Optimization problems with non-commuting variables and applications to quantum many-body system

Haoyu Hu

1 Introduction

A general optimization problem with commuting variables can be written as

$$\begin{aligned} p^\star = &\min_{x\in\mathbb{R}^n} h(x) \\ &\text{s.t.} g_i(x) \leq 0 \end{aligned}$$

where all the variables x_i are commuting with each other: $x_i x_j - x_j x_i = 0$. Here, we consider an optimization problem with non-commuting variables, i.e. $x = (x_1, ..., x_n)$ are not real numbers, but non-commuting variables with $x_i x_j - x_j x_i \neq 0$. This type of problems that encounters non-commuting variables is common in quantum physics, where the basic objects are operators and matrices which are in general non-commuting.

To elucidate the problem, we let $\mathcal{H}$ be the Hilbert space, $X = \{X_i\}_{i=1}^n$ be a set of bounded operators on $\mathcal{H}$ and $\phi \in \mathcal{H}$ be a state. Given the polynomial functions $f, g_i : \mathbb{R}^n \to \mathbb{R}$, we define the operators $f(X)$ and $g_i(X)$ by substituting x with X. A general non-commuting optimization problems is:

$$\begin{aligned} p^\star \quad = \quad & \min_{(\mathcal{H},X,\phi)} \langle \phi, h(X)\phi \rangle \\ & \text{s.t.} \quad g_i(X) \geq 0 \qquad\qquad i \in \{1, 2, ..., m\} \\ & \qquad \langle \phi, \phi \rangle = 1 \end{aligned} \tag{1}$$

where $g_i(X) \geq 0$ means that the operators $g_i(X)$ are positive semi-definite. It's worth to mention that the constraints guarantee that the non-commuting relations of X are satisfied. For example, if we want to enforce the non-commuting relation $x_1 x_2 + x_2 x_1 = 1$, then we should add two constraints $g_1(x) = x_1 x_2 + x_2 x_1 - 1 \geq 0$ and $g_2(x) = -x_1 x_2 - x_2 x_1 + 1 \geq 0$.

We are looking for the combination $(\mathcal{H}, X, \phi)$ that minimize $p = \langle\phi, h(X)\phi\rangle$ under the constraints. It's important to note the the dimension of the Hillber space is not fixed.

Usually, the Hilbert space $\mathcal{H}$ is a vector space over complex number field, i.e. $\mathcal{H} = \mathbb{C}^N$. Then $|\phi\rangle$ can be understood as a vector and $\langle\phi|$ is its Hermition conjugate, i.e. $|\phi\rangle = v_\phi \in \mathbb{C}^N, \langle\phi| = v_\phi^\dagger$. Similarly the operators X are matrices acting on $\mathcal{H}$, in other words, $X \in \mathbb{C}^{N\times N}$. Usually, the strategy is first constructing a vector space $\mathcal{H}$ and the corresponding matrices X. The matrices X should satisfy the non-commuting relations given by the constraints . Next, we need to find the vector $|\phi\rangle \in \mathcal{H}$, such that $\langle\phi||h(X)\phi\rangle = v_\phi^\dagger h(X) v_\phi$ is minimized. However, the dimension of the vector space can be as large as 2^L where L is the system size. The large size of the vector space makes the problem difficult to solve. However, we will show that, we can map the problem to a semi-definite programming(SDP) problem[1]. After truncating the problem, we can solve the optimization problem efficiently.

We outline the structure of this report as follows. In section 2, we introduce some notations of polynomials and representations. In section 3, we illustrate the mapping between non-commuting optimization problem and SDP problem. In section 4, we apply our mapping to a specific quantum many-body system, Fermi Hubbard model. In section 5,6,7,8, we demonstrate the mapping between Fermi Hubbard model and SDP problem in detail. In section 9, we describe the primal-dual interior point method which has been used to solve the SDP problem. In section 10, we discuss the results of the Fermi Hubbard model. In section 11, we summary our results.

2 Notations

We consider the polynomial rings $R = \mathbb{C}[x, x^\dagger]$ of polynomials the $2n$ non-commuting variables $\{x_1, ..., x_n, x_{n+1} = x_1^\dagger, ..., x_{2n} = x_n^\dagger\}$. $\dagger$ is Hermitian conjugation . We let R_d denote the subring of R:

$$R_d := \{f \in \mathbb{C}[x, x^*] | f \text{ is polynomials with degree smaller then } d\}.$$

For example, $R_0 = \mathbb{C}$, $R_1 = \{\sum_i (a_i x_i + b_i x_i^\dagger) | a_i, b_i \in \mathbb{C}\}$.

We then define the monomial rings as

$$W_d = \{w \in R | x \text{ is monomial with degree smaller then } d\}$$

For example, $W_0 = \{1\}$, $W_1 = \{x_1, x_2, ..., x_1^\dagger, x_2^\dagger, ...\}$, $W_2 = \{x_1x_2, x_2x_1, ...\}$.

We say a sequence of complex numbers $\{r_w \in \mathbb{C}\}_{w\in W_d}$ with index in W_d admits a representation, if there exists a vector space $\mathcal{H} = \mathbb{C}^N$, a vector $v_\phi \in \mathbb{C}^N$ and a sequence of matrices $\{X_i \in \mathbb{C}^{N\times N}\}_{i=1}^n$, such that

$$r_w = v_\phi^\dagger w(X) v_\phi \quad \forall w \in W_d \tag{2}$$

3 Semi-definite programming

With the notations introduced in the previous section, we can construct the semi-definite programming(SDP) problem that is equivalent to the problem in Eq. 1. . First, any polynomial functions of X can be expanded in terms of $w(X)$. In other word, for any $f(X)$ a function

$$f(X) = \sum_w f_w w(X)$$

Next, we assume $\{r_w\}_{w\in W_d}$ admits a representation. Then we have:

$$v_\phi^\dagger h(X) v_\phi = \sum_w h_w r_w \tag{3}$$

For the constraints, we notice that

$$\begin{aligned} & g(X) \geq 0 \\ \Leftrightarrow & v^\dagger g(X) v \geq 0 \quad \forall v \in \mathcal{H} \\ \Leftrightarrow & v_\phi^\dagger f(X)^\dagger g(X) f(X) v_\phi \geq 0 \quad \forall f(X) \in \mathbb{C}[x, x^*] \end{aligned}$$

Since any element $f(X)$ in $R = \mathbb{C}[x, x^*]$ is a superposition of the monomials in W_∞. So the above constraint is equivalent to

$$\begin{aligned} & M_g \geq 0 \\ & (M_g)_{w_1,w_2} := v_\phi^\dagger w_1(X)^\dagger g(X) w_2(X) v_\phi = \sum_w g_w r_{w_1^\dagger w w_2} \quad , \quad w_1, w_2 \in W_D \end{aligned} \tag{4}$$

where M_g is matrix with index taking value in W_∞. $M_g \geq 0$ means the matrix M_g is semi-definite.

Similarly, from the normalization condition $v_\phi^\dagger v_\phi = 1$, we have

$$\begin{aligned} & v_\phi^\dagger v_\phi = 1 \\ \Rightarrow & v_\phi^\dagger f(X)^\dagger f(X) v_\phi \geq 0 \quad \forall f(X) \in \mathbb{C}[x, x^\dagger] \\ \Rightarrow & M_1 \geq 0 \quad , \quad (M_1)_{w_1,w_2} := v_\phi^\dagger w_1(X)^\dagger w_2(X) v_\phi = r_{w_1^\dagger w_2} \quad , \quad w_1, w_2 \in W_D \end{aligned} \tag{5}$$

Then combined Eq.3, Eq.4 and Eq.5, we can reformulate the problem in 7 as

$$\begin{aligned} p^\star &= \min_{\{r_w\}} \sum_w h_w r_w \\ &\text{s.t.} \quad M_{g_i} \geq 0 \qquad i \in \{1, 2, ..., m\} \\ &\qquad M_1 \geq 0 \end{aligned} \tag{6}$$

with M_{g_i}, M_1 defined in Eq.4 and Eq.5. And ≥ 0 denotes the matrix are semi-definite. Here we only consider monomials w in W_d with some fixed number d. When d is larger than the upper limit of the problem D, problem in 7 and problem in 6 are equivalent. However, the value of D is usually very large and makes the problem 6 intractable. In practice, we can set d to be much smaller then D and solve the problem approximately.

4 Application to quantum many-body system

As we mentioned before, the non-commuting relation naturally arises in quantum physics. In this section, we map a quantum many-body problem to a SDP problem. Here, we consider the Fermi Hubbard model. The primal problem is.

$$\begin{aligned} p^\star &= \min_{v_\phi \in \mathcal{H}} v_\phi^\dagger h(C, C^\dagger) v_\phi \\ &\qquad g_i(C, C^\dagger) = 0 \end{aligned} \tag{7}$$

where

$$\begin{aligned} h(C, C^\dagger) &= -2t \sum_{i=0}^{L-1} (C_i^\dagger C_{i+1} + C_{i+1}^\dagger X_i) + U \sum_{i=0}^{L-1} C_i^\dagger C_i C_{i+1}^\dagger C_{i+1} \\ g_1(C, C^\dagger) &= C_i C_j^\dagger + C_j^\dagger C_i - \delta_{i,j} \quad , \quad 1 \leq i, j \leq L \\ g_2(C, C^\dagger) &= C_i^\dagger C_j^\dagger + C_j^\dagger C_i^\dagger \quad , \quad 1 \leq i, j \leq L \\ g_3(C, C^\dagger) &= C_i C_j + C_j C_i \quad , \quad 1 \leq i, j \leq L \end{aligned} \tag{8}$$

and $C_i, C_i^\dagger$ are the non-commuting variables. $\delta_{i,j}$ is the Dirac-delta function for the discrete variables:

$$\begin{aligned} \delta_{i,j} &= 0 \quad \text{if } i \neq j \\ \delta_{i,j} &= 1 \quad \text{if } i = j \end{aligned} \tag{9}$$

This model describes the behaviors of particles on the lattice. t is the hopping amplitudes of the particles, U is the repulsion interactions between particles and L is the lattice size.

The strategy of solving this system is first transform the problem to the standard SDP problem and solving the SDP problem by interior-point method. In the following section, we fill formulate the SDP problem step by step.

5 Primal variables and correlation functions

We first introduce the primal variables. As we discussed in section 3, our primal variables are $\{r_w\}_{w\in W_d}$ defined as

$$r_w := v_\phi^\dagger w(C, C^\dagger) v_\phi$$

with $w(C, C^\dagger)$ some monomials of $C, C^\dagger$. It is more convenience to study the Fourier series of $C, C^\dagger$:

$$\begin{aligned} a_p &= \frac{1}{L}\sum_{i=0}^{L-1} C_i e^{-i\frac{2\pi pi}{L}} \\ a_p^\dagger &= \frac{1}{L}\sum_{i=0}^{L-1} C_i^\dagger e^{i\frac{2\pi pi}{L}} \end{aligned} \tag{10}$$

We truncate the monomials to degree 2 and consider the following primal variables:

$$\begin{aligned} Q^1_{p,q} &:= v_\phi^\dagger a_p^\dagger a_q v \\ Q^2_{p,q} &:= v_\phi^\dagger a_p a_q^\dagger v \\ Q^3_{p,q} &:= v_\phi^\dagger a_p a_q v \\ Q^4_{p,q} &:= v_\phi^\dagger a_p^\dagger a_q^\dagger v \end{aligned} \tag{11}$$

and

$$
\begin{aligned}
F^1_{pq,mn} &= v_\phi^\dagger a_p^\dagger a_q^\dagger a_n a_m v_\phi \\
F^2_{pq,mn} &= v_\phi^\dagger a_p a_q a_n^\dagger a_m^\dagger v_\phi \\
F^3_{pq,mn} &= \begin{bmatrix} \left(v_\phi^\dagger a_p^\dagger a_q a_n^\dagger a_m v_\phi\right) & \left(v_\phi^\dagger a_p^\dagger a_q a_{m-L} a_{n-L}^\dagger v_\phi\right) \\ \left(v_\phi^\dagger a_{q-L} a_{p-L}^\dagger a_n^\dagger a_m v_\phi\right) & \left(v_\phi^\dagger a_{q-L} a_{p-L}^\dagger a_{m-L} a_{n-L}^\dagger v_\phi\right) \end{bmatrix} \\
F^4_{pq,mn} &= v_\phi^\dagger a_p^\dagger a_q^\dagger a_n^\dagger a_m v_\phi \\
F^5_{pq,mn} &= v_\phi^\dagger a_p^\dagger a_q a_n^\dagger a_m^\dagger v_\phi \\
F^6_{pq,mn} &= v_\phi^\dagger a_p a_q a_n^\dagger a_m v_\phi \\
F^7_{pq,mn} &= v_\phi^\dagger a_p^\dagger a_q a_n a_m v_\phi \\
F^8_{pq,mn} &= v_\phi^\dagger a_p a_q a_n a_m v_\phi \\
F^9_{pq,mn} &= v_\phi^\dagger a_p^\dagger a_q^\dagger a_n^\dagger a_m^\dagger v_\phi
\end{aligned}
\tag{12}
$$

Note that, both Q^α, F^γ are in matrices form. In addition, this matrices are called correlation functions in physics. We put all the matrix together to construct a block diagonalized matrix X:

$$
X := diag\{Q^1, Q^2, Q^3, Q^4, F^1, F^2, F^3, F^4, F^5, F^6, F^7, F^8, F^9\} \tag{13}
$$

We then introduce a matrix basis $\{B_i\}$ see Append A, such that

$$
\begin{aligned}
X &= \sum_i x_i B_i \\
x &= mat(X)
\end{aligned}
\tag{14}
$$

where $x = vec(X)$ is the primal variables in the vector form.

6 Symmetry of the system

Usually, the system we consider has certain symmetry. We can take the advantage of this symmetry to accelerate the algorithm and increase the accuracy. We use a matrix T to

represent a symmetry operation. A T transformation is defined as $a_i \to a'_i = T_{ij}a_j$. We say the system is invariant under T transformation or has T symmetry if and only if

$$\begin{aligned} h(a') &= h(a) \\ g(a') &= g(a) \end{aligned} \tag{15}$$

where h, g are the given functions defined in Eq. 7 and Eq. 8. Here we replace C with its Fourier series a. In general, we assume the system doesn't spontaneously break the symmetry, i.e. the correlation functions(primal variables F, Q) are also invariant under the transformation T. This invariant properties will simplify the problem. Here we consider two types of symmetry of the Fermi Hubbard model, named translational symmetry and $U(1)$ symmetry.

$U(1)$ **symmetry:** The correlation functions (Q, F) are invariant under transformation $a_p \to e^{i\phi}a_p, a_p^\dagger \to e^{-i\phi}a_p^\dagger$, i.e.

$$\begin{aligned} v_\phi^\dagger a_p a_q v &= v_\phi^\dagger(e^{i\phi})a_p(e^{i\phi})a_q v \\ \Rightarrow v_\phi^\dagger a_p a_q v &= 0 \\ \Rightarrow Q^3 &= 0 \end{aligned} \tag{16}$$

Similarly,we have $Q^4 = 0, F^4 = F^5 = F^6 = F^8 = F^9 = F^9 = 0$.

Translational symmetry: The correlation functions (Q, F) are invariant under transformation $a_p \to e^{i\frac{2\pi p}{L}}a_p, a_p^\dagger \to e^{-i\frac{2\pi p}{L}}a_p^\dagger$, i.e.

$$\begin{aligned} v_\phi^\dagger a_p^\dagger a_q v &= e^{-i\frac{2\pi(p-q)}{L}} v_\phi^\dagger a_p^\dagger a_q v \\ \Rightarrow v_\phi^\dagger a_p^\dagger a_q v &\neq 0 \quad \text{iff } p = q \end{aligned} \tag{17}$$

Similarly, we have

$$\begin{aligned} Q^1_{p,q} &\propto \delta_{p,q} \\ Q^2_{p,q} &\propto \delta_{p,q} \\ F^1_{pq,mn} &\propto \delta_{p+q,n+m} \\ F^2_{pq,mn} &\propto \delta_{p+q,n+m} \\ F^3_{pq,mn} &\propto \delta_{p-q,n-m} \end{aligned} \tag{18}$$

7 Constraints and non-commuting relations

In this section, we construct the constraint equations of the Fermi Hubbard mode from the non-commuting relations. The non-commuting constraints in the Fourier series a is

$$\begin{aligned}
&a_p^\dagger a_q + a_q^\dagger a_p = \delta_{pq} \\
\Leftrightarrow &v^\dagger (a_p^\dagger a_q + a_q^\dagger a_p - \delta_{p,q}) v = 0 \quad \forall v \in \mathcal{H} \\
\Leftrightarrow &v_\phi^\dagger f(X)^\dagger (a_p^\dagger a_q + a_q^\dagger a_p - \delta_{p,q}) f(X) v_\phi = 0 \quad \forall f(X) \in \mathbb{C}[X, X^\dagger]
\end{aligned}$$

where v_ϕ is a given vector. As we mentioned before, we truncate the polynomial rings $\mathbb{C}[X, X^\dagger]$ by only considering the polynomials with degree smaller than 2. Then the constraints we have are

$$\begin{aligned}
&v_\phi^\dagger (a_p^\dagger a_q + a_q^\dagger a_p - \delta_{p,q}) v_\phi = 0 \\
&v_\phi^\dagger a_m (a_p^\dagger a_q + a_q^\dagger a_p - \delta_{p,q}) a_n v_\phi = 0 \\
&v_\phi^\dagger a_m (a_p^\dagger a_q + a_q^\dagger a_p - \delta_{p,q}) a_n v_\phi = 0 \\
&v_\phi^\dagger a_m^\dagger (a_p^\dagger a_q + a_q^\dagger a_p - \delta_{p,q}) a_n v_\phi = 0 \\
&v_\phi^\dagger a_m (a_p^\dagger a_q + a_q^\dagger a_p - \delta_{p,q}) a_n^\dagger v_\phi = 0 \\
&v_\phi^\dagger a_m^\dagger (a_p^\dagger a_q + a_q^\dagger a_p - \delta_{p,q}) a_n^\dagger v_\phi = 0
\end{aligned}$$

In terms of correlation functions we have F, Q

$$\begin{aligned}
&Q^1_{p,q} + Q^2_{q,p} = \delta_{p,q} \\
&F^1_{pq,mn} = -\delta_{p,n}\delta_{q,m} + \delta_{q,m} Q^1_{p,m} + \delta_{p,n} Q^2_{m,q} + \delta_{q,m} Q^2_{n,p} + \delta_{p,m} Q^2_{n,q} - F^2_{nm,qp} \\
&F^3_{pq,mn} = \delta_{q,n} Q^1_{p,m} - F^1_{pn,mq} \quad , \quad 1 \le p, q, m, n \le L \\
&F^3_{pq,mn} = \delta_{p,m} Q^2_{q-L,n-L} - F^1_{(q-L)(m-L),(n-L)(q-L)} \quad , \quad L < p, q, m, n \le 2L \\
&F^3_{pq,mn} = \delta_{n,m} Q^1_{p,q} - F^3_{pq,(n-L)(m-L)} \quad , \quad 1 \le p, q \le L < m, n \le 2L
\end{aligned} \tag{19}$$

8 Equivalent SDP problem

Now, we are in the position of mapping the quantum many-body problem to a standard SDP problem. For the objective function, we have (same as Eq. 3)

$$v_\phi^\dagger h(C, C^\dagger) v_\phi = h^T x$$

where x is the primal variables defined in Eq. 14.

The semi-definite condition in Eq. 4 can be written as:

$$mat(x) = X = \sum_i x_i B_i \geq 0$$

where ≥ 0 denotes the matrix is semi-definite.

For the constraints, we put the constraints from symmetry (Eq. 16, Eq. 18) and non-commuting relation (Eq. 19) together and rewrite them in a closed form:

$$Ax = b$$

Combining all the equations above, we have the SDP problem

$$\begin{aligned} \min_x \quad & h^T x \\ s.t. \quad & \sum_i x_i B_i \text{ is semi-definite} \\ & Ax = b \end{aligned} \tag{20}$$

9 Primal-dual interior-point method, logarithmic barrier function and Nesterov-Todd scaling

In this section, we demons how to solve the solve the primal problem in 20 by primal-dual interior-point method with logarithmic barrier function and Nesterov-Todd scaling. The Lagrangian with logarithmic barrier is

$$L = h^T x + y^T (Ax - b) - \mu \log[\det(\sum x_i B_i)] \tag{21}$$

The KKT equations are:

$$0 = \nabla_x L = h + A^T y - \mu Tr[(\sum x_i B_i)^{-1} B]$$
$$0 = \nabla_y L = Ax - b$$

We let $z_i = \mu Tr[(\sum x_j B_j)^{-1} B_i]$, $X = mat(x) = \sum_i x_i B_i$ and $Z = mat(z) = \sum_i z_i B_i$. We obtain the new KKT equations

$$h + A^T y - z = 0$$
$$Ax - b = 0$$
$$XZ = \mu \mathbb{I} \tag{22}$$

Following the standard procedure, the interior algorithm with Nesterov-Todd scaling is [2][3]:

Step 1. Evaluate the residuals and gap

Calculate r_x, r_y, r_z, μ

$$r_x = A^T y - z + h$$
$$r_y = Ax - b$$
$$r_z = \mu(vec(\mathbb{I}))$$

where $\mu = \frac{1}{dim(x)} x^T z$. Stop if converge.

Step 2. Calculate scaling matrix

Find Nesterov-Todd scaling matrix W that satisfies:

$$W^{-T} x = W z$$

and calculate $\lambda = W^{-T} x$

We discuss the the algorithm of calculating scaling matrix W in Appendix B

Step 3. Affine direction

We solve the following Newton's equation to obatin the affine direction

$$\begin{bmatrix} 0 & A^T & -\mathbb{I} \\ A & 0 & 0 \\ \mathbb{I} & 0 & W^T W \end{bmatrix} \cdot \begin{bmatrix} \Delta x_a \\ \Delta y_a \\ \Delta z_a \end{bmatrix} = - \begin{bmatrix} A^T y - z + h \\ Ax - b \\ x \end{bmatrix} \tag{23}$$

Step 4. Step size

$$\alpha = \sup\{\alpha \in [0,1] | mat(x + \alpha\Delta x_a) > 0, mat(z + \alpha\Delta z_a) > 0\}$$
$$\rho = \frac{(x + \alpha\Delta x_a)^T (z + \alpha\Delta z_a)}{x^T z}$$
$$\sigma = \max\{0, \min\{1, \rho\}\}^3$$

Step 5. Combined direction

Find the combined direction through

$$\begin{bmatrix} 0 & A^T & -\mathbb{I} \\ A & 0 & 0 \\ \mathbb{I} & 0 & W^T W \end{bmatrix} \cdot \begin{bmatrix} \Delta x_a \\ \Delta y_a \\ \Delta z_a \end{bmatrix} = - \begin{bmatrix} (1-\sigma)(A^T y - z + h) \\ (1-\sigma)(Ax - b) \\ x - W^T \lambda \diamond (\sigma\mu(vec(\mathbb{I}))) \end{bmatrix} \tag{24}$$

$\lambda \diamond (u)$ is defined through $\lambda \diamond (u) = vec\Big(mat(u) \odot \Gamma\Big)$, where $\Gamma_{ij} = \frac{2}{mat(\lambda)_{ii} + mat(\lambda)_{jj}}$ and $\odot$ denotes the Hadamard (element-wise) product.

Step 6. Update

$$\alpha' = \sup\{\alpha \in [0,1] | mat(x + \alpha\Delta x) > 0, mat(z + \alpha\Delta z) > 0\}$$

Update

$$x \to x + 0.99\alpha'\Delta x$$
$$z \to z + 0.99\alpha'\Delta z$$

Go to step 1.

The Newton's equations appeared in step 3 and step 5 can be solved by the method in Appendix C.

10 Results

We use the algorithm demonstrated in the previous section to solve the Fermi Hubbard model in Eq. 7.

We first consider the case of $U = 0, t = 1.0$. In this case, we can analytically solve the problem. We compared the result obtained from numerical calculations and from analytical calculation. We show the result in the table 1

Lattice size	Iteration number	Time/s	p(Numerical)	p(Analytical)
2	6	0.124	-1.999995	-2.0
4	7	5.392	-1.999995	-2.0
6	8	62.310	-3.999879	-4.0
8	8	398.937	-4.828215	-4.828

Table 1: $t = 1.0, U = 0.0$

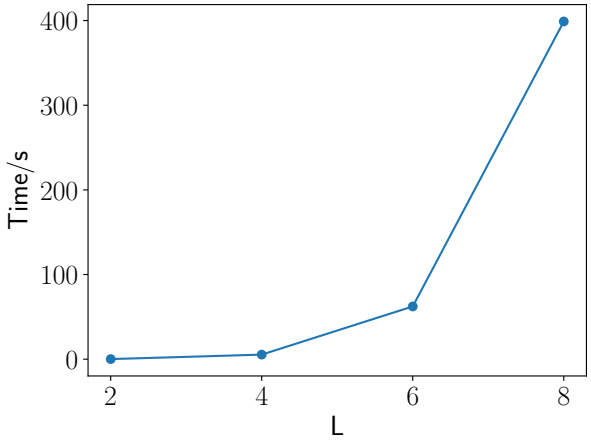

Figure 1: Calculation time v.s. system size L.

In the table 1, L denotes the system size. The matrix we consider in the optimization problem has the size of order $O(L^2)$. In addition, we plot the calculation time at different lattice size in Fig 1.

In table 1, $p = h^T x$ is value we want to minimized. By comparing the results from numerical calculation and analytical calculation, we conclude that our algorithm can efficiently solve the problem.

Now, we turn to the case with $U \neq 0$. Solving this type of model is one of the central problems in the area of condensed matter physic. There isn't any efficient numerical or analytical method to exactly solve this type of problem. We don't aim to find the exact solution, but try to approximately solve it and find the approximate correlation functions (primal variables). We demonstrate the results in the table 2.

U	Iteration number	Time/s	p
0.2	22	179.4	-2.448420
0.5	20	155.4	-1.996753
1.0	16	125.5	-1.994281
1.5	16	129.0	-1.991024

Table 2: $L = 6, t = 1.0$

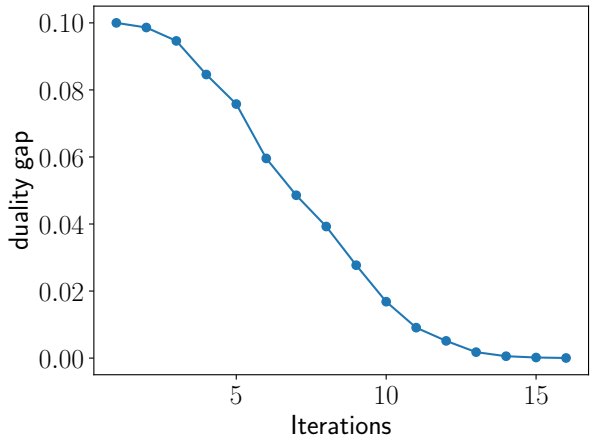

Figure 2: Duality gap v.s. iteration numbers for $L = 6, t = 1.0, U = 1.5$

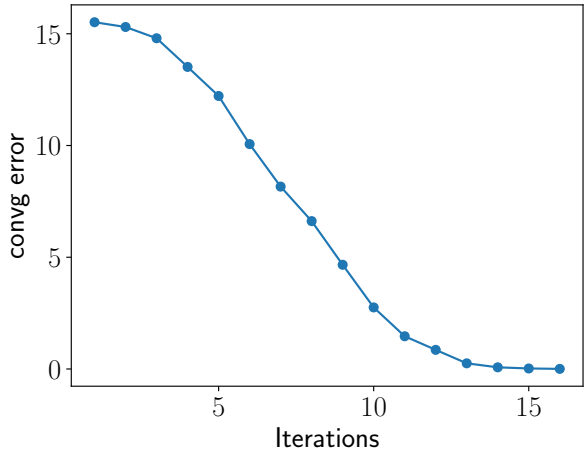

Figure 3: Residues v.s. iteration numbers for $L = 6, t = 1.0, U = 1.5$

To show the duality gap closes and the result converges, we plot the evolution of gap and convergence errors with iteration number for $L = 6, t = 1.0, U = 1.5$, in Fig 2 and Fig 3.

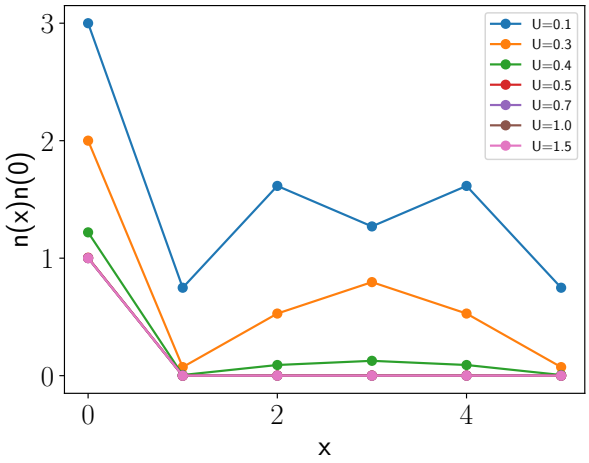

Figure 4: Density-density correlation function at $L = 6, t = 1.0$

We also plot the density-density correlation functions in the Fig 4 where the density-density correlation function is defined as

$$\begin{aligned} n(x)n(0) &= \frac{1}{L}\sum_i v_\phi^\dagger C_i^\dagger C_i C_{i+x}^\dagger C_{i+x} v_\phi \\ &= \frac{1}{L}\sum_{q,k,p} \cos(\frac{2\pi}{L}px)(v_\phi^\dagger a_q^\dagger a_{q-p} a_{k-p}^\dagger a_k v_\phi) \end{aligned}$$

From the Fig 4, we find the density-density correlation function only takes finite value at $x = 0$ when $U > 0.3$ which represents the particles are localized and the system is in the charge ordering phase. Once $U < 0.3$, the density-density correlation functions will take finite value at any x and the system is in a Fermi liquid phase. In summary, we find a quantum phase transition from a charge ordering phase to a Fermi liquid phase by decreasing U.

11 Summary and discussion

We discussed the optimization problem with non-commuting variables. The non-commuting properties naturally arise in the quantum many-body system. We transform the original quantum many-body problem to an equivalent SDP problem. We truncate the problem by only considering a subring of the polynnomial ring $\mathbb{C}[x, x^\dagger]$. The truncation allows us to solving the problem within polynomial time. We use the primal-dual interior point method

with logarithmic barrier and Nesterov-Todd scaling to solve the problem. In addition, we take the advantage of the symmetry property to accelerate the algorithm.

In particular, we study the Fermi Hubbard model. We first consider the case of $U = 0$ and compare the numerical and analytical results. It turns out that our algorithm are quite efficient. Next, we turn on the repulsive interaction $U \neq 0$. We identified a quantum phase transition between a charge ordered phase and a Fermi liquid phase by tuning U.

Our method can be applied to the other quantum many body systems, such as Heisenberg model, Bose Hubbard model and etc.. Also, our method has the ability to study the system with large system size. Using the super computer resource, we should be able to solve the problem with size $L \sim 100$ which are almost ten times larger than the other method, such as DMRG, quantum Monte Carlo.

12 Appendix A: matrix basis and vectorization

We construct the matrix basis we discussed in the main text. In the SDP problem, the primal variables are in the matrices form, but usually we can vectorize the matrix and bring it to the vector form by introducing matrix basis. We say the set of matrices $\{B_i\}$ is a basis if it satisfies:

$$\begin{aligned} Tr[B_i B_j] &= \delta_{ij} \\ B_i^\dagger &= B_i \end{aligned}$$

Here the Hermitina condition comes from the fact that all the matrices we consider in the SDP are Hermitian. We define the vectorization of the matrix X as following:

$$\begin{aligned} x &= vec(X) \\ X &= \sum_i x_i B_i \end{aligned}$$

Without loss of generality, we assume $X \in \mathbb{C}^{N\times N}$ with $X^\dagger X$. A natural way to introduce the basis is to take the unions of the generators of $SU(N)$ group and the identity matrix. For convenience, we take the linear combination of some generators and identity matrix and

show the basis $\{B^1_\alpha\}_{1\le\alpha\le N}\cup\{B^2_{\alpha\beta}\}_{1\le\alpha<\gamma\le N}\cup\{B^3_{\alpha\beta}\}_{1\le\alpha<\gamma\le N}$as follows:

$$
\begin{aligned}
(B^1_\alpha)_{ij} &= \delta_{i,j} \quad , \quad 1\le\alpha\le N \\
(B^2_{\alpha\gamma})_{ij} &= \frac{1}{\sqrt{2}}(\delta_{i,\alpha}\delta_{j,\beta}+\delta_{i,\beta}\delta_{j,\alpha}) \quad , \quad 1\le\alpha<\gamma\le N \\
(B^3_{\alpha\gamma})_{ij} &= \frac{i}{\sqrt{2}}(-\delta_{i,\alpha}\delta_{j,\beta}+\delta_{i,\beta}\delta_{j,\alpha}) \quad , \quad 1\le\alpha<\gamma\le N
\end{aligned}
$$

Then the relation between X and x is

$$
\begin{aligned}
mat(x) &= X \\
&= \begin{bmatrix}
x_1 & \frac{x_{N+1}-ix_{N+N(N-1)/2+1}}{\sqrt{2}} & \frac{x_{N+2}-ix_{N+N(N-1)/2+2}}{\sqrt{2}} & \dots \\
\frac{x_{N+1}+ix_{N+N(N-1)/2+1}}{\sqrt{2}} & x_2 & \dots & \dots \\
\frac{x_{N+2}+ix_{N+N(N-1)/2+2}}{\sqrt{2}} & \dots & \dots & \dots \\
\dots & \dots & \dots & x_N
\end{bmatrix}
\end{aligned}
$$

13 Appendix B: Nesterov-Todd scaling matrix

Given two matrix X and Z with $x=vec(X)$ and $z=vec(Z)$. A scaling matrix W is a linear transformation that preserve the cone and central path, i.e.

$$
\begin{aligned}
\tilde{x} = W^{-T}x \quad \tilde{z} = z \\
mat(x)\ge 0, mat(z)\ge 0, mat(x)\cdot mat(x)+mat(z)\cdot mat(x)=\mu\mathbb{I} \\
\Leftrightarrow mat(\tilde{x})\ge 0, mat(\tilde{z})\ge 0, mat(\tilde{x})\cdot mat(\tilde{z})+mat(\tilde{z})\cdot mat(\tilde{x})=\mu\mathbb{I}
\end{aligned}
$$

A Nesterov-Todd scaling point is characterized by a matrix W that satisfied $W^{-T}x=Wz$. The systematic way to construct the W is first doing Cholesky decomposition and SVD decomposition.

$$
\begin{aligned}
X &= L_1L_1^\dagger \\
Z &= L_2L_2^\dagger \\
L_2^\dagger L_1 &= U\Lambda V^\dagger \\
R &= L_1V\Lambda^{-\frac{1}{2}} = L_2^{-\dagger}U\Lambda^{\frac{1}{2}}
\end{aligned}
$$

Next, we calculate W from R by,

$$W_{ij}=Tr[R^\dagger B_jRB_i]$$

where $\{B_i\}$ is the matrix basis.

14 Appendix C: Solving Newton's equation

The interior-point method needs to solve the following Newton's equations

$$\begin{bmatrix} 0 & A^T & -\mathbb{I} \\ A & 0 & 0 \\ \mathbb{I} & 0 & W^TW \end{bmatrix} \cdot \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = - \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix}$$

The solution is

$$\begin{aligned} \Delta y &= \Big[AW^TWA\Big]^{-1}\Big[r_y - Ar_z - AW^TWr_x\Big] \\ \Delta z &= A^T\Delta y + r_x \\ \Delta x &= -W^TW\Delta z - r_z \end{aligned}$$

References

[1] Stefano Pironio, Miguel Navascues and Antonio Acin,
SIAM J. Optim. Volume 20, Issue 5, pp. 2157-2180 (2010)

[2] Stephen J. Wright,
Primal-Dual Interior-Point Methods(1987).

[3] Matthias Heinkenschloss,
Lecture notes(2019).