Reduced Density Matrix Cumulants: The Combinatorics of Size-Consistency and Generalized Normal Ordering

Jonathon P. Misiewicz, Justin M. Turney, and Henry F. Schaefer III*

Center for Computational Quantum Chemistry, University of Georgia, Athens, Georgia, 30602

E-mail: ccq@uga.edu

Abstract

Reduced density matrix cumulants play key roles in the theory of both reduced density matrices and multiconfigurational normal ordering, but the underlying formalism has remained mysterious. We present a new, simpler generating function for reduced density matrix cumulants that is formally identical to equating the coupled cluster and configuration interaction ansätze. This is shown to be a general mechanism to convert between a multiplicatively separable quantity and an additively separable quantity, as defined by a set of axioms. It is shown that both the cumulants of probability theory and reduced density matrices are entirely combinatorial constructions, where the differences can be associated to changes in the notion of "multiplicative separability" for expectation values of random variables compared to reduced density matrices. We compare our generating function to that of previous works and criticize previous claims of probabilistic significance of the reduced density matrix cumulants. Finally, we present the simplest proof to date of the Generalized Normal Ordering formalism to explore the role of reduced density matrix cumulants therein.

1 Introduction

Reduced density matrix cumulants are fundamental in both reduced density matrix (RDM) theories and multireference theories that use the generalized normal ordering formalism (GNO) of Kutzelnigg and Mukherjee.[1–5] To RDM theories, RDM cumulants are the size-consistent parts of the RDMs. This is one of the primary reasons why cumulants are either parameterized or varied directly in many RDM-based theories.[6–11] In GNO, second quantized operators are decomposed into linear combinations of operators "normal ordered" with respect to an arbitrary reference, Ψ, via the Generalized Wick's Theorem. This theorem gives the expansion coefficients of the linear combination in terms of contractions. This is analogous to the normal ordering procedure and contractions familiar from correlated single-reference wavefunction theory.[12,13] However, in the single-reference theory, the contractions are Kronecker deltas. In GNO, the contractions are Kronecker deltas and also the RDM cumulants of Ψ. The GNO formalism has also been used in many studies.[14–16]

Broadly speaking, there have been three approaches to defining reduced density matrix cumulants in the literature. The first definition is an explicit formula for them in terms of reduced density matrices.[17–19] Apart from one presentation of the two-body cumulant,[20] this presentation is *ad hoc*, and the connection to size-consistency is not established. The second definition begins by identifying the connected components of the perturbation expansion of the n-particle propagators.[21] The terms can then be related to terms of a perturbation expansion of the reduced density matrices,[22] and the size-consistent terms isolated. This definition was nearly immediately replaced with the third definition. The third definition is based on Kubo's presentation of cumulants in probability theory[23] and is now the exclusive formalism used to discuss size-consistency of the RDM cumulants.[20,24–28]

Given random variables $X_1, ..., X_n$, Kubo began by defining the moment generating function

$$M(t_1, ..., t_n) = \langle \exp(\sum_{i=0}^{N} t_i X_i) \rangle \tag{1}$$

and the cumulant generating function

$$K(t_1, ..., t_n) = \log M(t_1, ..., t_n) \quad . \tag{2}$$

Any moment, or expectation value, of a product of the variables can be written in the form $\langle \prod_{n=0}^{N} X_n^{i_n} \rangle$. That moment is the coefficient of $\prod_{n=0}^{N} (i_n!)^{-1} t_n^{i_n}$ in (1). Kubo defined the cumulant of random variables, which we call the probabilistic cumulant, as the coefficient of $\prod_{n=0}^{N} (i_n!)^{-1} t_n^{i_n}$ in (2). Kubo showed that the probabilistic cumulants so defined are "additively separable" with respect to variables that are "multiplicatively separable." Specifically, Kubo defined sets of random variables as being statistically independent if any moment of variables factors into a product of moments, one for each set. For example, if the sets $\{X\}$ and $\{Y, Z\}$ are statistically independent, then $\langle X^2 Y Z \rangle = \langle X^2 \rangle \langle Y Z \rangle$. Then any cumulant of variables from multiple independent sets is zero. This is the probabilistic analogue of the fact that any coupled cluster amplitude with orbital indices from multiple noninteracting systems is zero.

To adapt Kubo's definition of probabilistic cumulants to a definition of RDM cumulants, we must change expectation values of random variables to expectation values of second quantized operators. However, Kubo's proof of additive separability assumed that the random variables commute, but our second quantized operators do not. To define cumulants of non-commuting objects while keeping additive separability, Kubo proposed that the multiplication appearing in the power series of exp and log from (1) and (2) be replaced with a "multiplication" which does make the objects commute. This idea has been key in the third approach to defining RDM cumulants, via a generalization of Kubo's generating functions.

We are convinced that the current definitions of RDM cumulants from generating function have left important points unclear, and that these points leave hindered broader use of cumulants among electronic structure theorists:

1. It is not *a priori* obvious what the "correct" adjustment to the definition of multiplica-

tion in the exp and log series is. This has led to two distinct schemes to adapt Kubo's cumulants to RDM cumulants. Is there another way to generalize Kubo's approach? The complexity of both definitions makes a simpler definition desirable:

(a) The approach dominant in generalized normal ordering literature was pioneered by Kutzelnigg and Mukherjee[25] and later refined by Hanauer and Köhn.[24] Accordingly, we call it the KMHK approach. In this formalism, the analogue of random variables are the particle-conserving operators a^p_q, and exp and log must be redefined to use a modified normal order product. While this modified normal ordering works, its significance is unclear. Further, the presentation of Hanauer and Köhn uses six different product operations: the Grassmann product ($\wedge$), the alternative Grassmann product ($\otimes$), the normal order product ($\{\}$), the scalar product of tensors ($\cdot$), the antisymmetrized tensor product ($\times_A$), and the modified normal order product ($\{\}'$).

(b) The approach dominant in reduced density matrix literature was pioneered by Mazziotti.[26] In this formalism, the analogue of random variables are the creation and annihilation operators $a^\dagger_p$ and a_q, and the exponential is modified by applying an "ordering" operator. The analogue of the "formal variables", $t_1, ..., t_n$, are neither real nor complex numbers, but anticommuting numbers. Throughout the literature, it has been typical[20,26,27,29–32] to obtain RDM cumulants by differentiating the exponentiated analogue of (2), rather than using the log series. The required differentiation operators also anticommute. Furthermore, an n-electron RDM cumulant in this formalism is related not to the n-variable probabilistic cumulant, as in the KMHK approach, but to the $2n$-variable probabilistic cumulant, before reducing to a formula reminiscent of the n-variable probabilistic cumulant.

2. It remains opaque why RDM cumulants should be size-consistent, that is, why RDM cumulants with indices in noninteracting subsystems should vanish. This is so for two

reasons. First, the proof that the RDM cumulants are size-consistent using the KMHK definition is more complicated than in Kubo's case, because not all "random variables" can be assigned to one subsystem or the other.[24] (Mazziotti's approach does not have this drawback.) Second, when the formula for a cumulant is simplified to a polynomial, it is not clear why one polynomial is size-consistent while another is not. For example, why is $\gamma^{pq}_{rs} - \gamma^p_r \gamma^q_s + \gamma^p_s \gamma^q_r$ size-consistent but not $\gamma^{pq}_{rs} + \gamma^p_r \gamma^q_s - \gamma^p_s \gamma^q_r$? Neither generalization of Kubo's approach immediately offers insight.

3. It is an open question whether RDM cumulants of arbitrary rank have some further probabilistic interpretation, due to their similarity to the probabilistic cumulants. Kutzelnigg and Mukherjee tried to offer such an interpretation[25] but later said it did not apply to the "exclusion-principle violating" cumulants.[33] Kong and Valeev interpreted some RDM cumulant elements as probabilistic correlations of electron occupation, within some restrictive assumptions.[19] Hanauer and Köhn gave the same interpretation with looser assumptions, but could still not provide a definitive probabilistic interpretation for all RDM cumulant elements.[24] The latter paper was explicitly motivated by trying to understand the analogy between probabilistic cumulants and RDM cumulants.

4. We are aware of no attempt to explain why the RDM cumulants defined via this generating function should appear in GNO at all.

In this research, we propose a more intuitive definition of the RDM cumulants that starts not from the cumulant generating function of Kubo but by three axioms, inspired by Percus,[34] that characterize a solution to the general problem of breaking a multiplicatively separable second quantized quantity into additively separable parts. Accordingly, the relation between RDM elements and their cumulants is precisely the same as the relation between configuration interaction amplitudes and the coupled cluster amplitudes because both coupled cluster amplitudes and RDM cumulants obey the same three strict axioms.

This allows us to provide a "generating function" for the RDM cumulants that is simpler than the functions from the KMHK and the Mazziotti approaches, and can also be trivially adapted to construct an additively separable quantity from any multiplicatively separable one. All this will be covered in Section 2.

In Section 3, we shall compare our generating function with that of the KMHK approach and the Mazziotti approach to analyze how they generalize the idea of Kubo, and how all three generating functions can lead to the same answer. Section 3.1 shall review generating functions in detail. Section 3.2 will analyze the use of generating functions in the definition of the probabilistic cumulant. We intend to establish that the probabilistic cumulant obeys very similar axioms to the RDM cumulant, with the crucial difference of what they mean by "multiplicative separability," and discuss the implications of that difference for the generating functions that the two cumulants need. In Section 3.3, we discuss how the previous RDM cumulant generating functions of the KMHK and Mazziotti approaches simplify to ours and lead to the same answer. By this point in our argument, it will be clear that the analogy between the probabilistic and RDM cumulants is entirely a matter of combinatorics and the three axioms, and probability theory plays no role in the analogy. In Section 3.4, we shall criticize claims of a probabilistic interpretation of RDM cumulants. We refer readers interested in a detailed look at the connection between our generating function and the combinatorial problem of the three axioms to Appendix A.

Lastly, we shall consider why cumulants appear in the formalism of generalized normal ordering in Section 4 and give a proof of its various Generalized Wick Theorems that is far simpler than any prior proof in the literature.

2 Additive Separability from Multiplicative Separability

Suppose some family of antisymmetric tensors exists where each element is indexed by n creation operators and n annihilation operators, and the family contains tensors of variable n.

Further suppose the orbitals can be partitioned into subsets, called blocks, such that the orbitals of each block factor each tensor into a product of tensors, each indexed by the orbitals of a single block, multiplied by the usual permutational sign factor. For example, if p, q, s and t are in one block and the orbitals r and u are in another block, then $z^{pqr}_{stu} = z^{pq}_{st} z^{r}_{u}$ and $z^{rq}_{tu} = -z^{q}_{t} z^{r}_{u}$ and $z^{r}_{t} = 0$, and so on. If this occurs, we say the family of tensors is multiplicatively separable with respect to the partition.

There are two archetypal examples of multiplicatively separable families of tensors. In both, the relevant partitioning of orbitals into blocks is the partitioning of orbitals on non-interacting subsystems.

The first example is the exact configuration interaction amplitudes in intermediate normalization, i.e., when the reference determinant has overlap 1 with the full wavefunction. This is shown by the argument in Section 4.3.1 of Reference 35, upon recognizing that the product of intermediately normalized excitation operators is also an intermediately normalized excitation operator.

The second example is the reduced density matrix elements using the McWeeny normalization convention,[36] in which all RDM elements are just the expectation values of the corresponding second quantized operator, such as $\gamma^{pq}_{rs} = \langle \Psi | \, a^{pq}_{rs} \, | \Psi \rangle$. It may be shown, by tedious but straightforward anticommutation of second quantized operators of different systems, that the RDMs are multiplicatively separable.

We want to construct a family of tensors that will be *additively separable* with respect to a partition of orbitals if we are given a family of tensors that is *multiplicatively separable*

with respect to that same partition. Additive separability means that the tensor for the system with orbitals from all blocks of the partition is the sum of the tensors with orbitals from only a single block of the partition. In the usual case where blocks contain orbitals of different subsystems, this means that to obtain the additively separable tensor for a system of noninteracting subsystems, add the tensors from each non-interacting subsystem. The experience of quantum chemists has been that formulating an approximate theory in terms of additively separable quantities is crucial for maintaining the accuracy of a theory applied beyond the smallest systems.[12,13,35]

We can construct this manually for low-rank tensors. The key idea is to include products of tensors such that after applying any possible factorization, all products of tensors either have coefficient zero or vanish. For example, consider $z^{pq}_{rs} - z^p_r z^q_s + z^p_s z^q_r$. If p and r are on one subsystem and q and s are on a different noninteracting subsystem, the polynomial factors to $z^p_r z^q_s - z^p_r z^q_s + 0 * 0 = 0$. Likewise, if p and s are on a different subsystem from q and r, the polynomial becomes $-z^p_s z^q_r + 0 * 0 + z^p_s z^q_r = 0$. This is the general mechanism by which a polynomial of a multiplicatively separable quantity acquires additive separability.

With additional notation, we can condense our observations into axioms that dictate what we require of our family of tensors, $\mathcal{T}$, that will be additively separable with respect to an orbital partition if the tensor $\mathcal{C}$ is multiplicatively separable with respect to it. If $\mathcal{Z}$ is a family of tensors and $\mathcal{S}$ is an ordered set of n creation operators and n annihilation operators, $\mathcal{Z}(\mathcal{S})$ is the tensor in $\mathcal{Z}$ indexed by the orbitals of $\mathcal{S}$. The set of possible factorization patterns of a $\mathcal{Z}(\mathcal{S})$, called the set of fermionic partitions, is written as $\Pi_{\mathcal{F}}(\mathcal{S})$. Given fermionic partition ρ with jth block ρ_j, $\mathcal{Z}(\rho)$ is $\prod_j \mathcal{Z}(\rho_j)$. We regard $\mathcal{S}$ as the fermionic partition of only one block.

Our requirements of the family of tensors, $\mathcal{T}$, which must be additively separable with respect to any partition that makes $\mathcal{C}$ multiplicatively separable, are:

1. **Functional Form**

$$\mathcal{T}(\mathcal{S}) = \sum_{\rho:\rho\in\Pi_{\mathcal{F}}(\mathcal{S})} \mu_\rho \mathcal{C}(\rho) \tag{3}$$

2. **Normalization** $\mu_{\mathcal{S}} = 1$, where again, $\mathcal{S}$ is the fermionic partition consisting of one block

3. **Additive Separability** If $\mathcal{C}$ is multiplicatively separable with respect to an orbital partition of $\mathcal{S}$ (other orbitals not being relevant), $\mathcal{T}(\mathcal{S})$ is identically zero, for any choice of the remaining $\mathcal{C}(\rho)$.

These axioms closely mimic the axioms applied to probabilistic cumulants by Percus[34] and latter refined by Simon.[37]

To obtain an explicit formula for $\mathcal{T}(\mathcal{S})$, we need only determine the coefficients μ of (3) using the last two axioms. Further, there is at most one solution, and it is independent of the values of $\mathcal{C}(\rho)$. This holds because $\mu_{\mathcal{S}}$ is determined by the normalization axiom, and all other coefficients may be determined by the following induction over the number of blocks: Consider fermionic partition π other than $\mathcal{S}$. Suppose that each block defines orbitals of an independent system. Factorize (3) accordingly, and the coefficient of $\mathcal{C}(\pi)$ is $\sum_{\rho:\rho\supseteq\pi} \mu_\rho$, where $\rho \supseteq \pi$ means that each block of ρ is contained in a block of ρ, as shown in Figure 1. It does not compare numerical values. For $\mathcal{T}(\mathcal{S})$ to be zero for any choice of the $\mathcal{C}(\rho)$, as required by the third axiom, we need

$$\sum_{\rho:\rho\supseteq\pi\neq\mathcal{S}} \mu_\rho = 0 \quad . \tag{4}$$

But all terms other than μ_π correspond to fermionic partitions with fewer blocks than π and thus are already known.

While it is possible to give a closed-form solution for the coefficients using low-level combinatorics, as we show in Appendix A, there is a far more accessible solution. The coupled cluster amplitudes satisfy all three axioms. The coupled cluster and configuration interaction amplitudes are related by

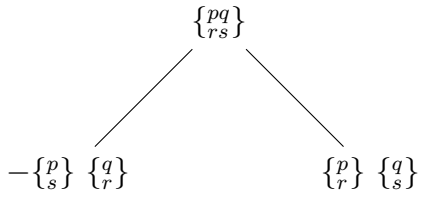

Figure 1: The fermionic partitions for a rank-two tensor. A vertical path between two elements ρ and π, where ρ is higher than π, means that $\rho \supseteq \pi$.

$$1 + C = \exp(T) \tag{5}$$

where

$$C = \sum_{i,a} \frac{1}{(1!)^2} c_a^i a_i^a + \sum_{i,j,a,b} \frac{1}{(2!)^2} c_{ab}^{ij} a_{ij}^{ab} + ... \tag{6}$$

and

$$T = \sum_{i,a} \frac{1}{(1!)^2} t_a^i a_i^a + \sum_{i,j,a,b} \frac{1}{(2!)^2} t_{ab}^{ij} a_{ij}^{ab} + ... \tag{7}$$

and the operators a_i^a, a_{ij}^{ab}, etc. are the usual second quantized excitation operators of many-fermion theory.[12,13,35]

The excitation operators in (6) and (7) perform two roles. First, they make the left and right hand sides of (5) operators that transform the reference Φ into the target state Ψ.[38] The need for an operator to act on a wavefunction is the usual rationale for the appearance of second quantized operators in (5).[12,13,35,38] For our purposes, this role is irrelevant. Second, equating the coefficients of the operators a_i^a, a_{ij}^{ab}, etc. on each side of (5) gives a c amplitude as a polynomial in the t amplitudes.[38,39] In this second role, (5) solves our problem of additive separability for the configuration interaction and coupled cluster amplitudes. By the previous arguments, this is the one and only solution.

Let us adapt this to solve the additive separability problem for any multiplicatively separable quantity. Define

$$\mathcal{C} = \sum_{p,q} \mathcal{C}\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) a_q^p + \sum_{p,q,r,s} \frac{1}{(2!)^2} \mathcal{C}\left(\begin{smallmatrix} pq \\ rs \end{smallmatrix}\right) a_{rs}^{pq} + ... \tag{8}$$

and

$$\mathcal{T} = \sum_{p,q} \mathcal{T}\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) a_q^p + \sum_{p,q,r,s} \frac{1}{(2!)^2} \mathcal{T}\left(\begin{smallmatrix} pq \\ rs \end{smallmatrix}\right) a_{rs}^{pq} + ... \quad . \tag{9}$$

We may attempt to use $1 + \mathcal{C} = \exp(\mathcal{T})$, but our second quantized operators need not commute, so we lose the property that $\exp(A + B) = \exp(A)\exp(B)$, which plays a central role in the logic that the cluster operators are additively separable.[12,13,35] Modifying an idea from Lindgren,[40] we redefine the multiplication in the exponential to be the vacuum-normal order product rather than the operator product; so for example, we use the multiplication $\{a_r^p a_s^q\} = a_{rs}^{pq}$ rather than $a_r^p a_s^q = a_{rs}^{pq} + \delta_r^q a_s^p$. (In notation such as $\{a_r^p a_s^q\}$, the braces denote redefining multiplication, not a function applied to $a_r^p a_s^q$. The latter approach leads to contradictions of the type discussed in References 41 and 42.) The normal product is always commutative for particle-conserving operators and reduces to the usual exponential when we only need excitation operators, as in coupled cluster. Normal ordered exponentials also appear explicitly in the KMHK approach to cumulants[24,25,28,43,44] and as the ordering operator in the Mazziotti approach to cumulants.[20,26,27,29–32] However, both second quantized operators and normal ordering appear for a different reason in our formalism compared to theirs, as we discuss in Section 3.3.

Therefore, the solution to our additive separability problem is given by

$$1 + \mathcal{C} = \{\exp(\mathcal{T})\} \tag{10}$$

or equivalently

$$\{\log(1+\mathcal{C})\} = \mathcal{T} \quad . \tag{11}$$

We have used the fact that the logarithm and exponential are inverses as long as the product operation commutes because they must be inverses as formal power series.[45–47]

It is easy to confirm from (11) that $\mathcal{T}$ satisfies the first two axioms, and it is easy to confirm from (10) that $\mathcal{C}$ satisfies the third axiom because

$$1 + \mathcal{C}_{A+B} = \{(1+\mathcal{C}_A)(1+\mathcal{C}_B)\} = \{\{\exp(\mathcal{T}_A)\}\{\exp(\mathcal{T}_B)\}\} = \{\exp(\mathcal{T}_A + \mathcal{T}_B)\} \quad . \tag{12}$$

It is also possible to confirm the additive separability of a given explicit formula derived from (11) by performing any possible factorization (which corresponds to a way to divide orbitals into multiplicatively separable groups) and observing that the resulting polynomials are identically zero. A derivation of the cumulants based on this is given in Appendix A.

We have thus solved the problem of converting *any* multiplicatively separable quantity into an additively separable one. Merely equate coefficients of the second quantized operators on both sides of (11), using the definitions (8) and (9) and the power series of the logarithm. This applies not only to converting configuration interaction amplitudes to coupled cluster amplitudes and to converting reduced density matrices to their cumulants, but also to more exotic quantities, such as the reduced transition matrices of Mazziotti[30,31] or the amplitudes of valence universal multireference coupled cluster.[42] The coefficients of (3) will be exactly the same in any case.

We can give yet more explicit formulas because the Taylor series expansion coefficients of exp and log are known, we may replace the $\frac{1}{n!}^2$ appearing in (8) and (9) by choosing one permutation of any given set of indices, and the $\frac{1}{k!}$ appearing in the degree k term of the exp and log series precisely cancels out the $k!$ ways to permute the $\mathcal{C}$ or $\mathcal{T}$ operators to get identical results. Denoting the number of blocks of partition ρ with $\#\rho$, we find:

$$\mathcal{C}(\mathcal{S}) = \sum_{\rho:\rho\in\Pi_{\mathcal{F}}(\mathcal{S})} \mathcal{T}(S) \tag{13}$$

and

$$\mathcal{T}(\mathcal{S}) = \sum_{\rho:\rho\in\Pi_{\mathcal{F}}(\mathcal{S})} (-1)^{\#\rho-1}(\#\rho-1)!\mathcal{C}(S) \quad . \tag{14}$$

We emphasize that the additive separability of $\mathcal{T}$ with respect to a partition of orbitals is not automatic, even if the orbitals can be divided into noninteracting subsystems. It is automatic if $\mathcal{C}$ is multiplicatively separable. So for example, the RDM cumulant of an approximate theory may fail to be additively separable if the RDM lacks the correct multiplicative separability. We believe this is what Kong and Valeev meant when remarking that the additive separability of the RDM cumulant "is not guaranteed for arbitrary wavefunctions."[19] Examples of this behavior, even for functions with size-consistent energies, include the orbital unrelaxed density matrices of coupled cluster,[48] the orbital optimized methods studied by Bozkaya and co-workers,[49–52] and the RDM formulation of CEPA given by Mazziotti and related to his parametric RDM method.[53,54]

There have been previous attempts to connect reduced density matrix cumulants and coupled cluster,[18,30,31,48,55] but we are aware of none that recognized that near identical "generating functions" can be produced for the two, or that this is a general solution to the problem of converting between multiplicative and additive separability.

3 Generating Functions

Equations (10) and (11) provide a way to construct a multiplicatively separable quantity from an additively separable one and vice versa. However, we have not yet established why a function should be so useful in solving what is an essence a combinatorial problem, how the differences between RDM cumulants and the probabilistic cumulants should be understood,

how the difference between our generating functions and those of the KMHK and Mazziotti approaches should be understood, or what this means for efforts to interpret RDM cumulants probabilistically. We address each of these questions in turn in the following subsections.

3.1 Mathematicians' Generating Functions

While RDM cumulant generating functions have been defined numerous times,[20,24–29,32,43,44] as have generating functions for the more general reduced transition matrix cumulants,[30,31] we are aware of no general discussion of generating functions in the chemistry literature. As this is crucial for this research, we provide one, emphasizing the underlying ideas in language accessible to quantum chemists rather than mathematical rigor. We refer readers interested in detailed mathematical treatments of generating functions to References 46, 56, 47, 57, and 58.

Combinatorialists frequently study arrays of numbers indexed by n natural numbers. For example, $a_{v,c}$ may count the graphs with v vertices and c connected components. This sequence may be encoded into a formal power series in n variables. A formal power series is a power series where the variables are associative and commutative, but are otherwise indeterminate. These are called formal variables. Formal variables cannot be evaluated at specific numbers, and accordingly, questions of convergence do not exist. The formal power series that a sequence is converted into is called a generating function.

Although generating functions have multiple uses, the one most relevant to the present work is that they convert combinatorial problems into algebraic ones. It is possible to define algebraic operations on formal power series that replicate familiar operations on functions and that also automate some combinatorially significant operation on the sequence. We can thus solve a problem algebraically and only afterwards rephrase the result in terms of the original combinatorial problem.

Let us illustrate the combinatorial significance of the familiar algebraic operation of multiplying functions of one variable. Suppose the sequence $\{a_n\}$ is encoded into a function

by $a(x) = \sum_{n=0}^{\infty} a_n x^n$. This construction is called an ordinary generating function. If $b(x)$ is the ordinary generating function of $\{b_n\}$, then term-by-term multiplication of $a(x)$ and $b(x)$ gives $\sum_{n=0}^{\infty}(\sum_{m=0}^{n} a_m b_{n-m})x^n$. Multiplying two ordinary generating functions creates a new sequence where term n is $\sum_{m=0}^{n} a_m b_{n-m}$. Combinatorially, this corresponds to combining a "degree m" object of the first sequence and a "degree $n-m$" object of the second sequence into a new "degree n" object. The summation exists because there are multiple ways to get a "degree n" object.

Alternatively, suppose the sequence $\{a_n\}$ was encoded into $a(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$, and likewise to produce $b(x)$ from $\{b_n\}$. This produces an exponential generating function. Term-by-term multiplication gives $\sum_{n=0}^{\infty}(\sum_{m=0}^{n} \binom{n}{m} a_m b_{n-m})\frac{1}{n!}x^n$. Multiplication of exponential generating functions thus means that term n of the new sequence is $\sum_{m=0}^{n} \binom{n}{m} a_m b_{n-m}$. This is commonly used in problems where not only do a "degree m" object and "degree $n-m$" object combine to produce a "degree n" object, but every one of the $\binom{n}{m}$ ways to assign the degrees of the "degree n" object to the original two objects produces a distinct object.

These ideas can be extended to a sequence d indexed by n natural numbers, d_1 through d_n. There are n formal variables x_1 through x_n, and the generating functions are written as $a(x_1, ..., x_n) = \sum_d a_d \prod_{i=1}^{\infty} \frac{x_i^{d_i}}{d_i!}$ for an exponential generating function and $a(x_1, ..., x_n) = \sum_d a_d \prod_{i=1}^{\infty} x_i^{d_i}$ for an ordinary generating function.

3.2 Probabilistic Cumulants

We are now prepared to address the probabilistic cumulants.

We require notation that mirrors our notation for the fermionic case. Given a set, $\mathcal{S}$, of n random variables, we define $\Pi(\mathcal{S})$ as the set of partitions, all ways to divide the variables into subsets (blocks) such that each variable is in exactly one set. κ refers to the set of cumulants, and m refers to the set of moments. $\kappa(\mathcal{S})$ refers to the cumulants using the variables of set $\mathcal{S}$, and $m(p)$ is the product of moments corresponding to each block of partition p.

Once again, we can define cumulants following the axioms of Percus[34] and Simon.[37]

1. **Functional Form**

$$\kappa(\mathcal{S}) = \sum_{p:p\in\Pi(\mathcal{S})} \mu_p m(p) \tag{15}$$

2. **Normalization** $\mu_{\mathcal{S}} = 1$, where again, $\mathcal{S}$ is the partition consisting of one block

3. **Connectedness** If m is multiplicatively separable with respect to a partition of $\mathcal{S}$ (other variables not being relevant), $\kappa(\mathcal{S})$ is identically zero, for any choice of the remaining $m(p)$.

Compared to the axioms that apply to RDM cumulants, the axioms for probabilistic cumulants differ in two primary ways. First, our additively separable quantity is no longer a linear combination of antisymmetric tensors, but of moments of our variables. Second, the set of partitions that our multiplicatively separable quantities can separate into have changed from fermionic partitions to the usual set partitions.

The corresponding cumulant generating function is given by (2). Again, we may check that when all variables in the moment are distinct, the cumulants so defined satisfy all three of the axioms.

Let us try to understand why (2) solves our combinatorial problem. First, the appearance of log is unsurprising, as we are converting a quantity that is multiplicatively separable with respect to a partition to a quantity that will be additively separable with respect to that partition. But unlike (11), we do not need to modify the multiplication appearing in the power series of the logarithm.

We can rationalize this as follows: To read off relations of form (3) and (15) from our generating functions, we need both sides of the equations to be coefficients of the same monomial. In both cases, the right-hand side will arise from repeated multiplication of a function with the multiplicatively separable quantities as coefficients. How should we "multiply" this function by itself to get the collection of terms we want?

The multiplication differs for these two cases because the partitions we want to sum differ for these two cases. For probabilistic cumulants, we sum over set partitions. All we need is to count how many times each variable has been included in our partition, and this is accomplished by attaching the moment $\langle \prod x_i^{n_i} \rangle$ to the monomial $\prod t_i^{n_i}$. Even after taking products, counting the degree of each formal variable tells how many times the random variable appeared, which is all we want. So the multiplication of formal variables as in Section 3.1 is perfectly acceptable, though we have yet to decide whether to use an ordinary or an exponential generating function.

For additively separable fermionic quantities, we sum over fermionic partitions. We need to count how many times each fermion appears as a creation or annihilation index, and also the overall phase factor. For efficiency, we should assume the same number of creation as annihilation operators. It is possible to adapt the formal variable approach to this, and as we shall discuss in Section 3.3, this is exactly what the KMHK and Mazziotti approaches to RDM cumulants do. However, quantum chemists already have a multiplication to count this: the normal ordered product of particle-conserving operators. This is the fundamental reason why the normal ordered product must be used rather than the operator product in equations (10) and (11).

Another obvious difference between the generating functions for the additively separable probabilistic (2) and fermionic (11) quantities is that the probabilistic multiplicatively separable generating function (1) uses an exponential that has no counterpart in the "generating function" for the fermionic multiplicatively separable quantity, (8). This is due to fermionic antisymmetry eliminating a technicality in the probabilistic cumulants.

For probabilities, it is perfectly legitimate to have a moment with a repeated variable, such as the cumulant $\kappa(XX)$. This cannot occur for fermionic quantities, because any "moments" with a repeated creation index or annihilation index must be zero by antisymmetry. We point out that because the fermionic partitions are defined with a sign factor, the formulas for moments and cumulants in terms of each other given by equations (13) and (14) preserve

antisymmetry.

The possibility of repeated variables in a probabilistic cumulant introduces an ambiguity in how we define the probabilistic cumulant. Do we define it by taking the formula for the cumulant given *distinct* variables and substitute in the repeated variables, or do we extract the term from the functions (2) and (1) with the repeated variables? Ideally, both approaches should produce the same polynomial.

Direct computation shows that when all variables are distinct, encoding the moments m as an ordinary generating function or an exponential generating function produces the same polynomial. However, for repeated variables, the two definitions *differ* using the ordinary generating function. For example, the ordinary generating function produces $\kappa(XX) = m(XX) - \frac{1}{2}m(X)m(X)$ and $\kappa(XY) = m(XY) - m(X)m(Y)$.

The remedy for the case of repeated variables is to choose an exponential generating function for the moment and cumulant generating functions. Taking the logarithm of our moment-generating function is then the composition of exponential generating functions. It is a well-known combinatorial fact that this encodes a sum over all set partitions for a single variable. (See Theorem 5.1.4 of Reference 58.) This interpretation hinges on repeated application of the multiplication of exponential generating functions we discussed in Section 3.1. The multivariable generalization of the same argument shows that the use of an exponential generating function maintains the desired sum over partitions structure, whether variables are repeated or not. Thus, we see that the exponential in the moment generating function is only necessary to treat repeated variables, which we do not have in the fermionic case.

We make two final observations about the probabilistic generating function. First, probabilistic concepts played no role in our construction of the cumulant. The generating function we need is determined only by combinatorial considerations of how a moment can separate. Accordingly, any probabilistic interpretation of the probabilistic cumulant must arise from the fact that the cumulant is a polynomial in moments. Second, the needed coefficients μ for the probabilistic cumulant and the cumulants arising for our original fermionic problem

are closely related. The coefficients for the fermionic case obey the recursion (4) where the sum is over fermionic partitions, and the same logic shows that the probabilistic cumulants obey the same recursion, but where the sum is now over set partitions. But once you have chosen a fermionic partition into n blocks, the set of elements appearing in the recursion is exactly the same as those from the probabilistic cumulant of n elements, as demonstrated in Figure 2. Because both cumulants also agree about the base case of 1 blocks, where the recursion does not apply, it follows that the corresponding coefficients are identical.

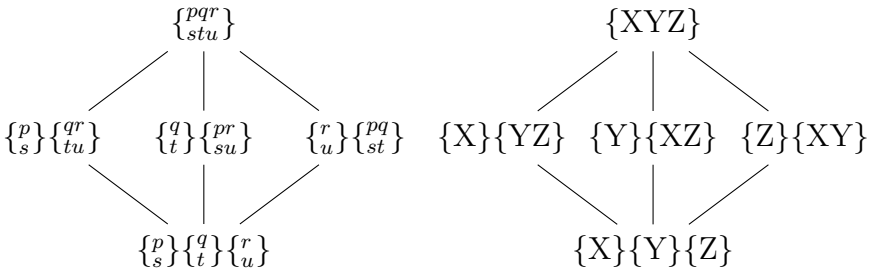

Figure 2: The fermionic partitions greater than or equal to $\{{}^{p}_{s}\}\{{}^{q}_{t}\}\{{}^{r}_{u}\}$, and the set of partitions of three objects. A vertical path between two elements ρ and π, where ρ is higher than π, means that $\rho \supseteq \pi$. The two subsets and their order relations are isomorphic.

3.3 Comparison with Previous Generating Functions

Both the KMHK and Mazziotti approaches acquire their added complexity by sticking too closely to formal variables. Products of formal variables reflect the factorizations of probabilistic cumulants but not fermionic cumulants, so products of formal variables are not an optimal tool for defining fermionic cumulants. We now describe how the concepts of our approach in Section 2 have been hidden in those previous. The comparison is summarized in Table 1.

Both the KMHK and Mazziotti approaches obtain their RDM generating functions by taking the normal ordered exponential of a sum of "minimal" second quantized operators

Table 1: A comparison of different generating functions of reduced density matrix cumulants.

Descriptor	KMHK Approach[a]	Mazziotti Approach[b]	This Work[c]
Moment Generating Function	$\langle\psi\vert\{\exp(k^p_q a^q_p)\}\vert\psi\rangle$	$\langle\psi\vert\{\exp(J_k a^\dagger_p + J^*_k a_p)\}\vert\psi\rangle$	$1+\sum C^{q\cdots}_{p\cdots} a^{p\cdots}_{q\cdots}$
Formal Variable	k^p_q	J_k, J^*_k	a^p_q
Product of Formal Variables	$k^p_q * k^r_s = k^p_q k^r_s \neq -k^p_s k^r_q$	$J_k * J_l = J_k J_l = -J_l J_k$	$\{a^p_q a^r_s\} = a^{pr}_{qs} = -a^{pr}_{sq}$
Particle-conserving variables only?	Yes	No	Yes
Role of a^p_q	Construct RDMs	Construct RDMs	Formal variables
Multiplication in exp/log	$\times_A$	Standard	$\{\}$
Match coefficients of...	Antisym. products of variables	Products of variables	Products of variables
Rank n cumulant needs	n variables	$2n$ variables	n variables

[a] Kutzelnigg, Mukherjee, Hanauer, and Köhn;;[24,25,28,43,44] [b] Mazziotti[26,29–32] and other reduced density matrix investigators;[20,27] [c] Section 2 of the present research

multiplied by formal variables indexed by the "minimal" operator. An expectation value is then taken. This constructs the moment-generating function. For probabilistic moments, where repeated variables exist, this is a very useful device to construct the moment-generating function and much easier to remember than the explicit factorials. However, for our fermionic quantities, their approach is more complicated than our own (8), which is an easy generalization of the familiar configuration interaction form (6). This is exactly as discussed in Section 3.2.

We now consider the two approaches separately.

First is the KMHK approach. As Hanauer and Köhn's presentation uses six different product operations,[24] we comment only on the elements that are apparent from the less detailed presentations of Kutzelnigg and Mukherjee.[25,28,43,44] In the KMHK approach, each formal variable is indexed by both a creation operator and an annihilation operator. This ensures that each term contains the same number of creation operators as annihilation operators. However, different products may be related by antisymmetry. Namely, $k^p_r k^q_s$ and $-k^p_s k^q_r$ both count the same thing. To resolve this, when extracting terms from the generating functions, the KMHK approach matches coefficients of antisymmetrized products of their formal variables, such as $k^p_r k^q_s - k^p_s k^q_r$, instead of simply matching coefficients of the formal variables. Our formalism avoids this entirely because $\{a^p_r a^q_s\} = -\{a^p_s a^q_r\}$. Instead of the normal ordered

logarithm that appears in our formalism, the KMHK approach uses an "antisymmetrized logarithm" to enforce that each product of their formal variables appearing in the Taylor series of $\log(1+X)$ is antisymmetric. In our formalism, this is unnecessary because the formal variables have been replaced with the fermionic second-quantized operators, which are already antisymmetric.

In the Mazziotti approach, each formal variable is indexed by a single operator, creation or annihilation. In that case, the formal variables are ordered in the same way as the creation and annihilation operators used to produce the reduced density matrix. In Mazziotti's moment-generating function, every string of "probe variables" can be replaced with a second-quantized operator to convert to our notation. The anticommutation of the probe variables so $J_p J_q J_s^\dagger J_r^\dagger = -J_q J_p J_s^\dagger J_r^\dagger$ is just the familiar equation in our formalism, $a_{rs}^{pq} = -a_{rs}^{qp}$. That the probe variables are ordered so that the ones associated with creation operators are on the left of those with annihilation operators again is more naturally stated in our formalism as $\{a_r^p a_s^q\} = a_{rs}^{pq}$. In Mazziotti's approach, a traditional exponential is used rather than a normal ordered one. However, the multiplication used by Mazziotti's approach is not that of a second-quantized operator product, but multiplication of formal variables. Further, different orderings of the operator are not treated as distinct, so $J_p J_q J_s^\dagger J_r^\dagger$ and $J_p J_r^\dagger J_q J_s^\dagger$ are treated as the same. This is again the behavior of the more familiar normal ordering our formalism uses, $a_{rs}^{pq} = \{a_r^p a_s^q\}$. However, we reiterate that Mazziotti's formalism generates terms with different numbers of creation and annihilation operators that must eventually vanish. This does not occur in our formalism, which is particle-conserving from the start.

So we see that both previous formalisms can be understood in terms of our simplified cumulants.

3.4 Probabilities and the RDM Cumulant

The arguments of the preceding sections establish that the cumulants are a fundamentally combinatorially entity that describe additive separability from multiplicative separability.

This has different forms for probabilities compared to fermionic quantities because they have different notions of multiplicative separability. Probabilistic cumulants have probabilistic significance only because they are polynomials in expectation values, which themselves have combinatorial significance. Accordingly, we revisit and correct the claims of Hanauer and Köhn[24] that there is a probabilistic interpretation of the RDM cumulant.

Hanauer and Köhn concluded that "in a natural orbital basis, the diagonal elements of λ_n are in fact the covariances of the occupation numbers of n spin orbitals" and stated that a paper by Kong and Valeev[19] made the same conclusion. The actual conclusion of Kong and Valeev was limited to the special cases of λ_2 and λ_3, but conspicuously made no statement for λ_n of higher ranks. For ranks higher than 3, the statement is false.

Hanauer and Köhn correctly claimed that a diagonal RDM element, where the creation and annihilation operators are the same, can be interpreted as the probability that the relevant orbitals are simultaneously occupied. We can thus say $\gamma^p_p = m(p), \gamma^{pq}_{pq} = m(pq)$, and so forth. Let us then take the RDM cumulant, λ, and see if it agrees with the probabilistic cumulant we obtain by regarding RDMs as probabilistic quantities, κ.

For the two-electron case, we have $\lambda^{pq}_{pq} = \gamma^{pq}_{pq} - \gamma^p_p\gamma^q_q + \gamma^p_q\gamma^q_p$ and $\kappa(pq) = m(pq) - m(p)m(q)$. The two cumulants λ and κ disagree by the non-diagonal terms. If we choose our orbitals to be the natural spin orbitals, γ_1 is diagonal by definition, so γ^p_q and γ^q_p vanish, and the two formulas then agree. The same argument shows equality for the λ_3 case. However, for λ_4, the argument fails because the RDM cumulant will contain terms such as $-\gamma^{pq}_{rs}\gamma^{rs}_{pq}$, which cannot be assumed to vanish. The RDM cumulant then disagrees with the probabilistic cumulant of the probabilistic interpretation of the RDM.

This disagreement is unsurprising from the framework of this article. The functional forms for the RDM cumulant, (3), and the probabilistic cumulant, (15), differ precisely by such terms. These represent valid multiplicative separations for fermionic quantities, which have n creation and n annihilation indices, but not for expectation values of variables, which simply have n variables.

Hanauer and Köhn further attempted to give a probabilistic interpretation for off-diagonal RDMs but struggled to make sense of negative RDM elements. The situation is in fact worse. The second quantized operators of off-diagonal RDMs are non-Hermitian. These quantities may be complex numbers, which cannot be a probability. For example, consider the hydrogen atom RDM element, $\langle p_{+1}| \, a^{p_x}_{p_y} \, |p_{+1}\rangle = \frac{i}{2}$.

Accordingly, we reject claims of a general probabilistic interpretation of the RDM cumulant. The similarities between RDM cumulants and probabilistic cumulants should be understood on the basis that they solve very similar problems of constructing an additively separable quantity from a multiplicatively separable one using very similar techniques.

4 Generalized Normal Ordering

Lastly, we present the GNO formalism[1–5] motivated by considering why RDM cumulants appear here. In brief, RDM elements appear so that the expectation values of normal ordered second-quantized operators (which will themselves be RDM elements) vanish. Cumulants then appear either by asking for the RDM to be broken into size-consistent pieces, or by generalizing the contraction patten of single-reference normal ordering, as we show in Section 4.1. While the use of cumulants over RDMs simplifies the remaining theorems discussed in Sections 4.2 and 4.3, these results have little to do with cumulants in particular but follow from the contraction pattern.

4.1 Wick Expansion

We seek to generalize the familiar single-reference Wick Theorem,[12,13] which says that an arbitrary string of creation and annihilation operators can be expanded into a scalar and a linear combination of operator strings "normal ordered" with respect to Φ, meaning their expectation value for wavefunction Φ is zero. We do so in two steps: we generalize this for vacuum-normal operator strings, and then extend this result to arbitrary operator strings.

First, let us assume a vacuum-normal operator string. In any such expansion, the scalar must be the expectation value of the string because all other terms in the expansion have zero expectation value. If the operator is particle-conserving, this expectation value is an RDM element; otherwise, it is zero.

Now, in the single-reference formalism, we write the scalar term as the sum of all possible "contractions." Contractions take a creation operator and an annihilator operator into a size-consistent tensor element, and multiple contractions are allowed. If we want size-consistent contractions in GNO, we must perform a cumulant expansion of the RDM, per (13), and say that each contraction is a cumulant. Accordingly, the rules for which contractions are allowed are dictated by the possible cumulant patterns in Equation (13). Multiple contractions are still allowed, but contractions now may take n creation and n annihilation operators for any n. As usual, there is a sign factor associated with anticommuting operators to bring operators together for a contraction.

Alternatively, we could have started by generalizing the rule that the scalar term is the sum of all possible complete contractions. If we equate this to the RDM, we find that when trying to normal-order a string of two creation and two annihilation operators, the contractions only treat the product of 1RDMs, and another contraction will be needed for the remainder, the cumulant. In this way, we also arrive at the fact that contractions are cumulants, multiple contractions remain allowed, and contractions must be able to take n creation and n annihilation operators. This was the heart of the approach with convolutions and Hopf algebras by Brouder and coworkers,[59] although they did not recognize the importance of the cumulants.

Writing a creation or annihilation operator as $\hat{q}$, we can write the Wick expansion of a

vacuum normal operator as

$$\hat{q}_p\hat{q}_q\hat{q}_r\hat{q}_q\hat{q}_s\hat{q}_t\hat{q}_u\cdots = \{\hat{q}_p\hat{q}_q\hat{q}_r\hat{q}_q\hat{q}_s\hat{q}_t\hat{q}_u\cdots\} + \sum\{\hat{q}_p\overbrace{\hat{q}_q\hat{q}_r}\hat{q}_s\hat{q}_t\hat{q}_u\cdots\} + \sum\{\hat{q}_p\overbrace{\hat{q}_q\hat{q}_r\hat{q}_s}\hat{q}_t\hat{q}_u\cdots\}$$
$$+\sum\{\hat{q}_p\overbrace{\hat{q}_q\hat{q}_r\hat{q}_s\hat{q}_t}\hat{q}_u\cdots\} + \sum\{\hat{q}_p\overbrace{\hat{q}_q\hat{q}_r\hat{q}_s}\hat{q}_t\hat{q}_u\cdots\} + \sum\{\hat{q}_p\overbrace{\hat{q}_q\hat{q}_r}\hat{q}_s\overbrace{\hat{q}_t\hat{q}_u}\cdots\} + \sum\{\hat{q}_p\overbrace{\hat{q}_q\hat{q}_r\hat{q}_s}\hat{q}_t\hat{q}_u\cdots\} + \ldots \tag{16}$$

where the sums range over all possible contractions, and there can be any number of contractions, and contractions can connect n creation and n annihilation operators for n. At this point in the argument, contractions are defined by $\overbrace{a^p a_q} = \lambda^p_q$, $\overbrace{a^p a^q a_s a_r} = \lambda^{pq}_{rs}$, and so forth. As usual, there is a permutational sign factor to bring non-adjacent operators in the string together. It is also possible to define a "quasi-normal order" where (16) holds, but the contractions are not RDM cumulants. Then it will *not* be true that the normal-ordered operators have zero expectation value with respect to Ψ, as only cumulants have this property. For now, we shall note that (16) alone is needed for all the remaining proofs.

Before proceeding to the general case, let us confirm that our procedure defined on operator strings is well-defined on operators. There are two ways by which different strings can refer to the same operator: the use of anticommutation relations and expanding one orbital as a linear combination of others. The only way to use anticommutation relations on a vacuum-normal order string to get another vacuum-normal order string is to anticommute creation and annihilation operators, so we need to check orbital invariance and antisymmetry. Both of these properties can be shown by a straightforward recursion on the minimum of the number of creation operators and the number of annihilation operators, assuming contractions are antisymmetric and orbital invariant. For RDM cumulants, they are.

Now let us define the Ψ-normal Wick expansion of an arbitrary operator by first bringing it into vacuum-normal order and then bringing the resulting operators into Ψ-normal order using (16). We are composing two maps that obey the anticommutation relations and are orbital invariant, so our final result obeys the anticommutation relations and is orbital invariant.

Our expansion still has the form of (16), but more contractions are possible. First, it is possible to have a contraction if creation operators are not all left of annihilation operators, by reordering them in the transformation to vacuum-normal ordering and then contracting them. This introduces contractions such as $\overbrace{a^p a_s a^q a_r} = -\lambda^{pq}_{rs}$. Second, the contractions of vacuum-normal ordering must also be accounted for. We do this by adding the Kronecker delta from the vacuum normal contraction to the contraction from applying (16) after the vacuum normal ordering step, so we have $\overbrace{a_q a^p} = -\lambda^p_q + \delta^p_q$.

We also note that a Ψ-normal ordered operator is antisymmetric with respect to *any* permutation of the operators in the operator string inside the normal ordering. This property is inherited from the vacuum-normal ordering.

4.2 Ψ-Normal from Vacuum-Normal

It remains to derive the rule for taking products of GNO operators, the analogue of what is usually called the Generalized Wick Theorem. Because we are already in Generalized Normal Ordering, we follow Mukherjee,[4] Evangelista,[5] and their coworkers in instead calling it the extended generalized Wick Theorem. Our derivation in Section 4.3 requires a lemma, the formula for a Ψ-normal ordered operator in terms of vacuum normal ordered operators. We are not aware of any previous presentation of this formula in the literature.

First, we will use the freedom to reorder the operators inside a Ψ-normal operator operator to place all creation operators left of annihilation operators, in vacuum-normal order form. Then

$$\{...a^\dagger_p a^\dagger_q a^\dagger_r a_s a_t a_u...\} = ...a^\dagger_p a^\dagger_q a^\dagger_r a_s a_t a_u... - \sum ...\overbrace{a^\dagger_p a^\dagger_q a^\dagger_r} a_s a_t a_u... - \sum ...a^\dagger_p \overbrace{a^\dagger_q a^\dagger_r a_s a_t a_u}... - \sum ...\overbrace{a^\dagger_p a^\dagger_q a^\dagger_r a_s a_t a_u}... + \sum ...a^\dagger_p \overbrace{a^\dagger_q a^\dagger_r a_s a_t} a_u... + \sum ...\overbrace{a^\dagger_p a^\dagger_q a^\dagger_r a_s a_t} a_u... - \sum ...\overbrace{a^\dagger_p a^\dagger_q a^\dagger_r a_s a_t a_u}... + ... \tag{17}$$

where a term with c contractions has phase $(-1)^c$, and all contraction patterns appear in

the sums. We prove this by induction on the minimum of the number of creation operator and annihilation operators, n. In the base case $n = 0$, no contractions are possible, and (17) reduces to (16).

We proceed to prove the case of $n = k + 1$ if (17) holds for all cases from 0 to k. We can rewrite (16) to reflect that our terms are vacuum-normal ordered and then solve for the completely normal ordered term to give:

$$\{...a_p^\dagger a_q^\dagger a_r^\dagger a_s a_t a_u...\} = ...a_p^\dagger a_q^\dagger a_r^\dagger a_s a_t a_u... - \sum\{...a_p^\dagger \overbrace{a_q^\dagger a_r^\dagger} a_s a_t a_u...\} - \sum\{...a_p^\dagger \overbrace{a_q^\dagger a_r^\dagger a_s a_t a_u}...\}$$
$$- \sum\{...\overbrace{a_p^\dagger a_q^\dagger a_r^\dagger a_s a_t a_u}...\} - \sum\{...a_p^\dagger \overbrace{a_q^\dagger a_r^\dagger a_s a_t} a_u...\} - \sum\{...a_p^\dagger \overbrace{a_q^\dagger a_r^\dagger} a_s a_t a_u...\} - \sum\{...a_p^\dagger \overbrace{a_q^\dagger a_r^\dagger a_s} a_t a_u...\} - ... \tag{18}$$

All the normal ordered terms on the right-hand side are previous cases in the induction, so we substitute in (17) and collect the terms with t contractions. Given a particular set of t contractions, it can be produced by any term in the right-hand side of (18) that has less than t explicit contractions. The remaining contractions will be supplied by substituting (17). Let the number of explicit contractions be denoted o. There are $\binom{t}{o}$ ways to choose which of the t contractions come from the substitution, giving a sign factor of $(-1)^{t-o}$. Thus, the overall coefficient of our set of t contractions is

$$-\sum_{o=1}^{t}(1)^o(-1)^{t-o}\binom{t}{o} = -\big((1-1)^t - (-1)^t\big) = (-1)^t \tag{19}$$

by binomial expansion. All terms with a product of t contractions appear with coefficient $(-1)^t$. This proves (17).

4.3 Extended Generalized Wick's Theorem: Products

We now prove the Extended Generalized Wick Theorem:

$$\{A\}\{B\} = \{AB\} + \sum\{\overbrace{AB}\} \tag{20}$$

where the sum is over all repeated contractions, provided each contraction contains at least one operator from both A and B.

Our strategy is to expand the GNO operators into vacuum normal operators with (17), multiply them, and then convert the result back into GNO operators with (16). This is similar in concept to the proof of Kong, Nooijen, and Mukherjee,[3] but (17) allows a much simpler proof.

Take two Ψ-normal operators, A and B. The expansion via (17) sums over all contractions on only one term, with a sign factor. We call these internal contractions. When we multiply and convert the result back using a Wick expansion, we sum over all possible contractions. This includes contractions of operators from both A and B, called cross-contractions. So the result is a sum over all possible contraction patterns with some coefficient. Let us choose a particular contraction pattern and find its coefficient.

Suppose our contraction pattern has i internal contractions and c cross-contractions. The cross-contractions must occur during the Wick expansion (16), but the internal contractions may originate from (16) or (17). (16) always contributes a sign factor of 1, but the terms with n contractions from (17) contribute a sign factor of $(-1)^n$. Further, there are $\binom{i}{f}$ ways to choose which f internal contractions come from (16). So our total coefficient is

$$\sum_{f=0}^{i}(-1)^f(1)\binom{i}{f} \tag{21}$$

We can change the exponent of 1 arbitrarily to $i-f$ to apply a binomial expansion again and get

$$\sum_{f=0}^{i}(-1)^f(1)^{i-f}\binom{i}{f} = (1-1)^i = \begin{cases} 1 & i=0 \\ 0 & \text{else} \end{cases} \tag{22}$$

In other words, all contraction patterns happen exactly once, which contain no internal contractions. This is precisely the Extended Generalized Wick Theorem, (20).

As first observed by Kong, Nooijen, and Mukherjee,[3] the fact that the contractions are cumulants plays little role in the proof. All that we require is (16), from which (17) follows and then (20). Contractions can be defined in a largely arbitrary manner and still maintain these properties, although care should be taken to ensure that orbital invariance and antisymmetry are preserved. This freedom has been used by Evangelista and coworkers[60,61] to define a variant of GNO where the contractions are the "cumulants" of a density matrix for an statistical ensemble of electronic states, for multistate chemistry.

Suppose the contractions are chosen exactly as in the "canonical" GNO, where the sum of all contractions of n creation and n annihilation operators equal some tensor on those operators. By the logic of Section 2, as long as that family of tensors is multiplicatively separable with respect to some partition, the resulting contractions will be additively separable with respect to it. So even in this more general setting, contractions can quite generally be size-consistent.

5 Conclusions

Despite the importance of reduced density matrix cumulants, we believe that longstanding questions regarding what they are have discouraged their use in new electronic structure theories. This research has answered them. In particular:

1. We have provided a simplified definition of reduced density matrix cumulants and a generating function to provide explicit formulas for them. We have shown that this is mathematically no more complicated than the familiar exponential relation between configuration interaction amplitudes and coupled cluster amplitudes. Furthermore, our definition begins not from obscure mathematics but the intuition that a reduced density matrix cumulant is the "size-consistent part" of a reduced density matrix. The approaches of Mazziotti and coworkers[26] as well as Kutzelnigg, Mukherjee, Hanauer, and Köhn[24,25] are shown to reduce to our solution. Of special importance is the fact

that this our solution is a general prescription to convert between multiplicative and additive separability for fermionic quantities, which can be of use to novel electronic structure methods.

2. Interpretive issues of cumulants have been resolved. The analogy between RDM cumulants and the probabilistic cumulants is based on the fact that they are both combinatorial objects to solve the problem of converting from multiplicative to additive separability. No further probabilistic meaning of the reduced density matrix cumulants is expected, and arguments to the contrary[24] have been refuted. In addition, our definition of cumulants via axioms provides a way to confirm the additive separability of cumulants from their polynomial form and understand why, for some approximate theories, the cumulants are not additive separability. This gives an elementary way to confirm size-consistency from any seemingly *ad hoc* definition.

3. We have also presented the shortest proof in the literature to date of the difficult Generalized Normal Ordering formalism to explain why cumulants appear there and make it more accessible for multireference theories, one of the most pressing problems in electronic structure theory. The key theorems are shown to follow from binomial expansions combined along with the form of allowed contractions in the formalism. In the original Generalized Normal Ordering formalism where normal ordered operators are required to have zero expectation value against some wavefunction, this form of the contractions is just the cumulant expansion of the RDMs. More general formulations are possible and have even been shown to be quite useful,[60,61] and we have shown that the contractions will remain additively separable if the expectation value and RDMs is replaced with some other multiplicatively separable family of tensors.

A Cumulants by Low-Level Combinatorics

To illustrate how the exp and log functions solve the combinatorial problem given in our axiomatic definition of the RDM cumulant (or any other additively separable quantity), we solve it without generating functions by combining the axioms of Percus[34] and Simon[37] with the Möbius inversion of Speed.[62] In brief, suppose a set where some elements are said to be greater than others, or more precisely, a *partially ordered set*. We denote this abstract "greater than" relation with $\supseteq$. Mathematicians prefer to use $\geq$, but the symbol $\supseteq$ suggests the specific relation we will use. Then given an equation of form

$$\sum_{x:y\supseteq x} f(x) = g(y) \quad , \tag{23}$$

Möbius inversion solves for f as a linear combination of the g by

$$\sum_{x:y\supseteq x} g(x)\mu(x,y) = f(y) \tag{24}$$

where the function μ is determined by the recursion relations

$$\sum_{x:z\supseteq x\supseteq y} \mu(y,x) = \delta_{y,z} \tag{25}$$

and

$$\sum_{x:z\supseteq x\supseteq y} \mu(x,z) = \delta_{y,z} \tag{26}$$

Equations (25) and (26) show that the values of μ depend on the set and the rules governing which elements are greater than others.

Readers interested in a detailed mathematical treatment of Möbius inversion are directed to Chapter 16 of reference 63, Chapter 8 of reference 47, Chapter 3 of reference 58, Chapter 3 of reference 64, and reference 65. We especially recommend reference 63.

We shall use the following facts about the set of fermionic partitions:

1. Given any two fermionic partitions ρ, σ, $\rho \supseteq \sigma$ means that each block of σ is contained in a block of ρ. This $\supseteq$ is a *partial order*, which means that we may use Möbius inversion. Figure 1 demonstrates this for a rank-two tensor.

2. Given any fermionic partition of n creation and annihilation operators, arbitrarily pair up creation and annihilation operators, and assign each pair to one of n distinct symbols. Then any fermionic partitions where each operator in the pair is in the same block can be mapped to a partition of n sets. Furthermore, if all pairs are in the same block for σ, all pairs will also be in the same block for any ρ where $\rho \supseteq \sigma$. By this map between fermionic partitions and set partitions, the set of π where $\rho \supseteq \pi \supseteq \sigma$ has exactly the same $\supseteq$ (partial order) structure as some subset of the set of partitions of n objects, which is known as the *partition lattice*. An example of this is shown in Figure 2. By this trick, if we show a statement is true on some subset of the partition lattice, we can show it is true for any "counterpart" of that subset in the fermionic lattice.

3. Suppose $\rho \supseteq \sigma$ and block i of ρ is split into b_i blocks in σ, then

$$\mu(\sigma, \rho) = \prod_i (-1)^{b_i - 1} (b_i - 1)! \quad . \tag{27}$$

   The same property holds on the set of fermionic partitions, because the recursions that determine μ, (26) and (25), depend only on the structure of the partially ordered set, which is the same between the two sets by Point 2.

   This property of the partition lattice is shown in Example 16.17 combined with Theorem 16.4 of Reference 63, proved in two ways in Example 3.10.4 of Reference 57 and Examples 3.3.4 and 3.5.5 of Reference 64, then proved in two more ways in Sections 16 and 18 of Reference 65.

4. Let $\rho \wedge \pi$ denote the partition in the partition lattice whose blocks are obtained by

intersecting the blocks of ρ and π. For any σ and for any π other than $\mathcal{S}$:

$$\sum_{\rho:\rho\wedge\pi=\sigma} \mu(\rho,\mathcal{S}) = 0 \quad . \tag{28}$$

This is proven in the course of Theorem 16.5 of reference 63 and by more sophisticated arguments in Proposition 3.5.4 of reference 64 and Corollary 3.9.3 of reference 57.

By the trick of Point 2, a very similar property holds for the fermionic partitions:

$$\sum_{\rho:\rho,\pi\sqsupseteq\sigma,\rho\wedge\pi=\sigma} \mu(\rho,\mathcal{S}) = 0 \quad . \tag{29}$$

Now, suppose a polynomial satisfying the fermionic axioms. It must have the form of (3). Consider an arbitrary fermionic partition, π.

For most fermionic partitions, π has multiple blocks. Factorize every $\mathcal{C}(\rho)$ in (3) so each tensor contains only indices of a single block of π. Given a partition, σ, the new coefficient of $\mathcal{C}(\sigma)$ after this factorization by π is

$$\mu_{\sigma,\pi} = \sum_{\rho:\rho,\pi\sqsupseteq\sigma,\rho\wedge\pi=\sigma} \mu_\rho \quad . \tag{30}$$

By the third axiom, for any such π, our polynomial is identically zero. Therefore, each coefficient must equal zero.

$$\mu_{\sigma,\pi} = 0 \tag{31}$$

Choosing the coefficients c so that (31) is satisfied is necessary and sufficient to define our cumulant. When $\sigma = \pi$, (30) simplifies to

$$\mu_{\pi,\pi} = \sum_{\rho:\rho\sqsupseteq\pi} \mu_\rho \tag{32}$$

Combining (31) and (32) yields

$$\sum_{\rho \sqsupseteq \pi} \mu_\rho = 0 \tag{33}$$

which is the equation (4) derived in the body of the article.

The above discussion has assumed π consists of multiple blocks, so we may apply the connectedness axiom. If π consists of only one block, $\pi = \mathcal{S}$, and the connectedness axiom does not apply, but $\sum_{\rho \sqsupseteq \pi} \mu_\rho = \mu_{\mathcal{S}} = 1$ by the normalization axiom.

In either case, we require

$$\sum_{\rho \sqsupseteq \pi} \mu_\rho = \delta_{\pi,\mathcal{S}} \quad . \tag{34}$$

But this equation is just (26) when $z = \mathcal{S}$, which is precisely the recursion that determines $\mu(\rho, \mathcal{S})$. Using (27), we may immediately conclude

$$\mu_\sigma = \mu(\rho, \mathcal{S}) = (-1)^{\#\rho - 1}(\#\rho - 1)! \tag{35}$$

where $\#\rho$ is the number of blocks of ρ. This is precisely in agreement with (13)

While (35) is necessary, the connectedness axiom still requires that (31) holds. With a formula for the coefficients just derived, (31) reduces to

$$\sum_{\rho : \rho, \pi \sqsupseteq \sigma, \rho \wedge \pi = \sigma} \mu(\rho, \mathcal{S}) = 0 \quad . \tag{36}$$

This equation is merely (29) and is thus guaranteed to hold. We have therefore shown a polynomial satisfying the fermionic additively separability axioms exists and is unique, and we have determined its coefficients by (35). This polynomial is the probabilistic cumulant.

With Möbius inversion, we can straightforwardly invert our formula to convert multiplicative separability to additive separability and obtain a formula for a multiplicatively separable quantity as a polynomial in additively separable ones. Given fermionic partition π, we may substitute the cumulant formula just found for the cumulants appearing in the

product $\mathcal{T}(\pi)$. We find

$$\mathcal{T}(\pi) = \sum_{\pi \supseteq \rho} \mathcal{C}(\rho)\mu(\rho, \pi) \tag{37}$$

but this is just (24) with $f(y) = \mathcal{T}(y)$ and $g(x) = \mathcal{C}(x)$. Because (24) is equivalent to (23), we have

$$\mathcal{C}(\pi) = \sum_{\pi \supseteq \rho} \mathcal{T}(\rho) \tag{38}$$

which is equivalent to relation (14). We have now derived the relations between probabilistic moments and cumulants entirely from combinatorics and the axiomatic definition.

The reader may wonder what any of this has to do with the exp and log functions of Section 2. The answer is that taking log of an exponential generating function precisely corresponds to performing Möbius inversion of the partition lattice, and taking exp of an exponential generating function undoes the Möbius inversion on the partition lattice, or sums over all partitions. (This is made precise by Theorem 5.1.11 and Example 5.1.13 of Reference 58.) As evidence of this, observe that the Taylor-series expansion coefficients of the log-series are precisely (35) when $\#\rho$ is replaced with the degree of the coefficient. We expect a similar relation holds for the set of fermionic patterns and "generating functions" based on the normal ordered exponential.

The use of generating functions entirely avoids this otherwise tedious and non-obvious problem of Möbius inversion.

Acknowledgement

We acknowledge support from the National Science Foundation, Grant No. CHE-1661604. We acknowledge helpful discussions with Professor Andreas Köhn and Dr. Mathias Hanauer regarding the interpretation of Reference 24, with Professor Francesco Evangelista on the

generalized normal ordering formalism, and with Dr. Chenyang Li regarding typesetting contractions in generalized normal ordering.

References

(1) Kutzelnigg, W.; Mukherjee, D. Normal order and extended Wick theorem for a multi-configuration reference wave function. *J. Chem. Phys.* **1997**, *107*, 432–449.

(2) Mukherjee, D. Normal ordering and a Wick-like reduction theorem for fermions with respect to a multi-determinantal reference state. *Chem. Phys. Lett.* **1997**, *274*, 561–566.

(3) Kong, L.; Nooijen, M.; Mukherjee, D. An algebraic proof of generalized Wick theorem. *J. Chem. Phys.* **2010**, *132*, 234107.

(4) Sinha, D.; Maitra, R.; Mukherjee, D. Generalized antisymmetric ordered products, generalized normal ordered products, ordered and ordinary cumulants and their use in many electron correlation problem. *Comput. Theor. Chem.* **2013**, *1003*, 62–70.

(5) Li, C.; Evangelista, F. A. Multireference Driven Similarity Renormalization Group: A Second-Order Perturbative Analysis. *J. Chem. Theory Comput.* **2015**, *11*, 2097–2108.

(6) Mazziotti, D. A. Contracted Schrödinger equation: Determining quantum energies and two-particle density matrices without wave functions. *Phys. Rev. A* **1998**, *57*, 4219–4234.

(7) Mazziotti, D. A. Two-Electron Reduced Density Matrix as the Basic Variable in Many-Electron Quantum Chemistry and Physics. *Chem. Rev.* **2011**, *112*, 244–262.

(8) Piris, M.; Ugalde, J. M. Perspective on natural orbital functional theory. *Int. J. Quant. Chem.* **2014**, *114*, 1169–1175.

(9) Peng, R.; Copan, A. V.; Sokolov, A. Y. Simulating X-ray Absorption Spectra with Linear-Response Density Cumulant Theory. *J. Phys. Chem. A* **2019**, *123*, 1840–1850.

(10) Hollett, J. W.; Hosseini, H.; Menzies, C. A cumulant functional for static and dynamic correlation. *J. Chem. Phys.* **2016**, *145*, 084106.

(11) Hollett, J. W.; Loos, P.-F. Capturing static and dynamic correlation with ΔNO-MP2 and ΔNO-CCSD. *J. Chem. Phys.* **2020**, *152*, 014101.

(12) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics*; Cambridge University Press, 2009.

(13) Paldus, J. Lecture Notes. University of Nijmegen, Holland. URL: https://www.math.uwaterloo.ca/ paldus/downloads/NijmegenLectures.pdf. Accessed, February 2020.

(14) Neuscamman, E.; Yanai, T.; Chan, G. K.-L. A review of canonical transformation theory. *Int. Rev. Phys. Chem.* **2010**, *29*, 231–271.

(15) Evangelista, F. A. Perspective: Multireference coupled cluster theories of dynamical electron correlation. *J. Chem. Phys.* **2018**, *149*, 030901.

(16) Li, C.; Evangelista, F. A. Multireference Theories of Electron Correlation Based on the Driven Similarity Renormalization Group. *Annu. Rev. Phys. Chem.* **2019**, *70*, 245–273.

(17) Harris, F. E. Cumulant-based approximations to reduced density matrices. *Int. J. Quant. Chem.* **2002**, *90*, 105–113.

(18) Nooijen, M.; Wladyslawski, M.; Hazra, A. Cumulant approach to the direct calculation of reduced density matrices: A critical analysis. *J. Chem. Phys.* **2003**, *118*, 4832–4848.

(19) Kong, L.; Valeev, E. F. A novel interpretation of reduced density matrix and cumulant for electronic structure theories. *J. Chem. Phys.* **2011**, *134*, 214109.

(20) Herbert, J. M.; Harriman, J. E. *Reduced-Density-Matrix Mechanics: With Application to Many-Electron Atoms and Molecules*; John Wiley & Sons, Inc., 2007; pp 261–292.

(21) Yasuda, K.; Nakatsuji, H. Direct determination of the quantum-mechanical density matrix using the density equation. II. *Phys. Rev. A* **1997**, *56*, 2648–2657.

(22) Schirmer, J. *Many-Body Methods for Atoms, Molecules and Clusters*; Springer International Publishing, 2018.

(23) Kubo, R. Generalized Cumulant Expansion Method. *J. Phys. Soc. Japan* **1962**, *17*, 1100–1120.

(24) Hanauer, M.; Köhn, A. Meaning and magnitude of the reduced density matrix cumulants. *Chem. Phys.* **2012**, *401*, 50–61.

(25) Kutzelnigg, W.; Mukherjee, D. Cumulant expansion of the reduced density matrices. *J. Chem. Phys.* **1999**, *110*, 2800–2809.

(26) Mazziotti, D. A. Approximate solution for electron correlation through the use of Schwinger probes. *Chem. Phys. Lett.* **1998**, *289*, 419–427.

(27) Ziesche, P. *Mathematical and Computational Chemistry*; Springer US, 2000; pp 33–56.

(28) Kutzelnigg, W. n-Electron problem and its formulation in terms ofk-particle density cumulants. *Int. J. Quant. Chem.* **2003**, *95*, 404–423.

(29) Mazziotti, D. A. 3,5-contracted Schrödinger equation: Determining quantum energies and reduced density matrices without wave functions. *Int. J. Quant. Chem.* **1998**, *70*, 557–570.

(30) Mazziotti, D. A. Comparison of contracted Schrödinger and coupled-cluster theories. *Phys. Rev. A* **1999**, *60*, 4396–4408.

(31) Mazziotti, D. A. *Mathematical and Computational Chemistry*; Springer US, 2000; pp 139–163.

(32) Mazziotti, D. A. *Reduced-Density-Matrix Mechanics: With Application to Many-Electron Atoms and Molecules*; John Wiley & Sons, Inc., 2007; pp 165–203.

(33) Kutzelnigg, W.; Mukherjee, D. Irreducible Brillouin conditions and contracted Schrödinger equations for n-electron systems. IV. Perturbative analysis. *J. Chem. Phys.* **2004**, *120*, 7350–7368.

(34) Percus, J. K. Correlation inequalities for Ising spin lattices. *Commun. Math. Phys.* **1975**, *40*, 283–308.

(35) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd, 2000.

(36) McWeeny, R. Some Recent Advances in Density Matrix Theory. *Rev. Mod. Phys.* **1960**, *32*, 335–369.

(37) Simon, B. *The Statistical Mechanics of Lattice Gases, Volume I (Princeton Legacy Library)*; Princeton University Press, 2016.

(38) Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **1966**, *45*, 4256–4266.

(39) Lehtola, S.; Tubman, N. M.; Whaley, K. B.; Head-Gordon, M. Cluster decomposition of full configuration interaction wave functions: A tool for chemical interpretation of systems with strong correlation. *J. Chem. Phys.* **2017**, *147*, 154105.

(40) Lindgren, I. A coupled-cluster approach to the many-body perturbation theory for open-shell systems. *Int. J. Quant. Chem.* **2009**, *14*, 33–58.

(41) Solomon, A. I. Normal Ordering and Abnormal Nonsense. *Found. Phys.* **2009**, *40*, 684–691.

(42) Jeziorski, B.; Paldus, J. Valence universal exponential ansatz and the cluster structure of multireference configuration interaction wave function. *J. Chem. Phys.* **1989**, *90*, 2714–2731.

(43) Kutzelnigg, W.; Mukherjee, D. *Reduced-Density-Matrix Mechanics: With Application to Many-Electron Atoms and Molecules*; John Wiley & Sons, Inc., 2007; pp 293–330.

(44) Kutzelnigg, W. *Explicitly Correlated Wave Functions in Chemistry and Physics*; Springer Netherlands, 2003; pp 3–90.

(45) Stillwell, J. *Naive Lie Theory*; Springer New York, 2008.

(46) Wilf, H. S. *Generatingfunctionology*; Elsevier, 1990.

(47) Cameron, P. J. *Notes on Counting: An Introduction to Enumerative Combinatorics*; Cambridge University Press, 2017.

(48) Korona, T. Two-particle density matrix cumulant of coupled cluster theory. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5698.

(49) Bozkaya, U.; Turney, J. M.; Yamaguchi, Y.; Schaefer, H. F.; Sherrill, C. D. Quadratically convergent algorithm for orbital optimization in the orbital-optimized coupled-cluster doubles method and in orbital-optimized second-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2011**, *135*, 104103.

(50) Bozkaya, U.; Sherrill, C. D. Orbital-optimized coupled-electron pair theory and its analytic gradients: Accurate equilibrium geometries, harmonic vibrational frequencies, and hydrogen transfer reactions. *J. Chem. Phys.* **2013**, *139*, 054104.

(51) Bozkaya, U. Analytic energy gradients for the orbital-optimized third-order Møller–Plesset perturbation theory. *J. Chem. Phys.* **2013**, *139*, 104116.

(52) Bozkaya, U.; Sherrill, C. D. Orbital-optimized MP2.5 and its analytic gradients: Approaching CCSD(T) quality for noncovalent interactions. *J. Chem. Phys.* **2014**, *141*, 204105.

(53) Mazziotti, D. A. Parametrization of the two-electron reduced density matrix for its direct calculation without the many-electron wave function: Generalizations and applications. *Phys. Rev. A* **2010**, *81*, 062515.

(54) DePrince, A. E.; Mazziotti, D. A. Connection of an elementary class of parametric two-electron reduced-density-matrix methods to the coupled electron-pair approximations. *Mol. Phys.* **2012**, *110*, 1917–1925.

(55) Lyakh, D. I.; Bartlett, R. J. Algebraic connectivity analysis in molecular electronic structure theory II: total exponential formulation of second-quantised correlated methods. *Mol. Phys.* **2013**, *112*, 213–260.

(56) Flajolet, P.; Sedgewick, R. *Analytic Combinatorics*; Cambridge University Press, 2009.

(57) Stanley, R. P. *Enumerative Combinatorics*; Cambridge University Press, 2009.

(58) Stanley, R. P.; Fomin, S. *Enumerative Combinatorics*; Cambridge University Press, 1999.

(59) Brouder, C.; Fauser, B.; Frabetti, A.; Oeckl, R. Quantum field theory and Hopf algebra cohomology. *J. Phys. A* **2004**, *37*, 5895–5927.

(60) Li, C.; Evangelista, F. A. Driven similarity renormalization group for excited states: A state-averaged perturbation theory. *J. Chem. Phys.* **2018**, *148*, 124106.

(61) Li, C.; Lindh, R.; Evangelista, F. A. Dynamically weighted multireference perturbation theory: Combining the advantages of multi-state and state-averaged methods. *J. Chem. Phys.* **2019**, *150*, 144107.

(62) Speed, T. P. Cumulants and Partition Lattices. **1983**, *25*, 378–388.

(63) Bóna, M. *A Walk Through Combinatorics*; WORLD SCIENTIFIC, 2016.

(64) Martin, J. L. Lecture Notes. University of Kansas, United States. URL: http://jlmartin.faculty.ku.edu/CombinatoricsNotes.pdf. Accessed, March 2020.

(65) Godsil, C. Lecture Notes. University of Waterloo, Canada. URL: https://arxiv.org/abs/1803.06664v1. Accessed, March 2020.

Graphical TOC Entry