributing with a plus exactly. This is not the case. There is still some leftover contribution.
We tried to get an idea of the variance of the Majorana operator estimator numerically, but due to the nested sum over four permutations and the high computational complexity of $O((n!)^4)$ we have not been able to get conclusive results.

7 Conclusion

In this bachelor thesis we have shown that *Boson Sampling* as well as approximating the permanent of complex valued matrices has a *Sign Problem* and that the computational complexity of approximating the permanent of complex valued matrices is solely due to the *Sign Problem*. As an open question we leave the relation of the approximation of permanents to fermionic physics. We have managed to introduce a fermionic computation of the permanent with the help of Majorana fermion operators. We did not succeed in showing that this fermionic simulation is time efficient, i.e. approximates the the permanent of a complex valued matrix with high probability in polynomial time.

8 References

[1] Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics. *ArXiv e-prints*, 2014.

[2] Scott Aaronson and Travis Hance. Generalizing and derandomizing gurvits's approximation algorithm for the permanent. *Quantum Info. Comput.*, 14(7&8):541–559, May 2014.

[3] Alexander Barvinok. Polynomial time algorithms to approximate permanents and mixed discriminants within a simply exponential factor. *Random Struct. Algorithms*, 14(1):29–61, January 1999.

[4] Markus Bläser. Noncommutativity makes determinants hard. *Information and Computation*, 243:133 – 144, 2015. 40th International Colloquium on Automata, Languages and Programming (ICALP 2013).

[5] Sergey Bravyi, Barbara M Terhal, and Bernhard Leemhuis. Majorana fermion codes. *New Journal of Physics*, 12(8):083039, 2010.

[6] Steve Chien, Lars Rasmussen, and Alistair Sinclair. Clifford algebras and approximating the permanent. *Journal of Computer and System Sciences*, 67(2):263 – 290, 2003. Special Issue on {STOC} 2002.

[7] Kevin P. Costello and Van Vu. Concentration of random determinants and permanent estimators. *SIAM J. Discrete Math.*, 23(3):1356–1371, 2009.

[8] R. Laflamme E. Knill and G.J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409:46–52, 2001.

[9] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6):467–488, 1982.

[10] Matthias Lehner Gerd Fischer and Angela Puchert. *Einfuehrung in die Stochastik*. Springer Spektrum, 2015.

[11] Christopher Gerry and Peter Knight. *Introductory Quantum Optics*. Cambridge University Press, 2004.

[12] C.D. Godsil and I. Gutman. On the matching polynomial of a graph. *Algebraic Methods in Graph Theory I-II*, pages 67–83, 1981.

[13] Richard Jozsa and Akimasa Miyake. Matchgates and classical simulation of quantum circuits. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 464(2100):3089–3106, 2008.

[14] Richard Jozsa, Akimasa Miyake, and Sergii Strelchuk. Jordan-wigner formalism for arbitrary 2-input 2-output matchgates and their classical simulation. *Quantum Info. Comput.*, 15(7-8):541–556, May 2015.

[15] N. Karmarkar R. Karp R. Lipton L. Lovasz and M. Luby. A monte-carlo algorithm for estimating the permanent. *Siam J. Comput*, 22:284–293, 1991.

[16] Alistair Sinclair Mark Jerrum and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *Journal of the ACM*, 51:671–697, 2004.

[17] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, 2000.

[18] Sergey Bravyi David P. DiVincenzo Roberto I. Oliveira and Barbara M. Terhal. The complexity of stoquastic local hamiltonian problems. *Quant. Inf. Comp.*, 8:361–385, 2008.

[19] Reuven Y. Rubinstein and Dirk P. Kroese. *Simulation And The Monte Carlo Method.* John Wiley & Sons, 1991.

[20] Barbara M. Terhal and David P. Divincenzo. Classical simulation of noninteracting-fermion quantum circuits. *Phys. Rev. A*, 65:0108010, 2002.

[21] Matthias Troyer and Uwe-Jens Wiese. Computational complexity and fundamental limitations to fermionic quantum monte carlo simulations. *Phys. Rev. Lett.*, 94, May 2005.

[22] L.G. Vailant. The complexity of computing the permanent. *Theoretical Comput. Sci.*, 8, 1979.

[23] C. Vignat. A generalized isserlis theorem for location mixtures of gaussian random vectors. *Statistics and Probability Letters*, 82(1):67 – 71, 2012.