Bosonic and fermionic particle-preserving coherent states, their dynamics and applications

Transfer Report

Michal Horanský

July 31, 2025

Abstract

Coherent states-based methods have been previously developed and utilised for particle-preserving bosonic systems, such as the Bose-Hubbard model. This work is divided into two parts: Firstly, a new method is proposed for the bosonic systems, where the basis coherent states are propagated separately to create a frozen dynamical basis, which lowers the time-complexity significantly. Second, an analogous method is proposed for particle-preserving fermionic systems, and particularly the electronic structure of a molecule, for which the goal of this method is to find the ground state. For this, multiple methods are outlined, and relevant properties of the fermionic coherent states are obtained.

Contents

Chapter 1

Background

To find the dynamics of a given system, one typically needs to solve the Schrodinger equation given by the system's Hamiltonian. For complex systems which can be found in nature, an analytic solution rarely exists. To solve this partial differential equation computationally, one needs to chose a basis for the Hilbert space and, since it typically is infinite or even continuous, sample it to obtain a basis sample onto which the initial state is decomposed and on which it is propagated. This sampling choice is always limiting and must be carefully justified, as omitting elements of the Hilbert space may render the analysis inaccurate.

Coherent states are particular states in the Hilbert space which follow classical trajectories, remain coherent, and minimise position-momentum uncertainty. Choosing them as the basis, a very small sample is typically suitable to capture the physics of a wide variety of systems and starting conditions [13]. This is an example of a semi-classical approximation.

1.1 Utility and scope of semi-classical methods

Semi-classical methods have long been used in computational quantum dynamics, mainly in systems which can be partitioned into multiple sub-systems, some to be treated classically, some to be treated with the machinery of quantum mechanics [2]. A common example is the molecule, for which, based on the Born-Oppenheimer approximation, the nuclei may be treated as if following classical trajectories, and the electronic structure following quantum dynamics.

More recently, systems which are "fully quantum" started being approached with classical approach in mind, as quantum dynamics may be formulated on top of classical behaviour embedded in a quantum-mechanical system [3].

An example of this approach is the method of coherent-state basis trajectories, which lies at the heart of my research. As elaborated on below in Sec. 1.2, states which follow classical-like trajectories and which remain localised in phase-space are embedded in every quantum system, and are inextricably linked to second-quantisation of position and momentum operators. This second-quantisation expresses the state of a system by the level of discrete excitations in an array of modes. As such, we have chosen three particular physical systems as subjects to the methods developed in this work: two analysing bosonic dynamics, and one analysing fermionic dynamics.

1.1.1 Bosonic system 1: Bose-Hubbard model

The Bose-Hubbard model describes the physics of bosonic particles trapped on a lattice, which is for our particular analysis chosen to be one-dimensional. The modes occupied by a fixed total number of bosons correspond to sites on this lattice, and the relevant components of the Hamiltonian are on-site energy U and hopping energy J, respectively. The model, in its general formulation, also incorporates an external potential being applied on the lattice, which we, however, ignore. As such, the full Hamiltonian becomes:

$$\hat{H} = -J \sum_{i=1}^{M-1} \left(\hat{b}_i^\dagger \hat{b}_{i+1} + \hat{b}_{i+1}^\dagger \hat{b}_i \right) + \frac{U}{2} \sum_{i=1}^{M} \hat{b}_i^{\dagger\,2}\, \hat{b}_i^2 \tag{1.1}$$

This system has previously been analysed using coherent-states-based methods by Qiao and Grossmann [13]. As such, the machinery we propose for bosonic systems shall be compared to this work as a benchmark.

The method used by Qiao and Grossmann is as follows: a basis of generalised field coherent states is constructed mathematically and then sampled around the initial wavestate. The elements of this basis are parametrised by complex vectors $\vec{\xi}_a$. The wavestate is decomposed into this basis like so:

$$|\Psi\rangle = \sum_a^N A_a \left| \vec{\xi}_a \right\rangle \tag{1.2}$$

where $A_a, \vec{\xi}_a$ are taken as dynamical coordinates. The equations of motion for these coordinates are expressed in a fully variational manner from the Schrodinger Lagrangian, and solved numerically as an initial-value problem for several particular choices of system parameters; importantly, for two-mode and three-mode lattices.

The method we propose as a faster alternative is as follows: after sampling the coherent state basis, each element of this basis is propagated along its classical-like trajectory. Then, the wavestate is decomposed onto the basis, and its decomposition coefficients are propagated on this basis, which is not constant in time, but follows pre-determined trajectories. As such, this is not a fully variational method, and the consideration for quantum effects is fully engrossed in the propagation of decomposition coefficients alone. We propose that for the type of initial wavestates as analysed by Qiao and Grossmann, which are either pure coherent states or superpositions of a handful of coherent states, this new method will converge easily with small basis samples and, by construction, entails drastically smaller time-complexity.

1.1.2 Bosonic system 2: Displaced harmonic trap

Another bosonic system is taken from the work of Green and Shalashilin [14], who studied a closed system of particles in a trap comprised of a displaced harmonic potential. This second application is of interest to us as a benchmark because it can be formulated with position and momentum operators, and the second-quantised formulation is in fact an alternative which requires a large number of modes (25–27) to converge. As such, the ability to solve this system with our new coherent-states-based method would prove its scalability for bosonic systems with a large number of modes.

The method utilised by Green and Shalashilin is similar to the one employed by us, but it is based on a less general formulation of coherent states, which are frozen gaussians in the phase space. As such, it serves as a usedul benchmark in its own right to measure the trade-off between the accuracy of classical-like trajectories and added complexity of the more general formulation used in this work.

1.1.3 Fermionic system: molecular electronic structure

An example of a fermionic system which preserves the total number of particles is the electronic structure of an inert molecule. This is a very important problem, as the ground state of the electronic structure is closely related to many chemical properties of molecules. The molecular orbitals of electrons can be well-approximated as linear combinations of the electron orbitals of the constituent atoms [4].

In this work, we propose a new method for obtaining the ground state, which works by restricting the Hilbert space to a small coherent-state subspace surrounding the lowest-configuration state, sampling this subspace, and manually diagonalising the Hamiltonian

expressed as a matrix on the basis sample. One method for this is to take the lowest-configuration state as a reference state in our coherent state construction, which means it in itself is a coherent state, and then propagating it along its classical-like trajectory and sampling it at fixed time-intervals. Using the overcompleteness of coherent states, we hope that even a sample of a size much smaller than the dimensionality of the full Hilbert space will be able to converge to the true ground state.

1.2 Coherent states: a hundred year-long history

The study of classical trajectories present in quantum dynamics dates back to Schrodinger's work in 1926, and remained an active area of study ever since. Multiple generalisations were constructed for the initial concept of coherent states, and the specific properties of coherent states of many particular systems were found and subsequently utilised in semi-classical methods. In this section, I outline a brief history of the theoretical treatment of coherent states leading up to the framework used for the $SU(M)$ bosonic/fermionic coherent states.

1.2.1 Schrodinger: the harmonic oscillator

In 1926, Schrodinger formulated so-called "classical states" of the simple harmonic oscillator [5]. Consider the Hamiltonian of the simple harmonic oscillator

$$\hat{H} = \omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \tag{1.3}$$

where ω is the characteristic scale of the potential. Then consider the following superposition of energy eigenstates $|n\rangle$, characterised by a single complex parameter α:

$$|\alpha\rangle = N(\alpha) \sum_{i=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{1.4}$$

where $N(\alpha)$ is a normalisation factor. This state has the following properties:

1. The expected values of $\hat{x}$ and $\hat{p}$ are proportional to the real and imaginary components of α, respectively.

2. These expected values evolve according to the "classical" Hamiltonian equation, which can therefore be succintly written as

$$\dot{\alpha} = -i \frac{\partial H(\alpha, \alpha^*)}{\partial \alpha^*} \tag{1.5}$$

   where $H(\alpha, \alpha^*)$ is the classical simple harmonic oscillator Hamiltonian as given by the expected values of $\hat{x}, \hat{p}$.

3. The position and momentum uncertainties of $|\alpha\rangle$ are minimal: $\Delta x \Delta p = \frac{1}{2}$.

4. The state remains "coherent", i.e. as it evolves, the parameter $\alpha(t)$ changes, but the wave-state remains describable by the construction in Eq. 1.4.

The significance of these states is immediately obvious: although fundamentally non-classical, the Schrodinger equation with the simple harmonic oscillator Hamiltonian permits "wave-groups" (as dubbed by Schrodinger) whose expectation values not only follow non-trivial classical trajectories, but which remain compact (both in space and in momentum representation!) and coherent.

1.2.2 Glauber: field coherent states

In 1963, Glauber generalised Schrodinger's construction to the Hamiltonian which describes the interaction between an atomic system and an electromagnetic field, forming field coherent states[1] [6]. According to Glauber, the following three properties all lead to the construction of field coherent states [10, p. 869]:

1. The coherent state is an eigenstate of the annihilation operator

$$\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle \tag{1.6}$$

2. The coherent state is obtained by applying a displacement operator on the vacuum state

$$|\alpha\rangle = \hat{D}(\alpha)\,|0\rangle \quad \text{where} \quad \hat{D}(\alpha) = \exp\big(\alpha\hat{a}^\dagger - \alpha^*\hat{a}\big) \tag{1.7}$$

3. The coherent state is a state with minimal position-momentum uncertainty.

However, as remarked by Zhang, Feng, and Gilmore in [10], the third property is not sufficient for a unique construction.

Field coherent states showed that "classical states" are present in systems beyond the simple harmonic oscillator, but Glauber's construction still was not applicable to general Hamiltonians–specifically, it fails when the Hilbert space is finite, as no eigenstate of the annihilation operator exists! This is particularly relevant when studying systems which preserve total particle number, and thus further generalisation was required.

1.2.3 Perelomov and Gilmore: generalised field coherent states

Such a generalisation of Glauber's coherent states for arbitrary Hamiltonians was indeed found independently by Perelomov [8] and Gilmore [9] in 1972. This construction discards the approach via annihilation operator eigenstates, and instead fully generalises the approach through the displacement operator (which, in retrospect, fully recovers Schrodinger's classical states and Glauber's field coherent states). Perelomov and Gilmore's construction follows these steps:

1. Firstly, the transition operators $\hat{T}_i$ of the Hilbert space $\mathcal{H}$ are identified. These are operators which, given any state in $\mathcal{H}$, can reach any other state in $\mathcal{H}$ by finitely-many sequential applications to the initial state. These operators must form a Lie algebra. Also, the Hamiltonian has to be expressable as a function of the transition operators (not necessarily a linear function).

2. Secondly, a reference state $|\phi_0\rangle$ is chosen. This choice is arbitrary.

[1] Also called Glauber coherent states.

3. Thirdly, two Lie groups are identified: the dynamical group G, which is the Lie group associated with the Lie algebra formed by the transition operators $\hat{T}_i$, and its stability subgroup $\hat{H}$, which is the group of all elements in G which leave $|\phi_0\rangle$ invariant up to a phase. The quotient group G/H then contains elements $\hat{D(z_j)} = \exp\left(\sum_j z_j \hat{T}_j\right)$, where the sum over j contains those transition operators which are not in the stability group Lie algebra. Thus we obtain a map which assigns one coherent state to every element of the quotient group:

$$|z\rangle = N(z, z^*) \exp\left(\sum_j z_j \hat{T}_j\right) |\phi_0\rangle \tag{1.8}$$

where N is a normalisation function, which is necessary since $\hat{D}(z)$ is not in general unitary: the parameters z_j are complex and the transition operators are not required to be Hermitian.

These states posses many of the aforementioned properties, most importantly, classical-like trajectories, as will be discussed in Sec. 1.3.

1.3 Mathematical approach to CS-based methods

In this section, we firstly discuss the mathematical properties of generalised field coherent states and their ensembles, and then the particular construction of $SU(M)$ coherent states.

1.3.1 Topology of the CS parameter space

This subsection is a summary of relevant information as included by Viscondi, Grigolo, and de Aguiar in their review of generalised field coherent states [11].

When applying the quotient-space displacement operator on the reference state, the resulting states are no longer normalised. We will adopt the notation of Viscondi, Grigolo, and de Aguiar, and denote unnormalised elements of the Hilbert space as bras and kets with curly brackets. Then, the unnormalised coherent state is directly obtained as

$$|z\} = \hat{D}(z)\,|\phi_0\rangle \tag{1.9}$$

and a normalised coherent state is constructed by scaling by a normalisation factor:

$$|z\rangle = N(z^*, z)\,|z\} \quad \text{where} \quad N(z_a^*, z_b) = \{z_a|z_b\}^{-\frac{1}{2}} \tag{1.10}$$

The quotient space can be characterised by the following metric:

$$g(z_a^*, z_b) = \frac{\partial^2 \ln\{z_a^*|z_b\}}{\partial z_b \partial z_a^*} \tag{1.11}$$

Then, the coherent states remain coherent under time evolution, with the parameter z evolving according to a Hamiltonian equation on a curved manifold:

$$\dot{z}_i = -i\xi_{ij}^T(z^*, z)\frac{\partial H(z^*, z)}{\partial z_j^*} \tag{1.12}$$

where $H(z^*, z) = \langle z|\hat{H}|z\rangle$ is the effective Hamiltonian and $\xi(z^*, z)$ is the inverse of the quotient-space metric $g(z^*, z)$.

1.3.2 Fully variational equations of motion

Here we employ the paradigm discussed in previous sections which encompasses the heart of this kind of semi-classical approaximation: a limited sample of classical trajectories. We will approximate our Hilbert space by a discrete sample of N coherent states, each characterised by its parameters z_a, which evolve in time. Onto this basis sample, we decompose an arbitrary wavestate $|\Psi(t=0)\rangle$ whose dynamics we are interested in:

$$|\Psi(t=0)\rangle = \sum_{a=1}^{N} A_a(t=0)\,|z_a(t=0)\} \tag{1.13}$$

Note that I opt to use the unnormalised coherent states as the basis, rather than their normalised counterparts. This does not affect the physicality of the solution, which shall remain normalised, and it serves to simplify the equations of motion by omitting the normalisation factors $N(z_a)$.

Now, we can apply the Schrodinger Lagrangian $\hat{L} = i\frac{\mathrm{d}}{\mathrm{d}t} - \hat{H}$ onto our system of free coordinates (A_a, z_a) to obtain a system of first-order equations of motion, which can be computationally solved to obtain the trajectories of the basis, the decomposition coefficient evolution, and hence the evolution of the initial wavestate. This approach is a slight modification of the approach used by Qiao and Grossmann in [13].

Let us begin by obtaining the derivatives of unnormalised coherent states:

$$\frac{\partial}{\partial z_j}|z\} = \frac{\partial}{\partial z_j}\exp\Big(z_i\hat{T}_i\Big)|\phi_0\rangle = \hat{T}_j|z\} \tag{1.14}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}|z\} = \frac{\mathrm{d}}{\mathrm{d}t}\exp\Big(z_i\hat{T}_i\Big)|\phi_0\rangle = \dot{z}_i\hat{T}_i|z\} \tag{1.15}$$

Then the Lagrangian of the trial wavestate is

$$L = \sum_{m,n}^{N}\left(iA_m^*\dot{A}_n\{z_m|z_n\} + iA_m^*A_n\sum_i \dot{z}_{n,i}\{z_m|\hat{T}_i|z_n\} - \{z_m|\hat{H}|z_n\}\right)$$

and the equations of motion obtained by differentiating with respect to $A_m^*, z_{m,i}^*$, respectively, are

$$i\sum_n^N\left(\dot{A}_n\{z_m|z_n\} + A_n\sum_i\dot{z}_{n,i}\{z_m|\hat{T}_i|z_n\}\right) = \sum_n^N A_n\{z_m|\hat{H}|z_n\} \tag{1.16}$$

$$iA_m^*\sum_n^N\left(\dot{A}_n\{z_m|\hat{T}_i^\dagger|z_n\} + A_n\sum_j\dot{z}_{n,j}\{z_m|\hat{T}_i^\dagger\hat{T}_j|z_n\}\right) = A_m^*\sum_n^N A_n\{z_m|\hat{T}_i^\dagger\hat{H}|z_n\} \tag{1.17}$$

Note that for $N = 1$ (i.e. the wavestate is a pure coherent state), these equations are equivalent to Eq. 1.12.

It is helpful to express this in matrix form, with the coordinate vector defined as

$$\vec{Q} = \begin{pmatrix}\vec{Q}^A\\ \vec{Q}^z\end{pmatrix} \qquad \vec{Q}^A_m = A_m \qquad \vec{Q}^z = \begin{pmatrix}\vdots\\ \vec{Q}^{z,i}\\ \vdots\end{pmatrix} \qquad \vec{Q}^{z,i}_m = z_{m,i} \tag{1.18}$$

where the first-order matrix equation of motion becomes

$$i\Omega\frac{\mathrm{d}}{\mathrm{d}t}\vec{Q} = \vec{R} \quad \text{where} \quad \vec{R} = \begin{pmatrix}\vec{R}^A\\ \vec{R}^z\end{pmatrix} \qquad \vec{R}^z = \begin{pmatrix}\vdots\\ \vec{R}^{z,i}\\ \vdots\end{pmatrix} \tag{1.19}$$

and where

$$\Omega = \begin{pmatrix} \Omega^{AA} & \Omega^{A,1} & \dots & \Omega^{A,j} & \dots \\ \Omega^{1,A} & \Omega^{1,1} & \dots & \Omega^{1,j} & \dots \\ \vdots & \vdots & \ddots & \vdots & \\ \Omega^{i,A} & \Omega^{i,1} & \dots & \Omega^{ij} & \dots \\ \vdots & \vdots & & \vdots & \ddots \end{pmatrix} \tag{1.20}$$

The particular elements of the dynamical matrix Ω and the weight vector $\vec{R}$ are then identified as

$$\Omega^{AA}_{mn} = \{z_m|z_n\} \tag{1.21}$$

$$\Omega^{A,i}_{mn} = A_n \{z_m|\hat{T}_i|z_n\} \quad \text{and} \quad \Omega^{i,A} = (\Omega^{A,i})^\dagger \tag{1.22}$$

$$\Omega^{ij}_{mn} = A^*_m A_n \{z_m|\hat{T}^\dagger_i \hat{T}_j|z_n\} \tag{1.23}$$

$$R^A_m = \sum_n^N A_n \{z_m|\hat{H}|z_n\} \tag{1.24}$$

$$R^{z,i}_m = A^*_m \sum_n^N A_n \{z_m|\hat{T}^\dagger_i \hat{H}|z_n\} \tag{1.25}$$

Note that A^*_m could have been factored out of the second group of equations, but was deliberately left in to demonstrate that Ω is Hermitian.

This form of the equation of motion is especially useful, since it yields itself readily to primitive methods of numerical integration (the method of choice in this work is the Dormand-Prince method, see below). To propagate the wavestate by a finite time-difference Δt, one simply inverts the dynamical matrix Ω at the current time. Therefore the time complexity of this approach is limited by the evaluation of R and Ω and the inversion of Ω, both of which are polynomial in the basis size and the number of transition operators.

1.3.3 $SU(M)$ coherent states

Consider a system characterised by M modes and S particles, with both of these numbers remaining constant. The most general transition operator is then

$$\hat{T}_{ij} = \hat{a}^\dagger_i \hat{a}_j \tag{1.26}$$

where we remain agnostic about whether $\hat{a}$ denotes the bosonic or the fermionic annihilation operator. In either case, applying this operator on an arbitrary occupancy basis state within the configuration space yields either zero or another basis state within the configuration space. It can also be shown by applying the bosonic and fermionic commutation relations

that for both of these cases, the transition operators satisfy the following commutation relations:

$$\left[\hat{T}_{ij}, \hat{T}_{i'j'}\right] = \hat{T}_{ij'}\delta_{i'j} - \hat{T}_{i'j}\delta_{ij'} \tag{1.27}$$

Since the commutator is contained in the vector space spanned by $\hat{T}$, the transition operators form a Lie algebra.

On transformations of the dynamical Lie algebra

Note that a coherent state with parameter z is specified by the linear combination $z_i\hat{T}_i$. Hence, we can specify a new dynamical Lie algebra $\hat{T}'_i$ formed as an arbitrary complex linear combination of $\hat{T}_i$. These two Lie algebras correspond to the same coherent state space, with the parametric maps related by the same linear transformation.

Firstly, we choose to transform the mode-number operators $\hat{a}_i^\dagger\hat{a}_i$ like so:

$$\hat{S} = \sum_i^M \hat{a}_i^\dagger\hat{a}_i \tag{1.28}$$

$$\hat{H}_i = \hat{a}_{i+1}^\dagger\hat{a}_{i+1} - \hat{a}_i^\dagger\hat{a}_i, \qquad i = 1, 2 \dots M-1 \tag{1.29}$$

The operator $\hat{S}$ is uninteresting, because every state in the configuration space is its eigenstate with eigenvalue S, and therefore $\hat{S}$ is always in the stability subgroup. We can therefore ignore it.

The transformed Lie algebra constitutes the generators of real traceless $(M \times M)$ matrices, and therefore the full dynamical group is $U(1) \times SL(M, \mathbb{R})$. Ignoring the $U(1)$ direct product, as it is always fully contained in the stability subgroup, we may denote the dynamical group simply as $SL(M, \mathbb{R})$.

On the difference between $SU(M)$ and $SL(M, \mathbb{R})$

When transforming the basis $\hat{T}_{ij}$, one may choose to complexify it and form a set of Hermitian and anti-Hermitian operators ($\hat{T}_{ij} + \hat{T}_{ji}$ and $i(\hat{T}_{ij} - \hat{T}_{ji})$), respectively. This complexification is allowed in the sense that the resulting coherent state space is equivalent. Identifying the resulting operators with the generalised Gell-Mann matrices yields the dynamical group $SU(M)$ [15]. Although resulting in a different quotient group G/H, the coherent states are identical to our $SL(M, \mathbb{R})$ construction. For the details of this construction, see Sec. 2.2.3 in the paper by Viscondi, Grigolo, and de Aguiar [11].

For the sake of simplicity and adherence to existing literature, we will refer to this construction as $SU(M)$ coherent states, but we will still use the exponential map with the $SL(M, \mathbb{R})$ generators, as it leads to much simpler mathematical treatment.

1.3.4 General approach to particle-preserving CS decomposition

This method is based on the approach in [16, App. E]. Suppose we have a reference state $|\phi_0\rangle$, and the quotient space of the dynamical group of some system is transversed by the exponential map of the operator $\hat{F}(z)$, so that

$$|z\} = e^{\hat{F}(z)} |\phi_0\rangle \tag{1.30}$$

and $\hat{F}(z)$ is a linear combination of transition operators which all destroy the vacuum state[2]. Define $\hat{\phi}_0$ as such an operator so that

$$\hat{\phi}_0 |\text{vac.}\rangle = |\phi_0\rangle \tag{1.31}$$

Then we can write

$$|z\} = e^{\hat{F}(z)} \hat{\phi}_0 e^{-\hat{F}(z)} |\text{vac.}\rangle \tag{1.32}$$

using the fact that $\hat{F}(z) |\text{vac.}\rangle = 0$. We now express the operator product using Hadamard's lemma:

$$e^{\hat{F}(z)} \hat{\phi}_0 e^{-\hat{F}(z)} = \sum_{r=0}^{\infty} \frac{1}{r!} \left[\hat{F}(z), \hat{\phi}_0\right]_r \tag{1.33}$$

where

$$\left[\hat{A}, \hat{B}\right]_r = \left[\hat{A}, \left[\hat{A}, \hat{B}\right]_{r-1}\right] \quad \text{and} \quad \left[\hat{A}, \hat{B}\right]_0 = \hat{B} \tag{1.34}$$

are the repeated commutators. The unnormalised coherent state is obtained by acting with this sum on the vacuum state.

[2]This is characteristic of dynamical groups which preserve total particle number.

Chapter 2

Current work

In this chapter I outline the application of the mathematical construction of $SU(M)$ coherent states in particular to bosonic systems (such as the ones described in Sec. 1.1.1, 1.1.2) and fermionic systems (such as the one described in Sec. 1.1.3).

2.1 Bosonic $SU(M)$ coherent states

2.1.1 Construction

In this subsection, I describe the properties of $SU(M)$ coherent states constructed on bosonic annihilation and creation operators.

While I use a slightly different ansatz than the one in ([13, 16]) regarding the construction of bosonic $SU(M)$ coherent states, the difference is minor (I do not impose an extra condition on the components of the CS parameter z, and include an explicit normalisation factor instead). The ansatz I use is the same as the one used by Grigolo, Viscondi, and de Aguiar in [17]. Therefore the content of this subsection is not to be considered original work by me, and is instead a continuation of the literature review in the previous chapter. It is included in this chapter, however, to draw contrast with the fermionic $SU(M)$ coherent states, whose construction does contain original work by me (namely, the operator sequence matrix elements).

Reference state and displacement operator

A suitable reference state for a bosonic system of M modes and S particles is

$$|\phi_0\rangle = |0, 0 \dots 0, S\rangle \tag{2.1}$$

Finding the stability subgroup of the dynamical group for this reference state, we ultimately obtain the displacement operator

$$\hat{D}(z) = \exp\left(\sum_{i=1}^{M-1} z_i \hat{b}_i^\dagger \hat{b}_M\right) \tag{2.2}$$

The bosonic $SU(M)$ coherent state is then

$$|z\rangle = N(z) \exp\left(\sum_{i=1}^{M-1} z_i \hat{b}_i^\dagger \hat{b}_M\right) |\phi_0\rangle \tag{2.3}$$

where z is a vector of $M-1$ complex parameters

Decomposition into the occupancy basis

As shown by Buonsante and Penna in [16], this coherent state can be decomposed into the occupancy basis as follows:

$$|\vec{z}\rangle = \frac{N(z)}{\sqrt{S!}} \left(\hat{a}_M^\dagger + \sum_i^{M-1} z_i \hat{a}_i^\dagger\right)^S |0, 0 \dots 0\rangle \tag{2.4}$$

This result can be obtained by following the method outlined in Sec. 1.3.4. Applying the multinomial theorem gives the explicit superposition of occupancy basis states equivalent to the coherent state:

$$|z\rangle = N(z) \sum_{k_1+k_1+\cdots+k_M=S} \sqrt{\frac{S!}{k_1!k_2!\ldots k_M!}} z_1^{k_1} z_2^{k_2} \ldots z_{M-1}^{k_{M-1}} |k_1, k_2 \ldots k_M\rangle \tag{2.5}$$

Overlap and normalisation

Using Eq. 2.4 and Wick's theorem, the overlap of two bosonic CSs can be shown [13] to be

$$\langle z_a|z_b\rangle = N(z_a)N(z_b)\left(1 + \sum_i^{M-1} z_{a,i}^* z_{b,i}\right)^S \tag{2.6}$$

This also determines the normalisation function as $N(z) = (1 + \sum z_i^* z_i)^{-\frac{S}{2}}$

Matrix element of bosonic operator sequences

Consider a normal-ordered product sequence of x bosonic creation and x bosonic annihilation operators. Its matrix element for two arbitrary CSs can be shown [13] to be

$$\langle z_a|\hat{b}_{\rho_1}^\dagger \ldots \hat{b}_{\rho_x}^\dagger \hat{b}_{\sigma_1} \ldots \hat{b}_{\sigma_x}|z_b\rangle = z_{a,\rho_1}^* \ldots z_{a,\rho_x}^* z_{b,\sigma_1} \ldots z_{b,\sigma_x} \left\langle z_a^{(x)} \middle| z_b^{(x)} \right\rangle \tag{2.7}$$

where the *reduced overlap* is defined as

$$\left\langle z_a^{(x)} \middle| z_b^{(x)} \right\rangle = N(z_a)N(z_b)\left(1 + \sum_i^{M-1} z_{a,i}^* z_{b,i}\right)^{S-x} \tag{2.8}$$

i.e. the reduced unnormalised CSs behave as if their associated total particle number was reduced by x.

This is an important mathematical property to know in our methodology, as in the equations of motion operator sequences up to the third order are present (the partial derivatives of Hamiltonian matrix elements).

2.1.2 Time complexity of evaluation

Using standard techniques, the evaluation of the overlap in Eq. 2.6 has time complexity $O(M)$, as does the evaluation of the bosonic operator sequence matrix element in Eq. 2.7. Given a second-order Hamiltonian, its matrix element for two states $|z_a\rangle, |z_b\rangle$ may be evaluated with time complexity $O(M^5)$, and the full Hamiltonian matrix with time complexity

$O(M^5N^2)$ for a sample basis of N coherent states. The dynamical matrix Ω in Eq. 1.20 has dimensions $(MN \times MN)$, it may be evaluated in $O(M^3N^2)$, and inverted in $O(M^3N^3)$. Therefore the main bottlenecks are evaluating the Hamiltonian matrix (for large M) and inverting the dynamical matrix (for large N). However, in the case of the Bose-Hubbard model, the time complexity of evaluating the Hamiltonian matrix is lowered by a factor of M^2, as it is in tridiagonal form.

2.1.3 Results

The method proposed by us in Sec. 1.1.1 has been implemented and tested on the Bose-Hubbard model with the results by Qiao and Grossmann as a benchmark.

Qiao and Grossmann tested their fully variational method on a two-mode and a three-mode system of 50 particles, with the initial wavestate being a pure coherent state.

In the following figures, the "true" solution is obtained by explicit propagation on the full occupancy basis unless stated otherwise.

The particular Hamiltonian parameters in these results match the ones chosen by Qiao and Grossmann in [13].

Initial wavestate as a pure coherent state

When the initial wavestate is set to be a pure coherent state, our method becomes a massive improvement, since the trajectory of a single coherent state is semi-classical, and we can find it with a basis sample of size $N = 1$. In Fig. 2.1.3 I show the convergence of this minimal basis sample to the true solution for both two-mode and three-mode systems.

Initial wavestate as a superposition of coherent states

We wish to show that the propagation of each basis state separately yields a suitable time-dependent frozen basis for an arbitrary initial wavefunction. Therefore I extended the pool of initial wavestates to superpositions of a small number of coherent states, and chose the basis for each setup as a set of all present coherent states. Fig. 2.1.3 demonstrates the validity and scope of this approach.

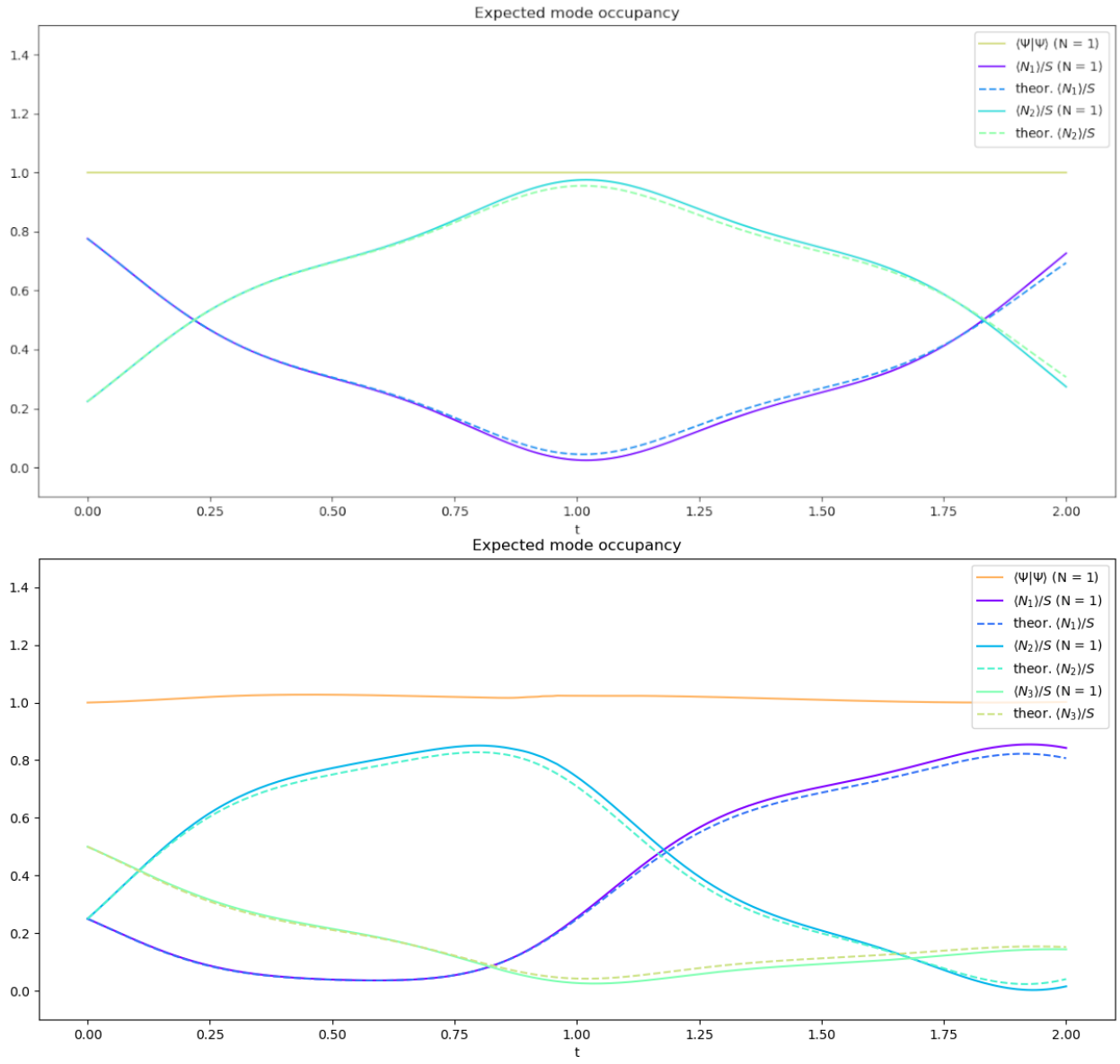

Figure 2.1: These graphs show the calculated trajectories of bosonic pure coherent states using the single CS equations of motion as compared to the decomposition onto the full occupancy basis. Both trajectories are calculated using the Dormand-Prince method. The top graph shows a system with $M = 2, S = 20$ and the bottom graph shows a system with $M = 3, S = 20$. We see convergence for both the two-mode and the three-mode system, although the three-mode system is slightly less numerically stable (notice the wiggle in the total wavefunction norm). The CS solution is obtained in the order of 100 ms; for the two-mode system, the full occupancy solution is obtained in the order of 1 s; for the three-mode system, the full occupancy solution is obtained in the order of 1 h.

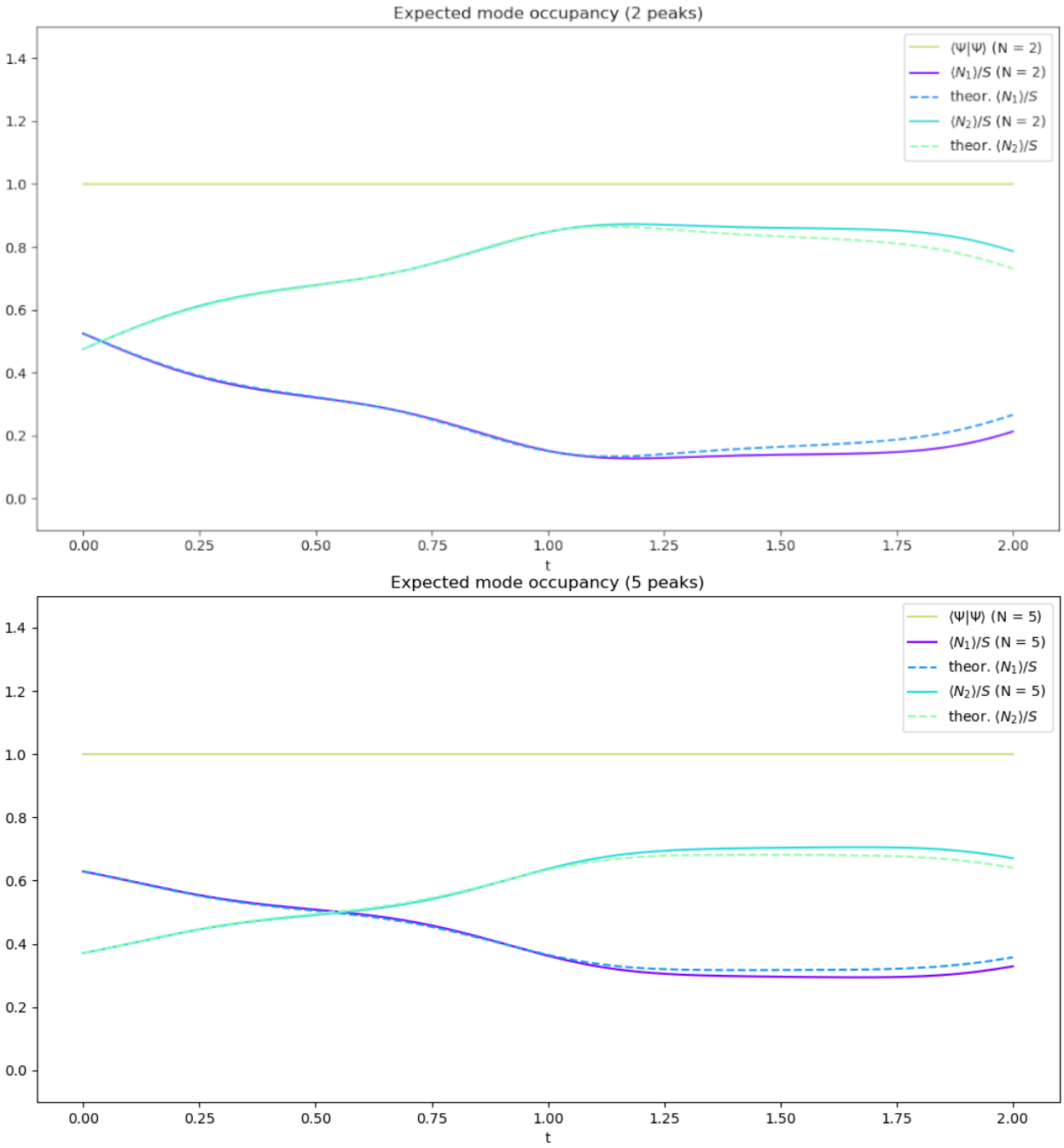

Figure 2.2: These graphs both show the evolution on a two-mode system with $S = 20$; the upper graph has the initial wavefunction a superposition of 2 pure coherent states, and the bottom graph a superposition of 5 pure coherent states. For both graphs, the basis is chosen to be equal to the components of the initial wavefunction superposition, and the evolution converges to the true solution. Choosing an initial wavefunction a superposition of a larger number of coherent states (say 20) will, however, prevent the model from converging easily, which shows the limit of this naive sampling method. For such highly-mixed initial wavefunctions, more general sampling methods are required (such as magnitude-squared weighted sampling from the parameter space). Nevertheless, this shows the validity of our method beyond pure-CS initial wavefunctions, extending it beyond the scope on which Qiao and Grossmann demonstrated their fully variational method.

2.2 Fermionic $SU(M)$ coherent states

2.2.1 Construction of the unnormalised state

In this subsection, I describe the properties of $SU(M)$ coherent states constructed on fermionic annihilation and creation operators. This construction is well-known in literature; while not equivalent, it is analogous to the construction described by Grigolo, Viscondi, and de Aguiar in [17, Sec. 4].

Reference state and displacement operator

Anticipating the application of this ansatz to finding the ground state of a complex system, the reference state is chosen as the Slater determinant of the S lowest-energy modes; in second quantisation, this is succintly written as

$$|\phi_0\rangle = |1, \dots 1, 0, \dots 0\rangle \quad \text{where the first } S \text{ modes are occupied} \tag{2.9}$$

The only transition operators $\hat{T}_{ij} = \hat{f}_i^\dagger \hat{f}_j$ which do not leave the reference state invariant up to a scalar factors are the ones with $i \in \{S+1 \dots M\}, j \in \{1 \dots S\}$. For convenience we define two sets of indices

$$\pi_1 = \{1, 2 \dots S\} \qquad \pi_0 = \{S+1, S+2 \dots M\} \tag{2.10}$$

Then the displacement operator becomes

$$\hat{D}(Z) = \exp\left(\sum_{i\in\pi_0}\sum_{j\in\pi_1} Z_{ij}\hat{f}_i^\dagger \hat{f}_j\right) \tag{2.11}$$

and the fermionic $SU(M)$ coherent state is

$$|Z\rangle = N(Z)\exp\left(\sum_{i\in\pi_0}\sum_{j\in\pi_1} Z_{ij}\hat{f}_i^\dagger \hat{f}_j\right)|\phi_0\rangle \tag{2.12}$$

where Z is a complex $(M - S \times S)$ matrix.

Decomposition into the occupancy basis

We use Eq. 1.33. Firstly, the reference state operator is

$$\hat{\phi}_0 = \hat{f}^\dagger_{\langle\pi_1\rangle} \tag{2.13}$$

where $\langle \pi_1 \rangle$ denotes the ascending sequence formed from elements of the set π_1. We also have the operator

$$\hat{F} = \sum_{i \in \pi_0} \sum_{j \in \pi_1} Z_{ij} \hat{f}_i^\dagger \hat{f}_j \tag{2.14}$$

Then the repeated commutator may be found to be

$$\left[\hat{D}(Z), \hat{\phi}_0\right]_x = (-1)^{\frac{1}{2}x(x+1)} x! \sum_{\langle a \rangle \in \Gamma_x \langle \pi_0 \rangle} \sum_{\langle b \rangle \in \Gamma_x \langle \pi_1 \rangle} \det\left(Z_{\langle a \rangle, \langle b \rangle}\right) \hat{f}^\dagger_{\langle \pi_1 - \{b\} + \{a\} \rangle} \tag{2.15}$$

where $\Gamma_x \langle S \rangle$ denotes all increasing subsequences of length x of index sequence $\langle S \rangle$. In other words, the unnormalised coherent state

$$|Z\} = \sum_{r=0}^{\min(S, M-S)} (-1)^{\frac{1}{2}r(r+1)} \sum_{\langle a \rangle \in \Gamma_r \langle \pi_0 \rangle} \sum_{\langle b \rangle \in \Gamma_r \langle \pi_1 \rangle} \det\left(Z_{\langle a \rangle, \langle b \rangle}\right) |\pi_1 - \{b\} + \{a\}\rangle \tag{2.16}$$

is a superposition of occupancy basis states such that the component with particles excited from initially occupied indices $\langle b \rangle$ to initially unoccupied indices $\langle a \rangle$ is weighted by the minor of the parameter matrix obtained by selecting the rows labelled by $\langle a \rangle$ and columns labelled by $\langle b \rangle$, up to a sign factor.

2.2.2 Overlap and normalisation

The overlap of two unnormalised fermionic coherent states may be expressed as

$$\{Z_a | Z_b\} = \sum_{r=0}^{\min(S, M-S)} \sum_{\langle a \rangle \in \Gamma_r \langle \pi_0 \rangle} \sum_{\langle b \rangle \in \Gamma_r \langle \pi_1 \rangle} \det\left((Z_a^\dagger)_{\langle b \rangle, \langle a \rangle}\right) \det\left((Z_b)_{\langle a \rangle, \langle b \rangle}\right) \tag{2.17}$$

where we may use the identity

$$\det\left((Z_a^\dagger)_{\langle b \rangle, \langle a \rangle}\right) \det\left((Z_b)_{\langle a \rangle, \langle b \rangle}\right) = \det\left((Z_a^\dagger Z_b)_{\langle b \rangle, \langle b \rangle}\right) = \det\left((Z_b Z_a^\dagger)_{\langle a \rangle, \langle a \rangle}\right) \tag{2.18}$$

to express this as a sum of all principal minors of a square matrix with sign alternating between their ranks. Using Lemma A.1.1 yields the overlap

$$\{Z_a | Z_b\} = \det\left(I + Z_a^\dagger Z_b\right) = \det\left(I + Z_b Z_a^\dagger\right) \tag{2.19}$$

which also determines the normalisation factor

$$N(Z) = \frac{1}{\sqrt{\det(I + Z^\dagger Z)}} \tag{2.20}$$

which agrees with the result in [17, Eq. 2.31].

2.2.3 Fermionic operator sequence matrix element

We now wish to evaluate the matrix element

$$S = \{Z_a|\hat{f}^\dagger_{\rho_1}\dots\hat{f}^\dagger_{\rho_x}\hat{f}_{\rho'_1}\dots\hat{f}_{\rho'_x}|Z_b\} \tag{2.21}$$

I was not able to find a general result of this form in existing literature, and thus I now present what is, according to the best of my knowledge, my original work and main theoretical contribution to the mathematical construction of $SU(M)$ fermionic coherent states.

First, we express the sequences ρ, ρ' as permutations of strictly descending and ascending sequences, respectively: $\rho = P_-\langle\rho\rangle^-, \rho' = P_+\langle\rho'\rangle^+$. Replacing the original sequences with the ordered versions introduces a factor of $\mathrm{sgn}(P_-)\mathrm{sgn}(P_+)$. We now partition the sequences into segments constructed of elements of π_1 and π_0:

$$S = \mathrm{sgn}(P_-)\mathrm{sgn}(P_+)\,\{Z_a|\hat{f}^\dagger_{\langle\tau\rangle^-}\hat{f}^\dagger_{\langle\sigma\rangle^-}\hat{f}_{\langle\sigma'\rangle}\hat{f}_{\langle\tau'\rangle}|Z_b\} \tag{2.22}$$

where

$$\rho = \sigma\cup\tau \quad \text{where} \quad \sigma\in\pi_1, \tau\in\pi_0 \quad \text{and} \quad \rho' = \sigma'\cup\tau' \quad \text{where} \quad \sigma'\in\pi_1, \tau'\in\pi_0 \tag{2.23}$$

Let us use $\eta_x(\chi)$ to denote the number of elements in set χ smaller than x. This is particularly useful to keep track of the Jordan-Wigner string. Taking the matrix element as an overlap of $\hat{f}^\dagger_{\langle\sigma\rangle^+}\hat{f}^\dagger_{\langle\tau\rangle^+}|Z_a\}$ and $\hat{f}_{\langle\sigma'\rangle}\hat{f}_{\langle\tau'\rangle}|Z_b\}$ and using the decomposition into the occupancy basis, we obtain

$$\begin{aligned}
S = \mathrm{sgn}(P_-)\mathrm{sgn}(P_+) &\sum_{r=|\tau|}^{\min(S-|\sigma|,M-S)}(-1)^{\frac{1}{2}r(r+1)}\sum_{r'=|\tau'|}^{\min(S-|\sigma'|,M-S)}(-1)^{\frac{1}{2}r'(r'+1)}\\
&\sum_{\langle a\rangle\in\Gamma_{r-|\tau|}\langle\pi_0-\tau\rangle}\ \sum_{\langle b\rangle\in\Gamma_r\langle\pi_1-\sigma\rangle}\ \sum_{\langle a'\rangle\in\Gamma_{r'-|\tau'|}\langle\pi_0-\tau'\rangle}\ \sum_{\langle b'\rangle\in\Gamma_{r'}\langle\pi_1-\sigma'\rangle}(-1)^{|\tau|(S-r)+\frac{1}{2}|\tau|(|\tau|-1)+\sum_i\eta_{\tau_i}(\langle a\rangle)}\\
&(-1)^{-|\sigma|+\sum_i(\sigma_i+\eta_{\sigma_i}(\langle b\rangle))}(-1)^{|\tau'|(S-r')+\frac{1}{2}|\tau'|(|\tau'|-1)+\sum_i\eta_{\tau'_i}(\langle a'\rangle)}(-1)^{-|\sigma'|+\sum_i(\sigma'_i+\eta_{\sigma'_i}(\langle b'\rangle))}\\
&\det\big((Z^\dagger_a)_{\langle b\rangle,\langle a\cup\tau\rangle}\big)\det\big((Z_b)_{\langle a'\cup\tau'\rangle,\langle b'\rangle}\big)\,\langle\pi_1\cup a-b\cup\sigma|\pi_1\cup a'-b'\cup\sigma'\rangle
\end{aligned} \tag{2.24}$$

The occupancy basis overlap is equivalent to

$$\langle\pi_1\cup a-b\cup\sigma|\pi_1\cup a'-b'\cup\sigma'\rangle = \delta_{\langle a\rangle,\langle a'\rangle}\delta_{\langle b\cup\sigma\rangle,\langle b'\cup\sigma'\rangle}\delta_{r-|\tau|,r'-|\tau|}\delta_{r+|\sigma|,r'+|\sigma|} \tag{2.25}$$

Note that, since $|\sigma|+|\tau| = |\sigma'|+|\tau'|$ unless the overlap vanishes due to mismatched total number of particles, the final two Kronecker deltas for r, r' are equivalent.

We now take

$$\gamma = r-|\tau| = r'-|\tau'| \quad \text{so that} \quad r = \gamma+|\tau|, r' = \gamma+|\tau'| \tag{2.26}$$

$$\langle\alpha\rangle\in\Gamma_\gamma\langle\pi_0-\tau\cup\tau'\rangle \quad \text{so that} \quad \langle a\rangle = \langle a'\rangle = \langle\alpha\rangle \tag{2.27}$$

$$\langle\beta\rangle\in\Gamma_{\gamma+|\tau|-|\sigma'-\sigma\cap\sigma'|}\langle\pi_1-\sigma\cup\sigma'\rangle \quad \text{so that} \quad \langle b\rangle = \langle\beta\cup\sigma'-\sigma\cap\sigma'\rangle, \langle b'\rangle = \langle\beta\cup\sigma-\sigma\cap\sigma'\rangle \tag{2.28}$$

where $|\tau| - |\sigma'| = |\tau'| - |\sigma|$ and the construction of $\langle b\rangle, \langle b'\rangle$ omits $\sigma \cap \sigma'$, since the terms with $\langle b\rangle$ containing any element in σ vanish (same for $\langle b'\rangle$ and σ').

Substituing $r, r', \langle a\rangle, \langle a'\rangle, \langle b\rangle, \langle b'\rangle$ and using simple algebraic manipulation we can show that, for terms with non-vanishing Kronecker deltas, the total sign simplifies significantly. Denoting $\varsigma = \sigma - \sigma \cap \sigma', \varsigma' = \sigma' - \sigma \cap \sigma'$, the overlap can be written as

$$= (-1)^{S(|\tau|+|\tau'|)+(|\varsigma|-1)(|\varsigma'|-1)+1+\sum\langle\varsigma\rangle+\sum\langle\varsigma'\rangle+\sum_i \eta_{(\sigma\cap\sigma')_i}(\langle\varsigma\cup\varsigma'\rangle)} \sum_{\gamma=0} \sum_{\langle\alpha\rangle\in\Gamma_\gamma\langle\pi'_0-\tau\cup\tau'\rangle} \sum_{\langle\beta\rangle\in\Gamma_{\gamma+|\tau|-|\varsigma'|}\langle\pi'_1-\sigma\cup\sigma'\rangle}$$
$$(-1)^{\sum_i \eta_{\varsigma_i}(\langle\beta\rangle)+\sum_i \eta_{\varsigma'_i}(\langle\beta\rangle)+\sum_i \eta_{\tau_i}(\langle\alpha\rangle)+\sum_i \eta_{\tau'_i}(\langle\alpha\rangle)} \det\left((Z_a^\dagger)^{(\mathrm{r}.\sigma,\mathrm{c}.\tau'-\tau\cap\tau')}_{\langle\beta\cup\varsigma'\rangle,\langle\alpha\cup\tau\rangle}\right) \det\left((Z_b)^{(\mathrm{r}.\tau-\tau\cap\tau',\mathrm{c}.\sigma')}_{\langle\alpha\cup\tau'\rangle,\langle\beta\cup\varsigma\rangle}\right) \tag{2.29}$$

where the superscript $(\mathrm{r}.X), (\mathrm{c}.X)$ means omitting the rows or columns specified by the set of indices X, and where the summation over $\gamma, \langle\alpha\rangle, \langle\beta\rangle$ is such that all square submatrices of $(Z_a^\dagger)^{(\mathrm{r}.\sigma)}, (Z_b)^{(\mathrm{c}.\sigma')}$ are present in the sum, as denoted by the apostrophed π_1, π_0, which represents the omission of indices corresponding to the removed rows and columns.

We now choose to permute the rows and columns of $(Z_a^\dagger)^{(\mathrm{r}.\sigma)}, (Z_b)^{(\mathrm{c}.\sigma')}$ as to bring the rows and columns which are included in every submatrix in every term of the sum to the lowest-index position. This introduces an extra sign factor to the determinant, which cancels the second sign term in the sum above. Formally

$$X = \begin{pmatrix} (Z_a^\dagger)_{\langle\varsigma'\rangle,\langle\tau\rangle} & (Z_a^\dagger)^{(\mathrm{c}.\tau\cup\tau')}_{\mathrm{r}.\langle\varsigma'\rangle} \\ (Z_a^\dagger)^{(\mathrm{r}.\sigma\cup\sigma')}_{\mathrm{c}.\langle\tau\rangle} & (Z_a^\dagger)^{(\mathrm{r}.\sigma\cup\sigma',\mathrm{c}.\tau\cup\tau')} \end{pmatrix} \tag{2.30}$$

$$Y = \begin{pmatrix} (Z_b)_{\langle\tau'\rangle,\langle\varsigma\rangle} & (Z_b)^{(\mathrm{c}.\sigma\cup\sigma')}_{\mathrm{r}.\langle\tau'\rangle} \\ (Z_b)^{(\mathrm{r}.\tau\cup\tau')}_{\mathrm{c}.\langle\varsigma\rangle} & (Z_b)^{(\mathrm{r}.\tau\cup\tau',\mathrm{c}.\sigma\cup\sigma')} \end{pmatrix} \tag{2.31}$$

The resulting expression is exactly in the form which is treated by Lemma A.1.3. Hence, if $|\tau| \leq |\tau'|$, we have

$$\langle Z_a|\hat{f}^\dagger_{\langle\tau\rangle-}\hat{f}^\dagger_{\langle\sigma\rangle-}\hat{f}_{\langle\sigma'\rangle}\hat{f}_{\langle\tau'\rangle}|Z_b\rangle = N(Z_a)N(Z_b)(-1)^{S(|\tau|+|\tau'|)+(|\varsigma|-1)(|\varsigma'|-1)+1+\sum\langle\varsigma\rangle+\sum\langle\varsigma'\rangle+\sum_i \eta_{(\sigma\cap\sigma')_i}(\langle\varsigma\cup\varsigma'\rangle)}$$
$$(-1)^{|\tau'|(1+|\tau'|-|\tau|)} \det\begin{pmatrix} 0_{|\tau'|,|\tau|} & (Z_b)_{\langle\tau'\rangle,\langle\varsigma\rangle} & (Z_b)^{(\mathrm{c}.\sigma\cup\sigma')}_{\mathrm{r}.\langle\tau'\rangle} \\ (Z_a^\dagger)_{\langle\varsigma'\rangle,\langle\tau\rangle} & (Z_a^\dagger)^{(\mathrm{c}.\tau\cup\tau')}_{\mathrm{r}.\langle\varsigma'\rangle}(Z_b)^{(\mathrm{r}.\tau\cup\tau')}_{\mathrm{c}.\langle\varsigma\rangle} & (Z_a^\dagger)^{(\mathrm{c}.\tau\cup\tau')}_{\mathrm{r}.\langle\varsigma'\rangle}(Z_b)^{(\mathrm{r}.\tau\cup\tau',\mathrm{c}.\sigma\cup\sigma')} \\ (Z_a^\dagger)^{(\mathrm{r}.\sigma\cup\sigma')}_{\mathrm{c}.\langle\tau\rangle} & (Z_a^\dagger)^{(\mathrm{r}.\sigma\cup\sigma',\mathrm{c}.\tau\cup\tau')}(Z_b)^{(\mathrm{r}.\tau\cup\tau')}_{\mathrm{c}.\langle\varsigma\rangle} & I + (Z_a^\dagger)^{(\mathrm{r}.\sigma\cup\sigma',\mathrm{c}.\tau\cup\tau')}(Z_b)^{(\mathrm{r}.\tau\cup\tau',\mathrm{c}.\sigma\cup\sigma')} \end{pmatrix} \tag{2.32}$$

If $|\tau| \geq |\tau'|$, we have $|\varsigma| \leq |\varsigma'|$, which allows us to apply Lemma A.1.3 to obtain the expression

$$\langle Z_a|\hat{f}^\dagger_{\langle\tau\rangle_-}\hat{f}^\dagger_{\langle\sigma\rangle_-}\hat{f}_{\langle\sigma'\rangle}\hat{f}_{\langle\tau'\rangle}|Z_b\rangle = N(Z_a)N(Z_b)(-1)^{S(|\tau|+|\tau'|)+(|\varsigma|-1)(|\varsigma'|-1)+1+\sum\langle\varsigma\rangle+\sum\langle\varsigma'\rangle+\sum_i \eta_{(\sigma\cap\sigma')_i}(\langle\varsigma\cup\varsigma'\rangle)}$$
$$(-1)^{|\varsigma'|(1+|\varsigma'|-|\varsigma|)}\det\begin{pmatrix} 0_{|\varsigma'|,|\varsigma|} & (Z_a^\dagger)_{\langle\varsigma'\rangle,\langle\tau\rangle} & (Z_a^\dagger)^{(\mathrm{c}.\tau\cup\tau')}_{\mathrm{r}.\langle\varsigma'\rangle} \\ (Z_b)_{\langle\tau'\rangle,\langle\varsigma\rangle} & (Z_b)^{(\mathrm{c}.\sigma\cup\sigma')}_{\mathrm{r}.\langle\tau'\rangle}(Z_a^\dagger)^{(\mathrm{r}.\sigma\cup\sigma')}_{\mathrm{c}.\langle\tau\rangle} & (Z_b)^{(\mathrm{c}.\sigma\cup\sigma')}_{\mathrm{r}.\langle\tau'\rangle}(Z_a^\dagger)^{(\mathrm{r}.\sigma\cup\sigma',\mathrm{c}.\tau\cup\tau')} \\ (Z_b)^{(\mathrm{r}.\tau\cup\tau')}_{\mathrm{c}.\langle\varsigma\rangle} & (Z_b)^{(\mathrm{r}.\tau\cup\tau',\mathrm{c}.\sigma\cup\sigma')}(Z_a^\dagger)^{(\mathrm{r}.\sigma\cup\sigma')}_{\mathrm{c}.\langle\tau\rangle} & I+(Z_b)^{(\mathrm{r}.\tau\cup\tau',\mathrm{c}.\sigma\cup\sigma')}(Z_a^\dagger)^{(\mathrm{r}.\sigma\cup\sigma',\mathrm{c}.\tau\cup\tau')} \end{pmatrix} \tag{2.33}$$

Note that for the case $|\tau| > |\tau'|$, we can simply take

$$\langle Z_a|\hat{f}^\dagger_{\langle\tau\rangle_-}\hat{f}^\dagger_{\langle\sigma\rangle_-}\hat{f}_{\langle\sigma'\rangle}\hat{f}_{\langle\tau'\rangle}|Z_b\rangle = \langle Z_b|\hat{f}^\dagger_{\langle\tau'\rangle_-}\hat{f}^\dagger_{\langle\sigma'\rangle_-}\hat{f}_{\langle\sigma\rangle}\hat{f}_{\langle\tau\rangle}|Z_a\rangle^* \tag{2.34}$$

so that we can always use Eq. 2.32 as the standard expression.

2.2.4 Connection to molecular electronic structure

An example of a fermionic particle-preserving system, as stated in Sec. 1.1.3, is the molecular electronic structure when no chemical process is occuring. Formally, we may index all the atomic orbitals of the constituent atoms such that the occupied orbitals are included in π_1, unoccupied orbitals are included in π_0, and within both of these sets, the orbitals are ordered by their self-energy (it is important to take notice of spin degeneracy, as each orbitals actually corresponds to two modes). This is a gross-structure model, governed by a second-order Hamiltonian

$$\hat{H} = V^{(1)}_{\alpha\beta}\hat{f}^\dagger_\alpha\hat{f}_\beta + \frac{1}{2}V^{(2)}_{\alpha\beta\gamma\delta}\hat{f}^\dagger_\alpha\hat{f}^\dagger_\beta\hat{f}_\gamma\hat{f}_\delta \tag{2.35}$$

where $V^{(1)}$ is the single-body energy tensor, here comprising of the kinetic term and electron-nucleus interaction, and $V^{(2)}$ is the two-body energy tensor, comprising of the electron-electron exchange integrals (care must be exercised, as in Mulliken notation, the quadratic term is not normal-ordered, and thus an extra diagonal term is introduced). These tensors are typically obtained by numerical calculation; I use the framework of Python, and specifically the PySCF package to calculate them for an arbitrary molecule with an arbitrary geometry and arbitrary basis for the atomic orbitals.

2.2.5 How to calculate the ground state

Once the second-quantised framework of $SU(M)$ coherent states is applied to our desired molecule, we have multiple options on how to quickly approximate the ground state. Below are listed a few methods we are in the process of implementing for the purpose of benchmarking. All of these methods use the state $|Z = 0\rangle$, i.e. the "atomic orbitals populated

as if outside of a molecule" state, as a starting point. This state shall be, for convenience, referred to as the "null state".

The first two methods specifically build on the idea of restricting the full Hilbert space into a much smaller subspace by quick energy considerations, on which full diagonalisation is then performed. The null state is taken as belonging to this small subspace, and the ground state is assumed to be "close" to the null state (formally, their overlap is assumed to be close to unity). In the first two of the three methods presented, the low-energy subspace is formed on a basis consisting of fermionic $SU(M)$ coherent states, and the ground state estimate is expressed as their superposition (which is not necessarily a pure coherent state in itself).

Random walk around the null state

The simplest algorithm for sampling the low-energy subspace is a random walk around the null state, onto which conditioning may be imposed. Formally, in each step a candidate state is considered by randomly sampling the coherent-state parameter space in the vicinity of the null state, and the candidate is either included in the full sample or rejected based on an arbitrary criterion (for example, whether its expected energy is under a certain treshold compared to the null state energy, or whether the eigenvalues of the Hamiltonian matrix on the full sample change in a favourable way). Then, this step is repeated either until the sample reaches a certain size, the estimated ground state energy no longer decreases dramatically, or a certain number of rejections occured. Taking the sample $|Z_a\rangle, a = 1 \dots N$ as an un-orthogonal basis of the low-energy subspace, constructing the Hamiltonian matrix $H_{ij} = \langle Z_i|\hat{H}|Z_j\rangle$, diagonalising it, choosing the eigenvector with the lowest associated eigenvalue and expressing it in the occupancy basis then yields the estimate of the ground state.

Sampling the trajectory of the null state

Another method for selecting a basis of the low-energy subspace exploits the assumption that the null state has low energy. Then its trajectory under proper time evolution traces states with the same low energy. Sampling this trajectory at regular time intervals then yields the basis of a low-energy subspace, diagonalising the Hamiltonian on which then yields the estimate of the ground state.

Imaginary time propagation

This is a well-known method [18] based on the fact that a superposition of energy eigenstates propagated in time has the phase of each component change at a rate proportional to the associated energy, with the amplitude remaining constant. Propagating along the imaginary time $-it$ then instead has the amplitude of each component decay exponentially at a rate given by the component's energy. Eventually, after a sufficiently prolonged propagation, only

the lowest-energy state is present in the decomposition. Assuming the null state contains the ground state as a component, propagating the null state along $-it$ should then converge to the ground state. This method is theoretically guaranteed to converge to the true ground state, but the rate of convergence may be slow if the low-excitation energy eigenstates have comparable energy levels to the ground state.

Chapter 3

Outlook

In this chapter the scope of work which has been done is summarised and an outline of the next steps of the research is presented.

3.1 Summary of original work done

During the first year of this doctorate degree, I have done the following original work:

- I have generalised the matrix form of the fully-variational equations of motion on a discrete CS sample as quoted by Qiao and Grossmann [13, App. B], which is also applicable to fermionic coherent states.
- Based on the suggestion by my supervisor Prof Dmitry Shalashilin, I have built a framework in Python 3 for bosonic $SU(M)$ dynamics on an arbitrary second-order Hamiltonian for both the fully variational method and the new separate-basis-trajectory method, and implemented the Bose-Hubbard and the displaced harmonic trap systems in this framework.
- Using this framework, I have achieved convergence for two- and three-mode systems for the Bose-Hubbard model on a scale which matches that of the work of Qiao and Grossmann, demonstrating the utility of the separate-basis-trajectory method. I have also extended the scope of this investigation onto problems with initial wavefunctions which are not pure bosonic coherent states.
- I have derived the general matrix element of a product sequence of fermionic creation and annihilation operators for fermionic $SU(M)$ coherent states, which allows the evaluation of matrix elements of arbitrary particle-preserving operators, including the Hamiltonian.

- I have built the skeleton of a framework in Python 3 using PySCF to translate the single-electron and electron-exchange energy integrals for an arbitrary molecule into a mode system onto which fermionic $SU(M)$ coherent state-based methods may be applied.

3.2 Research plan for the duration of the doctorate programme

I propose the following tasks for the remaining duration of my doctorate programme to be done within this research project:

- Finishing the bosonic dynamics for separate basis trajectories (estimated duration: 2 months). This entails
  - Implementing arbitrary initial wavefunction sampling techniques and testing various conditioning methods to maximise numerical stability
  - Scaling the mode and particle numbers to see the limits of this method
  - For high mode and particle numbers, it is no longer possible to calculate the "true" solution on the full occupancy basis; comparison data must be collected by other techniques, e.g. MCSCF.
- Obtaining a succesful estimate for the ground state of a complex molecule using $SU(M)$ fermionic coherent states (estimated duration: 2 months).
- Further investigation of the mathematical properties of $SU(M)$ fermionic coherent states, with the focus on lowering the time complexity of calculating the Hamiltonian matrix elements (estimated duration: 3 months).
- Benchmarking different methods for obtaining the molecular ground state (estimated duration: 3 months).
- Investigating alternative constructions for obtaining the molecular electronic ground state, such as the boson-analogous construction recently proposed by Prof Dmitry Shalashilin (estimated duration: 6 months).
- Applying the fermionic $SU(M)$ methods for dynamics to other particle-preserving fermionic systems (e.g. qubit systems, electron gas with high chemical potential etc) (estimated duration: 6 months).

- Investigating the underlying connection between bosonic and fermionic $SU(M)$ coherent states and generalising it to other dynamical groups, which massively enlargens the scale of the proposed methods in terms of fermionic physical systems which may be analysed (estimated duration: 1 year in parallel with other work)

3.3 Data management plan

The project is purely theoretical/computational, and thus no collection of laboratory data is required.

Collection of numerical data from the frameworks developed by me has so far been done on my personal computer and stored on a github repository. As I move ahead to upscale the data collection, especially for large-mode bosonic systems and complex molecules, I will create a designated workflow for collecting data on the MacBook machine provided to me by the Faculty, so that long-term data collection may be put into place without interrupting the development workflow.

In case the computational complexity of the largest systems we choose to analyse exceeds what may reasonably covered by a single machine, parallelisation or employing high-performance computing systems will be considered; however, given that the utility we are trying to demonstrate is that of a low-time-complexity approach, this will probably not be necessary.

Bibliography

[1] Shalashilin, D. V. (2011), Multiconfigurational Ehrenfest approach to quantum coherent dynamics in large molecular systems. *Faraday Discussions*, **153**, pp. 105–116

[2] Crespo-Otero, R., Barbatti, M. (2018), Recent Advances and Perspectives on Nonadiabatic Mixed Quantum–Classical Dynamics. *Chem. Rev.*, **118**, pp. 7026–7068

[3] Mattos, R. S., Mukherjee, S., Barbatti, M. (2024), Quantum Dynamics from Classical Trajectories. *J. Chem. Theory Comput.*, **20**, pp. 7728–7743

[4] Mulliken, R. S. (1932), Electronic Structures of Polyatomic Molecules and Valence. II. General Considerations. *Phys. Rev.*, **14**, pp. 49–71

[5] Schrodinger, E. (1926), Der stetige übergang von der Mikro-zur Makromechanik. *Naturwiss.* 14, pp. 664-666. Translated into English in Schrodinger, E. (1928), *Collected Papers in Wave Mechanics*. 1st edn. London: Blackie & Son, pp. 41–44

[6] Glauber, R. J. (1963), Coherent and Incoherent States of the Radiation Field. *Phys. Rev.*, **131**, 6, pp. 2766–2788

[7] Klauder, J. R., Skagerstam, B. S. (1985), *Coherent States: Applications in Physics and Mathematical Physics*. 1st eng. edn. Singapore: World Scientific.

[8] Perelomov, A. M. (1972), Coherent states for arbitrary Lie group. *Commun. Math. Phys.*, **26**, pp. 222–236

[9] Gilmore, R. (1972), Geometry of symmetrized states. *Ann. Phys.*, **74**, 2, pp. 391–463

[10] Zhang, W. M., Feng, D. H., Gilmore, R. (1990), Coherent states: Theory and some applications. *Rev. Mod. Phys.*, **62**, pp. 867–927

[11] Viscondi, T. F., Grigolo, A., de Aguiar, M. A. M. (2015), Semiclassical Propagator in the Generalized Coherent-State Representation. arXiv:1510.05952 [**quant-ph**]

[12] Block, I., Dalibard, J., Zwerger, W. (2008), Many-body physics with ultracold gases. *Rev. Mod. Phys.*, **80**, pp. 885–964

[13] Qiao, Y., Grossmann, F. (2021), Exact variational dynamics of the multimode Bose-Hubbard model based on $SU(M)$ coherent states. *Phys. Rev. A*, **103**, 042209

[14] Green, J. A., Shalashilin, D. V. (2019), Simulation of the quantum dynamics of indistinguishable bosons with the method of coupled coherent states. *Phys. Rev. A*, **100**, 013607

[15] Bertlmann, R. A., Krammer, P. (2008), Bloch vectors for qudits. arXiv:0806.1174 [**quant-ph**]

[16] Buonsante, P., Penna, V.(2008), Some remarks on the coherent-state variational approach to nonlinear boson models. *J. Phys. A: Math. Theor.*, **41**, 175301

[17] Grigolo, A., Viscondi, T. F., de Aguiar, M. A. M. (2016), Multiconfigurational quantum propagation with trajectory-guided generalized coherent states. *J. Chem. Phys.*, **144**, 094106

[18] Pederiva, F., Roggero, A., Schmidt, K. E. (2017), "Variational and Diffusion Monte Carlo Approaches to the Nuclear Few- and Many-Body Problem" in *An Advanced Course in Computational Nuclear Physics: Bridging the Scales from Quarks to Neutron Stars.* Springer International Publishing, Cham, pp. 401–476

[19] Horn, R. A., Johnson, C. R. (2012), *Matrix Analysis.* 2nd edn. Cambridge University Press, Cambridge. doi.org/10.1017/CBO9781139020411

[20] Chapman, A., Miyake, A. (2018), Classical simulation of quantum circuits by dynamical localization: analytic results for Pauli-observable scrambling in time-dependent disorder. Available at arxiv.org/pdf/1704.04405v2 [quant-ph]

Appendix A

Auxilary theorems

This appendix presents mathematical results which were used to derive the mathematical properties of bosonic and fermionic coherent states, but which were chosen to be omitted from the main body of text due to their abstract mathematical nature.

The author does not claim originality of these theorems or their proofs, but chooses to include them, as he was unable to find them in existing literature. The work presented in this appendix is the author's, except for mathematical identities which are stated explicitly.

A.1 Lemmas for counting matrix minors under constraints

In the following section, $\langle x \rangle$ for integer x denotes the sequence $\langle 1, 2 \dots x \rangle$, and $\langle u \rangle \oplus \langle v \rangle$ denotes the increasing sequence constructed from all the elements present in the sequences $\langle u \rangle, \langle v \rangle$.

Lemma A.1.1 *For a square matrix M, the sum of all principal minors of all ranks which contain the first n rows/columns is equal to $\det\big(I^{(n)} + M\big)$, where $I^{(n)}$ is the identity matrix with the first n elements along the diagonal set to zero.*

Proof. We will proceed by induction. Firstly, for $n = 0$, the sum is the sum of all principal minors for each rank r, denoted E_r. These sums form the coefficients of the characteristic polynomial of M like so: [19, Th. 1.2.16]

$$\det(tI - M) = \sum_{r=0}^{x} (-1)^r t^{x-r} E_r \tag{A.1}$$

where x is the number of rows of M. Evaluating this sum at $t = -1$ yields

$$\det(-I - M) = \sum_{r=0}^{x} (-1)^r (-1)^{x-r} E_r$$
$$\det(I + M) = \sum_{r=0}^{x} E_r \tag{A.2}$$

which shows the lemma holds for $n = 0$. Now: assume the lemma holds for n. For $n+1$, we have the sum of all principal minors of M which contain the first $n+1$ rows/columns. We identify this as equivalent to the sum of all principal minors of M which contain the first n rows/columns, from which we subtract the sum of all principal minors of M which contain the first n rows/columns and do *not* contain the $(n+1)$-th row/column. I.e. we can write

$$\sum_{r=n+1}^{x} \sum_{\langle a \rangle \in \Gamma_r \langle x^{(n+1)} \rangle} \det\left(M_{\langle n+1 \rangle \oplus \langle a \rangle, \langle n+1 \rangle \oplus \langle a \rangle}\right) = \tag{A.3}$$
$$\sum_{r=n}^{x} \sum_{\langle a \rangle \in \Gamma_r \langle x^{(n)} \rangle} \left[\det\left(M_{\langle n \rangle \oplus \langle a \rangle, \langle n \rangle \oplus \langle a \rangle}\right) - \det\left(M'_{\langle n \rangle \oplus \langle a \rangle, \langle n \rangle \oplus \langle a \rangle}\right)\right]$$

where M' is obtained by setting all elements in the $(n+1)$-th row/column to zero. To the right side of the equation we apply the lemma, since it is assumed it holds for n:

$$\sum_{r=n+1}^{x} \sum_{\langle a \rangle \in \Gamma_r \langle x^{(n+1)} \rangle} \det\left(M_{\langle n+1 \rangle \oplus \langle a \rangle, \langle n+1 \rangle \oplus \langle a \rangle}\right) = \det\left(I^{(n)} + M\right) - \det\left(I^{(n)} + M'\right) \tag{A.4}$$

Now, taking the Laplace expansion of $\det\left(I^{(n)} + M\right)$ along the $(n+1)$-th row, the element in the $(n+1)$-th column is $1 + M_{n+1,n+1}$ and its cofactor is $\det\left(I^{(n)} + M'\right)$. Subtracting 1 from this element and adding the cofactor to the full Laplace expansion preserves the determinant, revealing

$$\det\left(I^{(n)} + M\right) = \det\left(I^{(n+1)} + M\right) + \det\left(I^{(n)} + M'\right) \tag{A.5}$$

Substituing Eq. A.5 into Eq. A.4 yields

$$\sum_{r=n+1}^{x} \sum_{\langle a \rangle \in \Gamma_r \langle x^{(n+1)} \rangle} \det\left(M_{\langle n+1 \rangle \oplus \langle a \rangle, \langle n+1 \rangle \oplus \langle a \rangle}\right) = \det\left(I^{(n+1)} + M\right) \tag{A.6}$$

which finishes the proof.

Lemma A.1.2 *Consider two matrices X, Y of shapes (m, n) and (n, m), respectively. The sum*

$$\sum_{r=0}^{\min(m,n)} \sum_{\langle a\rangle\in\Gamma_r\langle m^{(u)}\rangle} \sum_{\langle b\rangle\in\Gamma_r\langle n^{(v)}\rangle} \det\left(X_{\langle u\rangle\oplus\langle a\rangle,\langle v\rangle\oplus\langle b\rangle}\right)\det\left(Y_{\langle v\rangle\oplus\langle b\rangle,\langle u\rangle\oplus\langle a\rangle}\right) \tag{A.7}$$

where $\langle x^{(y)}\rangle$ signifies the sequence $\langle x\rangle$ with the first y elements omitted, is equal to

$$(-1)^v \det\left(I^{(u+v)} + \begin{pmatrix} 0 & Y_{r.v} \\ X_{c.v} & X^{(c.v)}Y^{(r.v)}\end{pmatrix}\right) = (-1)^u \det\left(I^{(u+v)} + \begin{pmatrix} 0 & X_{r.u} \\ Y_{c.u} & Y^{(c.u)}X^{(r.u)}\end{pmatrix}\right) \tag{A.8}$$

where subscript $r.\ z, c.\ z$ specifies the rows or columns of a submatrix by inclusion of the index sequence z, and the superscript $(r.\ z), (c.\ z)$ specifies the rows or columns of a submatrix by omission of the index sequence z.

Proof. By applying the modified Cauchy-Binet formula [20, App. C] we can contract either of the two sequences $\langle a\rangle, \langle b\rangle$. Contracting sequence $\langle a\rangle$ yields

$$= (-1)^v \sum_{r=0}^{\min(m,n)} \sum_{\langle a\rangle\in\Gamma_r\langle(m+v)^{(u+v)}\rangle} \det\left(M_{\langle u+v\rangle\oplus\langle a\rangle,\langle u+v\rangle\oplus\langle a\rangle}\right) \quad \text{where} \quad M = \begin{pmatrix} 0 & Y_{\mathrm{r}.v} \\ X_{\mathrm{c}.v} & X^{(\mathrm{c}.v)}Y^{(\mathrm{r}.v)}\end{pmatrix} \tag{A.9}$$

Applying lemma A.1.1 directly yields the first result in the theorem. Contracting sequence $\langle b\rangle$ first and then applying the lemma yields the second result in the theorem. The proof is thus finished.

Lemma A.1.3 *Consider the generalisation of A.1.2, where the transpose of X no longer has the same dimensions as Y, and where the constraints on row/column inclusion for the two minors in each term of the sum are asymmetrical:*

$$S = \sum_{r} \sum_{\langle a\rangle\in\Gamma_{r_a}\langle m\rangle} \sum_{\langle b\rangle\in\Gamma_{r_b}\langle n\rangle} \det\left(X_{u_x\oplus\langle a\rangle,v_x\oplus\langle b\rangle}\right)\det\left(Y_{v_y\oplus\langle b\rangle,u_y\oplus\langle a\rangle}\right) \tag{A.10}$$

where $u_x + v_y = v_x + u_y$ and $v_x \leq v_y$, r_a and r_b are taken such that the submatrices in each term are square, the sum over r includes all possible square minors of X, Y which satisfy the constraints, and where $D_{k\oplus\langle l\rangle,m\oplus\langle n\rangle}$ is the smallest submatrix of D which contains the upper left block of D with shape (k, m), as well as the submatrix of the lower right block of D given by the row index sequence $\langle l\rangle$ and column index sequence $\langle n\rangle$. The sum evaluates to

$$S = (-1)^{v_y(1+v_y-v_x)} \det\left(I^{(u_x+v_y)} + \begin{pmatrix} 0_{v_y,v_x} & Y_{r.v_y} \\ X_{c.v_x} & X^{(c.v_x)}Y^{(r.v_y)}\end{pmatrix}\right) \tag{A.11}$$

Proof. We start with a few observations. Firstly, the constraint $u_x + v_y = v_x + u_y$ is not arbitrary, and it is in fact necessary for the sum to be constructable, which can be seen from inspecting the dimensions of X, Y–this also makes the choice of r_a, r_b, and the summation limits of r unique. Secondly, the notation $D_{k\oplus\langle l\rangle, m\oplus\langle n\rangle}$ can be thought of as $D_{\langle x\rangle,\langle y\rangle}$, where $\langle x\rangle$ is formed from indices $1\dots k$ concatenated with the sequence $\langle l\rangle$ where each element was increased by k, with construction of $\langle y\rangle$ being analogous. Thirdly, the fact only one of two possible contractions is presented follows from the minimal assumption $v_x \leq v_y$, which loses no generality. Should the opposite be true, relabelling X, Y and $\langle a\rangle, \langle b\rangle$ reduces the problem to its original form; should we instead specify $u_x \leq u_y$, relabelling X, Y *or* $\langle a\rangle, \langle b\rangle$ once again reduces the problem to the statement above.

Denote $\Delta v = v_y - v_x$. Consider the matrix $X' = I_{\Delta v} \oplus X$, i.e.

$$X' = \begin{pmatrix} I_{\Delta v} & \\ & X \end{pmatrix} \quad \text{hence} \quad \det\big(X'_{u_x+\Delta v\oplus\langle a\rangle, v_x+\Delta v\oplus\langle b\rangle}\big) = \det\big(X_{u_x\oplus\langle a\rangle, v_x\oplus\langle b\rangle}\big) \tag{A.12}$$

We can use the modified Cauchy-Binet formula to contract $\langle b\rangle$ like so:

$$S = (-1)^{v_y} \sum_{r=0}^{\min(m,n)} \sum_{\langle a\rangle\in\Gamma_{r_a}\langle m\rangle} \det\big(M_{u_x+v_y\oplus\langle a\rangle, u_y+v_y\oplus\langle a\rangle}\big) \quad \text{where} \quad M = \begin{pmatrix} 0_{v_y,v_y} & Y_{\mathrm{r}.v_y} \\ X'_{\mathrm{c}.v_y} & X'^{(\mathrm{c}.v_y)}Y^{(\mathrm{r}.v_y)} \end{pmatrix} \tag{A.13}$$

We now observe that

$$X'_{\mathrm{c}.v_y} = \begin{pmatrix} I_{\Delta v} & 0_{\Delta v, v_x} \\ 0_{v_x,\Delta v} & X_{\mathrm{c}.v_x} \end{pmatrix} \tag{A.14}$$

and since the first Δv rows of $X'^{(\mathrm{c}.v_y)}$ are zero, we also have

$$X'^{(\mathrm{c}.v_y)}Y^{(\mathrm{r}.v_y)} = \begin{pmatrix} 0_{\mathrm{r}.\Delta v} \\ X^{(\mathrm{c}.v_x)}Y^{(\mathrm{r}.v_y)} \end{pmatrix} \tag{A.15}$$

Taking the Laplace expansion of M for all $(\Delta v, \Delta v)$ minors along the first Δv columns, there is only one non-zero contribution, which is $\det(I_{\Delta v}) = 1$. Considering its cofactor for any arbitrary constrained minor of M, we can rewrite the sum as

$$S = (-1)^{v_y(1+\Delta v)} \sum_{r=0}^{\min(m,n)} \sum_{\langle a\rangle\in\Gamma_{r_a}\langle m\rangle} \det\Big(M'_{u_x+v_y\oplus\langle a\rangle, v_x+u_y\oplus\langle a\rangle}\Big) \quad \text{where} \quad M' = \begin{pmatrix} 0_{v_y,v_x} & Y_{\mathrm{r}.v_y} \\ X_{\mathrm{c}.v_x} & X^{(\mathrm{c}.v_x)}Y^{(\mathrm{r}.v_y)} \end{pmatrix} \tag{A.16}$$

Since $u_x + v_y = v_x + u_y$, this sum can be reducing by direct application of Lemma A.1.1, which finishes the proof.

Note. We shall now explicitly state the construction of r_a, r_b and the summation limits on r. Since we demand the submatrices of X, Y be square, we have

$$u_x + r_a = v_x + r_b = r \geq 0 \tag{A.17}$$
$$u_y + r_a = v_y + r_b = r \geq 0 \tag{A.18}$$

To satisfy both equations and inequalities, we take

$$\begin{aligned} r_a = r - \min(u_x, u_y), \qquad r_b = r - \min(v_x, v_y), \\ r = \max(\min(u_x, u_y), \min(v_x, v_y)) \ldots \min(X_{\text{rows}} - u_x + \min(u_x, u_y), X_{\text{cols}} - v_x + \min(v_x, v_y)) \end{aligned} \tag{A.19}$$