Combinatorics of chemical reaction systems

Nicolas Behr[a], Gérard H. E. Duchamp[b] and Karol A. Penson[c]

We propose a concise stochastic mechanics framework for chemical reaction systems that allows to formulate evolution equations for three general types of data: the probability generating functions, the exponential moment generating functions and the factorial moment generating functions. This formulation constitutes an intimate synergy between techniques of statistical physics and of combinatorics. We demonstrate how to analytically solve the evolution equations for all six elementary types of single-species chemical reactions by either combinatorial normal-ordering techniques, or, for the binary reactions, by means of Sobolev-Jacobi orthogonal polynomials. The former set of results in particular highlights the relationship between infinitesimal generators of stochastic evolution and parametric transformations fo probability distributions. Moreover, we present exact results for generic single-species non-binary reactions, hinting at a notion of compositionality of the analytic techniques.

1 Introduction

Intended as an invitation to interdisciplinary researchers and in particular to combinatorists, we present in this work an extension of the early work of Dellbrück [1] on probability generating functions for chemical reaction systems to a so-called stochastic mechanics framework. While the idea to study chemical reaction systems in terms of probability generating functions is thus not new and on the contrary one of the standard techniques of this field (see e.g. [2] for a historical overview), we believe that the reformulation of these techniques in terms of the stochastic mechanics formalism could lead to fruitful interaction of a broader audience of theoretical researchers. In the spirit of the ideas presented by M. Doi in his seminal paper [3], the main motivation for such a reformulation lies in a

clear conceptual separation of (i) the state space of the system and (ii) the linear operators implementing the evolution of the system. Combined with insights obtained in a recent study of stochastic graph rewriting systems [4–6], one may add to this list (iii) the linear operators that implement observable quantities such as moments of number counts on states. It is only through combining this Ansatz with the standard notions of combinatorial generating functions that we find the true strengths of the stochastic mechanics approach: providing an avenue to obtain exact solutions to dynamical evolution equations. Combinatorists will recognize in our formulation of evolution equations intrinsic notions of normal-ordering problems, and indeed certain semi-linear normal-ordering techniques [7–10] will prove immensely fruitful in this direction. Chemists and other practitioners might appreciate that our solutions not only provide asymptotic information on the time-evolution of the reaction systems, but on the contrary full information on the evolution of reaction systems from any initial state at time $t = 0$ to any desired time $t = T$ (with $T > 0$). While many individual results on such time-evolutions are known in the literature [2, 11], we hope that our concise formalism may help to consolidate the knowledge on the mathematical

[a] *Institut de Recherche en Informatique Fondamentale (IRIF), Université Paris-Diderot (Paris 07), France; E-Mail: nicolas.behr@irif.fr*
[b] *Laboratoire d'Informatique de Paris-Nord (LIPN), Institut Galilée, CNRS UMR 7030, Université Paris 13, Sorbonne Paris Cité, Villetaneuse, France; E-Mail: ghed@lipn.univ-paris13.fr*
[c] *Laboratoire de Physique Theorique de la Matière Condensée (LPTMC), CNRS UMR 7600, Sorbonne Universités, Université Pierre et Marie Curie (Paris 06), France; E-Mail: penson@lptl.jussieu.fr*

methods involved in deriving such exact results, and in some cases even extend the current knowledge. Amongst the novel technical results in this paper is an exact solution to single-species binary chemical reaction systems, which is obtained via a necessary modification of an Ansatz by McQuarrie [12] via so-called Sobolev-Jacobi orthogonal polynomials, which have been introduced in the work of Kwon and Littlejohn [13–16].

For clarity of exposition, we will first present the general theory of so-called continuous-time Markov chains (CTMCs) [17] in our framework of stochastic mechanics. This approach to the theory of CTMCs has the advantage of giving clear intuitions as familiar from the general theory of statistical physics to the types of computations that arise, and moreover admits an entry point to non-specialists that we think is somewhat more accessible than the vast standard literature on the topic. We will then present a derivation of a framework to describe chemical reaction systems, and illustrate how to obtain exact closed-form solutions to the evolution equations of certain types of reaction systems.

2 The stochastic mechanics formalism

As a preparation for the main part of this paper, we begin with a short survey of some of the standard notions of *continuous time Markov chain (CTMC)* theory [17] and related notions of probability theory. In particular, the notations we will use are inspired by those common in quantum mechanics, which suggests to call the present framework by analogy *stochastic mechanics* [18].

We will be concerned with the study of dynamical stochastic systems whose space of pure states (i.e. concrete configurations) is described in the following manner:

Definition 1. *For a stochastic dynamical system, let S denote the vector space (over $\mathbb{R}$) spanned by the*

pure states *(basis vectors)* $|s\rangle \in S$. *Typically, S is countably infinite dimensional. A mixed state* $|\psi\rangle \in Prob(S)$ *of the system is a* probability distribution *over S, with Prob(S) defined as*

$$Prob(S) := \left\{ |\Psi\rangle = \sum_s \psi_s |s\rangle , \right.$$
$$\left. \psi_s \in \mathbb{R}_{\geq 0} \text{ and } \sum_s \psi_s = 1 \right\}. \quad (1)$$

Elements of $Prob(S)$ may thus in particular be seen as special elements of the space $\ell^1_{\mathbb{R}}(S)$ of $\ell^1_{\mathbb{R}}$-summable sequences over S,

$$\ell^1(S) := \left\{ |\Psi\rangle = \sum_s \psi_s |s\rangle , \right.$$
$$\left. \psi_s \in \mathbb{R} \text{ and } \sum_s |\psi_s| < \infty \right\}. \quad (2)$$

For later convenience, we introduce an operation

$$\langle \mathscr{R}| : \ell^1_{\mathbb{R}}(S) \to \mathbb{R}$$

that implements the *summation of coefficients* of a linear combination or probability distribution over S, and which is defined via its action on basis vectors as

$$\langle \mathscr{R}|s \rangle := 1 \quad \text{for all } |s\rangle \in S. \quad (3)$$

For example, applying $\langle \mathscr{R}|$ to a probability distribution $|\Psi\rangle \in Prob(S)$ results in

$$\langle \mathscr{R}|\Psi\rangle = \sum_s \psi_s \langle \mathscr{R}|s\rangle = \sum_s \psi_s = 1 .$$

The reader is cautioned *not* to take the analogy to the notations of quantum mechanics too far in interpreting the notation $\langle \mathscr{R}|s\rangle$ as a form of scalar product. In fact, since this is a rather crucial point, we provide in Appendix A a direct comparison of the notions of stochastic mechanics as opposed to the notions of quantum mechanics. For the purposes of stochastic mechanics, we thus merely

treat $\langle \mathscr{R} |$ as a convenient (and standard*) shorthand notation.

We now turn to one of the core concepts of the stochastic mechanics framework:

Definition 2 (cf. e.g. [17]). *Let S denote a vector space of pure states, and let Prob(S) denote the space of probability distributions over S. Then a* continuous time Markov chain (CTMC) *over S is defined in terms of providing an* input state

$$|\Psi(0)\rangle \in Prob(S) \tag{4}$$

and an evolution operator *aka* evolution semi-group

$$\mathscr{E} : [0, \infty) \times Prob(S) \to Prob(S), \tag{5}$$

with the following defining properties (for all non-negative real parameters $s, t \in \mathbb{R}_{\geq 0}$):

$$\mathscr{E}(0) = \mathbb{1} \tag{6a}$$
$$\mathscr{E}(s+t) = \mathscr{E}(s)\mathscr{E}(t) = \mathscr{E}(t)\mathscr{E}(s) \tag{6b}$$
$$\frac{d}{dt}\mathscr{E}(t) = H\mathscr{E}(t) = \mathscr{E}(t)H. \tag{6c}$$

Here, the first two equations express the semi-group property of $\mathscr{E}$, while (6c) *expresses the Markov property (aka memorylessness) itself, with H the* infinitesimal generator *of $\mathscr{E}$. Then the time-dependent state $|\Psi(t)\rangle \in Prob(S)$ of the CTMC is given by*

$$|\Psi(t)\rangle = \mathscr{E}(t)|\Psi(0)\rangle. \tag{7}$$

Owing to (6c), *$|\Psi(t)\rangle$ satisfies the so-called* Master equation *aka* Kolmogorov backwards equation,

$$\frac{d}{dt}|\Psi(t)\rangle = H|\Psi(t)\rangle. \tag{8}$$

---

* Some alternative notations frequently encountered in the literature [3, 19, 20] include

$$\langle \mathscr{R} | \equiv \langle | \equiv \langle P | \equiv \langle - |.$$

Since the infinitesimal generator H thus bears a certain formal resemblance in its role to the Hamiltonian in the quantum mechanics setting, we will (by an abuse of terminology as illustrated in Appendix A) refer to it as Hamiltonian also in the stochastic mechanics setting.

The definition of a continuous-time Markov chain thus comprises the information on what are the *pure states* of the stochastic dynamical system, and on what are the possible *transitions* of the system. The transitions in turn would typically implement the notions of, say, chemical reactions or other changes of configurations of states. In practice, the precise nature of these transitions will determine the structure of the infinitesimal generator H. We will provide an in-detail discussion of the infinitesimal generators in the case of chemical reaction systems in the main part of this paper.

It might be worth pointing out that the systems thus defined are in a certain sense *dimensionless*, i.e. there is no notion of physical space understood from the definitions. However, as for example realized in the famous Doi-Peliti formalism [3, 20], it is possible to extend the concept of pure states to include a notion of space e.g. in the form of configurations over a lattice of positions in physical space, and even to consider taking a certain limit of zero lattice spacing to recover a notion of continuous space. For the purposes of this paper, we will focus solely on the dimensionless abstraction.

As we will see in the sequel, the Hamiltonian H of a CTMC can in general turn out to be an *unbounded* operator acting on a countably infinite-dimensional space, whence considerable care is necessary in order to understand the precise mathematical meaning of the CTMC definitions and of the evolution semigroup. While in the case of a finite state-space we simply have the relation (see e.g.

[21] and references therein)

$$\mathscr{E}(t) = e^{tH},$$

for unbounded operators H such an expression would possibly not be mathematically well-posed. In the general cases, one therefore needs to resort to more intricate concepts such as the *Hille-Yosida theory* of functional analysis. We refer the interested readers to the standard literature on the subject (see e.g. [22–24]) for the full mathematical details, and content ourselves here with stating a number of abstract properties of the infinitesimal generators H that will prove useful to our framework in the following.

Lemma 1. *Given a CTMC over some state space S with initial state $|\Psi(0)\rangle \in Prob(S)$, evolution semi-group $\mathscr{E}$ and infinitesimal generator H, we have that*

$$\langle \mathscr{R} | H = 0. \tag{9}$$

Proof. Recall that since for all $t \geq 0$ the state $|\Psi(t)\rangle$ of the CTMC is a probability distribution, it satisfies

$$\langle \mathscr{R} | \Psi(t) \rangle = \sum_s \psi_s = 1.$$

This entails in particular that

$$0 = \frac{d}{dt} \langle \mathscr{R} | \Psi(t) \rangle = \frac{d}{dt} \langle \mathscr{R} | \mathscr{E}(t) | \Psi(0) \rangle$$
$$\overset{(6c)}{=} \langle \mathscr{R} | H \mathscr{E}(t) | \Psi(0) \rangle.$$

But since this must hold for arbitrary times $t \geq 0$, the claim (9) follows. $\square$

2.1 Advantages of the stochastic mechanics formulation

We will now present some of the key strengths of the formulation of CTMCs in terms of the stochastic mechanics formalism, a point which we will elaborate further when entering the discussion of chemical reaction systems.

One of the main advantages of the formalism is the ability to compute naturally (and in many cases even in closed form) quantities of interest in practical applications, namely the *moments* of observables to be computed in the given problem. Since part of the computations for chemical reactions will involve additional mathematical structures, we choose to first present some of the general concepts here.

Suppose then that we are studying a continuous-time Markov chain (CTMC) with space of pure states S, initial state $|\Psi(0)\rangle \in Prob(S)$, evolution semi-group $\mathscr{E}(t)$ and infinitesimal generator H as before.

Definition 3. *An* observable *O for a CTMC is a diagonal linear operator on S, whence for all pure states $|s\rangle \in S$*

$$O|s\rangle = \omega_s(O)|s\rangle \quad (\omega_s(O) \in \mathbb{R}). \tag{10}$$

We denote the space of observables on S by $\mathscr{O}(S)$.

Since the observables are thus by definition *diagonal* in the basis of pure states, any two observables $O_1, O_2 \in \mathscr{O}(S)$ *commute*:

$$[O_1, O_2] = O_1 O_2 - O_2 O_1 = 0. \tag{11}$$

Consider now a fixed (finite) set of observables $\{O_1, \ldots, O_n\}$ $(O_i \in \mathscr{O}(S))$. Let us introduce the convenient *multi-index notation*

$$\underline{x} := (x_1, \ldots, x_n)$$

for vectors and

$$\underline{\lambda} \cdot \underline{O} := \sum_{i=1}^n \lambda_i O_i \tag{12}$$

for linear combinations of observables. Here, we will either consider real coefficients $\lambda_i \in \mathbb{R}$ or *formal* coefficients (such as e.g. in Definition 5). We will also make use of the shorthand notation

$$\underline{O}^{\underline{m}} := \prod_{i=1}^n O_i^{m_i}, \tag{13}$$

with exponents $m_i \in \mathbb{Z}_{\geq 0} < \infty$ (i.e. finite non-negative integers).

Denoting by $|\Psi(t)\rangle = \mathscr{E}(t)|\Psi(0)\rangle$ the *time-dependent state of the system*, we may now define in complete analogy to standard probability theory the concept of *expectation values* aka *moments* of observables:

Definition 4. *For a given choice of exponents $m_1, \ldots, m_n \in \mathbb{Z}_{\geq 0} \leq \infty$, we define*

$$\langle \underline{O}^{\underline{m}} \rangle (t) := \langle \mathscr{R} | O_1^{m_1} \cdots \ldots \cdots O_n^{m_n} | \Psi(t) \rangle . \quad (14)$$

For example, the expectation value of a single observable O_i in state $|\Psi(t)\rangle$, in the traditional probability theory literature sometimes denoted $\mathbb{E}_{|\Psi(t)\rangle}(O_i)$, would thus read

$$\begin{aligned}
\mathbb{E}_{|\Psi(t)\rangle}(O_i) \equiv \langle O_i \rangle (t) &= \langle \mathscr{R} | \mathscr{O}_i | \Psi(t) \rangle \\
&= \sum_s \psi_s(t) \langle \mathscr{R} | O_i | s \rangle = \\
&\overset{(10)}{=} \sum_s \psi_s(t) \omega_s(O_i) \langle \mathscr{R} | s \rangle \\
&= \sum_s \psi_s(t) \omega_s(O_i) .
\end{aligned}$$

Definition 5. *With notations as above, the* exponential generating function (EGF) $\mathscr{M}(t;\underline{\lambda})$ *for the set of observables $\{O_1, \ldots, O_m\}$ of the system with time-dependent state $|\Psi(t)\rangle$ is defined as*

$$\begin{aligned}
\mathscr{M}(t;\underline{\lambda}) &:= \left\langle e^{\underline{\lambda} \cdot \underline{O}} \right\rangle (t) \\
&= \sum_{m \geq 0} \frac{1}{m!} \langle \mathscr{R} | (\underline{\lambda} \cdot \underline{O})^m | \Psi(t) \rangle .
\end{aligned} \quad (15)$$

Arbitrary moments may be computed from $\mathscr{M}(t;\underline{\lambda})$ according to

$$\begin{aligned}
&\left\langle O_1^{m_1} \cdots \ldots \cdots O_n^{m_n} \right\rangle (t) \\
&\left[\frac{\partial^{m_1}}{\partial \lambda_1^{m_1}} \cdots \frac{\partial^{m_n}}{\partial \lambda_1^{m_n}} \mathscr{M}(t;\underline{\lambda}) \right] \Bigg|_{\underline{\lambda} \to \underline{0}} .
\end{aligned} \quad (16)$$

In other words, knowing the exponential moment-generating function $\mathscr{M}(t;\underline{\lambda})$ for the observables of interest in a given system amounts typically to knowing all relevant stochastic dynamical properties of the system.

Finally, we are now in a position to assemble all concepts introduced thus far into one of the cornerstones of this paper (see [4] for the original version of the result, formulated in the framework of rule algebras invented by the first named author, and [6] for its most general variant):

Theorem 1. *For a CTMC with space of pure states S, initial state $|\Psi(0)\rangle \in Prob(S)$, evolution semigroup $\mathscr{E}(t)$ and infinitesimal generator H, the exponential moment-generating function $\mathscr{M}(t;\underline{\lambda})$ for a given (finite) set of observables $\{O_1, \ldots, O_n\}$ ($O_i \in \mathscr{O}(S)$) fulfills the evolution equation*

$$\frac{d}{dt} \mathscr{M}(t;\underline{\lambda}) = \sum_{q \geq 1} \frac{1}{q!} \langle \mathscr{R} | \left(ad_{(\underline{\lambda} \cdot \underline{O})}^{\circ q} H \right) e^{\underline{\lambda} \cdot \underline{O}} | \Psi(t) \rangle . \quad (17)$$

Here, for two linear operators A and B the notation $ad_A B$ stands for the adjoint action *aka* commutator,

$$ad_A B := [A, B] = AB - BA ,$$

while $ad_A^{\circ q} B$ denotes the q-fold commutator ($q > 1$),

$$\begin{aligned}
ad_A^{\circ q} B &:= \underbrace{(ad_A \circ \ldots \circ ad_A)}_{q \text{ times}}(B) \\
&= \underbrace{[A, [A, \ldots, [A, B] \ldots]]}_{q \text{ times}} .
\end{aligned}$$

Proof. See Appendix C.1. $\qquad\square$

3 Chemical reaction systems of one species

According to the standard literature on chemical reaction systems [1, 12, 25–28] (see in particular [2] for a historical review of the early work on the subject), a chemical reaction system of one species of

5

particles is a stochastic transition system whose basic ("one-step") transitions are of the form

$$iA \xrightarrow{r_{i,o}} oA. \qquad (18)$$

The notation encodes a transition in which i copies of a particle of species A are transformed into o copies of the same species, while* $r_{i,o} \in \mathbb{R}_{\geq 0}$ denotes the transition's *base rate*. As sketched in Figure 1, one should think of the physical system to consist of a finite number of indistinguishable particles of type A, with the transitions implementing operations of removing and adding particles to the system. In this standard description, the positions and impulses of the individual particles are completely neglected, and all physical parameters such as temperature, pressure, volume, chemical potentials etc. are collected into the values of the base rates.
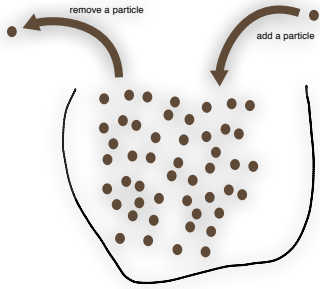
Fig. 1 Sketch of the action of the *creation operator* $a^\dagger$ (adding a particle) and of the *annihilation operator* a (removing a particle, in all possible ways) on an urn with a precise number n of particles (aka a *pure state* $|\underline{n}\rangle$). See equation (26) below for the concrete definitions.

The precise details of the dynamics of the system include of course the way in which transitions occur. Following the general principles of stochastic

---

mechanics introduced in the previous section, we would like to know how a transition influences the time-dependent state $|\Psi(t)\rangle$ of the system. We identify in the single-species case as *pure states* configurations with a precise (and finite) number of particles n, which we denote by $|n\rangle$. Then according to the general formalism, $|\Psi(t)\rangle$ is a probability distribution over these pure states with time-dependent coefficients,

$$|\Psi(t)\rangle = \sum_{n \geq 0} \psi_n(t) |n\rangle. \qquad (19)$$

We quote from the literature (see e.g. [2]) the following result: for a reaction such as in (18) that consumes i particles and produces o particles at rate $r_{i,o}$, the change to the coefficient $\psi_n(t)$ in $|\Psi(t)\rangle$ reads

$$\frac{d}{dt}\psi_n(t) = r_{i,o}\left((n+i-o)_i \psi_{n+i-o}(t) - (n)_i \psi_n(t) \right), \qquad (20)$$

where we use the notational convention

$$(n)_i := i!\binom{n}{i} \equiv \Theta(n-i)\frac{n!}{(n-i)!}, \qquad (21)$$

whence $(n)_i = 0$ for $n < i$.

Before proceeding with the analysis, it is interesting to note that the coefficient $(n)_i$ has in fact a natural *combinatorial* interpretation: since we consider all particles as indistinguishable, $(n)_i$ may be interpreted as the number of ways to pick i particles out of a pool of n particles without replacement.

As standard in the literature, (20) may be expressed in a more tractable form via the introduction of the concept of *probability generating functions*: given the formula for the time-dependent state $|\Psi(t)\rangle$ of the system as in (19), one may construct the *formal power series* $P(t;x)$ (with formal variable x):

$$P(t;x) := \sum_{n \geq 0} \psi_n(t) x^n. \qquad (22)$$

6

We further introduce the linear operators ∂_x and $\hat{x}$, defined to act on arbitrary formal power series $f(x)$ according to

$$\partial_x f(x) := \tfrac{\partial}{\partial x} f(x), \quad \hat{x} f(x) := x f(x). \qquad (23)$$

Since in particular

$$(\partial_x)^i x^n = (n)_i x^{n-i},$$

this allows us to recast the information encoded in (20) in the form of an *evolution equation* for $P(t;x)$:

$$\frac{\partial}{\partial t} P(t;x) = r_{i,o} (\hat{x}^o - \hat{x}^i)(\partial_x)^i P(t;x). \qquad (24)$$

The readers familiar with the theory of the Heisenberg-Weyl (HW) algebra (cite!) will recognize the operators $\hat{x}$ and ∂_x as the generators of the HW algebra in the *Bargmann-Fock representation*, whose defining relation is computed (for $f(x)$ an arbitrary formal power series) from

$$[\partial_x, \hat{x}] f(x) = f(x). \qquad (25)$$

Contact with the standard notations of stochastic mechanics is made by employing instead the isomorphic so-called *canonical representation* of the HW algebra, via the one-to-one correspondence

$$\begin{aligned}
x^n &\leftrightarrow |n\rangle \\
\hat{x} &\leftrightarrow a^\dagger \\
\partial_x &\leftrightarrow a,
\end{aligned} \qquad (26)$$

with $|n\rangle$ the pure state of *n* particles. The operator $a^\dagger$ is called the *creation operator* and a the *annihilation operator*. One obtains via the one-to-one correspondence the following concrete properties of this representation*

$$a^\dagger |n\rangle = |n+1\rangle, \quad a|n\rangle = (n)_1 |n-1\rangle. \qquad (27)$$

* Note that this particular normalization is sometimes referred to in the literature as the "mathematics" variant [29, 30] of bosons, as opposed to the "physics variant", in which the normalization is chosen more symmetrically.

We are finally in a position to formulate the Hamiltonian *H* of a chemical reaction system in the form proposed by M. Doi [3]:

Definition 6. *Consider a chemical reaction system of one species, specified in terms of an* initial state,

$$|\Psi(0)\rangle = \sum_{n \geq 0} \psi_n(t) |n\rangle,$$

and a set of one-species reactions,

$$iA \xrightarrow{r_{i,o}} oA.$$

Then this data induces a notion of stochastic transition system, namely a continuous-time Markov chain on the space of probability distributions over the number vectors and with infinitesimal generator

$$H = \sum_{i,o} r_{i,o} \left(\left(a^\dagger\right)^o - \left(a^\dagger\right)^i \right) a^i. \qquad (28)$$

Referring to the standard chemistry literature for the precise details, in practice only reactions with at most two input or output particles and a total number of at most three particles have to be considered, whence

$$0 \leq i, o \leq 2, \ 1 \leq i + o \leq 3.$$

The corresponding formulae are presented in the first column of Table 1.

Following the stochastic mechanics formalism, it then remains to construct a *dual reference state* $\langle \mathscr{R} |$ and to identify *observables*.

Definition 7. *The* dual number vectors† $\langle m |$ *(m $\in$ $\mathbb{Z}_{\geq 0}$) are defined via their actions on basis vectors,*

$$\langle m | n \rangle := \delta_{m,n} n!. \qquad (29)$$

† Note that "dual" is to be understood as "formal dual", i.e. via action on basis vectors; we have to resort to this technicality since $\ell_1^{\mathbb{R}}(\mathscr{S})$ does not admit a proper topological dual.

One may then define the resolution of the identity,

$$\mathbb{1} = \sum_{n=0}^{\infty} \frac{1}{n!} |n\rangle \langle n| , \qquad (30)$$

as well as the concrete realization of the dual reference vector $\langle \mathscr{R}|$ on the state space spanned by the number vectors,

$$\langle \mathscr{R}| \equiv \sum_{n=0}^{\infty} \frac{1}{n!} \langle n| = \langle 0| e^a . \qquad (31)$$

We refer the reader to Appendix A.1 for some comparison of these conventions to the analogous concepts in quantum mechanics (and in particular for a comparison of the definition of $\langle \mathscr{R}|$ to that of coherent states).

The dual reference vector $\langle \mathscr{R}|$ is thus in particular a *left Eigenvector with Eigenvalue* 1 *of the creation operator $a^{\dagger}$*,

$$\langle \mathscr{R}| a^{\dagger} = \langle \mathscr{R}| , \qquad (32)$$

which may be seen as follows:

$$\langle \mathscr{R}| a^{\dagger} = \langle 0| e^a a^{\dagger} = \sum_{m=0}^{\infty} \frac{1}{m!} \langle 0| a^m a^{\dagger}$$
$$= \sum_{m=1}^{\infty} \frac{(m)_1}{m!} \langle 0| a^{m-1} = \langle \mathscr{R}| .$$

It remains finally to identify a realization of *observables* for single-species chemical reaction systems. It is straightforward to identify the *number operator*

$$\hat{n} := a^{\dagger} a \qquad (33)$$

as an observable, since due to

$$\hat{n} |n\rangle = n |n\rangle \qquad (34)$$

every pure state $|n\rangle$ is an eigenvector of $\hat{n}$ as required for $\hat{n}$ to qualify as an observable according to Definition 3. Referring to equation (142) in Appendix A.3 for the precise details, every element of the HW algebra may be expressed as a linear combination of so-called *normal-ordered* terms $a^{\dagger r} a^s$,

thus there exists a whole family of operators that are diagonal in the basis spanned by the vectors $|n\rangle$:

$$\forall k \in \mathbb{Z}_{\geq 0}: \quad a^{\dagger k} a^k |n\rangle = (n)_k |n\rangle . \qquad (35)$$

However, there is an important relationship well-known from the combinatorics literature between powers of $\hat{n}$ and members of the above family:

Theorem 2. *For arbitrary $k, \ell \in \mathbb{Z}_{\geq 0}$, we have that*

$$a^{\dagger k} a^k = \sum_{\ell=0}^{k} s_1(k,\ell) \hat{n}^{\ell}, \quad \hat{n}^{\ell} = \sum_{k=0}^{\ell} S_2(\ell,k) a^{\dagger k} a^k, \qquad (36)$$

where $s_1(k,\ell)$ are the (unsigned) Stirling numbers of the first kind, *while $S_2(\ell,k)$ denotes the* Stirling numbers of the second kind *(see Appendix B for the explicit definitions).*

Proof. See for example [8]. □

What this theorem entails is that for single-species reaction systems we may describe the entire statistics in terms of either the expectation value and higher order moments of the number operator $\hat{n}$, or alternatively in terms of the expectation values of the operators $a^{\dagger k} a^k$. The latter are referred to in the statistics literature as *factorial moments*.

With these preparations, it becomes evident that the combinatorics of normal ordering and related techniques can be of use in the study of chemical reaction systems in two main ways: either to explicitly compute the time-dependent state $|\Psi(t)\rangle$ (see Section 6), or alternatively in order to formulate and solve evolution equations for observable generating functions. The former can be accomplished in a satisfyingly concise way:

Theorem 3. *Given a one-species chemical reaction system according to Definition 6, the* exponential moment generating function $\mathscr{M}(t;\lambda)$ *of the number operator $\hat{n}$,*

$$\mathscr{M}(t;\lambda) := \left\langle \mathscr{R} \left| e^{\lambda \hat{n}} \right| \Psi(t) \right\rangle , \qquad (37)$$

fulfills the evolution equation

$$\frac{\partial}{\partial t}\mathcal{M}(t;\lambda) = \mathbb{D}(\lambda,\partial_\lambda)\mathcal{M}(t;\lambda)$$

$$\mathcal{D}(\lambda,\partial_\lambda) = \sum_{i,o} r_{i,o} \left(e^{\lambda(o-i)}-1\right) \sum_{\ell=0}^{i} S_1(i,\ell)\left(\frac{\partial}{\partial\lambda}\right)^{\ell}.$$

(38)

Proof. See Appendix C.2 □

Another interesting variant of generating function evolution equations is provided upon considering the *factorial moment generating function* $\mathcal{F}(t;\nu)$, defined as

$$\mathcal{F}(t;\nu) := \sum_{n=0}^{\infty} \frac{\nu^n}{n!} \left\langle \mathcal{R} \left| a^{\dagger n} a^n \right| \Psi(t) \right\rangle.$$

(39)

Because the operation $\langle\mathcal{R}|\ldots|\Psi(t)\rangle$ of taking the expectation value in the state $|\Psi(t)\rangle$ is *linear* in its argument, $\mathcal{F}(t;\nu)$ may be seen as a kind of "change of basis" for the observables of the stochastic system, i.e. from moments of the number operator to factorial moments.

Theorem 4. *The generating function $\mathcal{F}(t;\nu)$ has the following evolution equation:*

$$\frac{\partial}{\partial t}\mathcal{F}(t;\nu) = \mathbb{d}(\nu,\partial_\nu)\mathcal{F}(t;\nu)$$

$$\mathbb{d}(\nu,\partial_\nu) = \sum_{i,o} r_{i,o} \left((\nu+1)^o - (\nu+1)^i\right)\left(\frac{\partial}{\partial\nu}\right)^i.$$

(40)

Proof. See Appendix C.3 □

In summary, chemical reaction systems of one species may be described equivalently in terms of either of the three formulae:

(i) the explicit formula for the *probability distribution*,
$$|\Psi(t)\rangle = \mathcal{E}(t)|\Psi(0)\rangle,$$

(ii) the evolution equation for the *exponential moment generating function $\mathcal{M}(t;\lambda)$ of the number operator $\hat{n}$,*
$$\frac{\partial}{\partial t}\mathcal{M}(t;\lambda) = \mathbb{D}(\lambda,\partial_\lambda)\mathcal{M}(t;\lambda),$$

(iii) the evolution equation for the *factorial moment generating function $\mathcal{F}(t;\nu)$,*
$$\frac{\partial}{\partial t}\mathcal{F}(t;\nu) = \mathbb{d}(\nu,\partial_\nu)\mathcal{F}(t;\nu).$$

We summarize the explicit formulae for individual transitions $iA \xrightarrow{r_{i,o}} oA$ in Table 1, both for generic transitions and for the ones of empirical relevance in the theory of chemical reaction systems.

4 Multi-species reaction systems

The key ingredient will be the canonical representation of the multi-species Heisenberg-Weyl (HW) algebra (over the field $\mathbb{R}$):

Definition 8. *Let* **S** *be a set of species (i.e. in general a denumerable set). Define a set of basis vectors*

$$|\underline{n}\rangle \equiv |n_{i_1}, n_{i_2}, \ldots\rangle,$$

(41)

called the multi-species number vectors. *Since each of these vectors will be interpreted to represent a pure state, whence a realizable configuration of the system, we require in addition that all n_i must be non-negative integers, and that only finitely many of the n_i are non-zero. Then the* canonical representation of the multi-species Heisenberg-Weyl algebra *is defined via*

$$a_i^\dagger |\underline{n}\rangle := |\underline{n}+\underline{\Delta}_i\rangle$$
$$a_j |\underline{n}\rangle := (n_j)_1 |\underline{n}-\underline{\Delta}_j\rangle,$$

(42)

with $(n_j)_1$ denoting the falling factorials

$$(n_j)_1 = \Theta(n_j-1)\frac{(n_j)!}{(n_j-1)!}.$$

We have moreover introduced the convenient shorthand notation $\underline{\Delta}_i$ for a vector of non-negative integers whose only non-zero coordinate is $n_i = 1$.

Table 1 Contributions $H_{(i,o)}$ to the evolution operator H, $\mathbb{D}_{(i,o)}$ to the differential operator $\mathbb{D} \equiv \mathbb{D}(\lambda, \partial_\lambda)$ and $\mathrm{d}_{(i,o)}$ to $\mathrm{d} \equiv \mathrm{d}(\nu, \partial_\nu)$ of individual chemical reactions $iA \xrightarrow{r_{i,o}} oA$

Parameters	$H_{(i,o)}$	$\mathbb{D}_{(i,o)}$	$\mathrm{d}_{(i,o)}$
(i,o)	$r_{i,o}\left(a^{\dagger o} - a^{\dagger i}\right)a^i$	$r_{i,o}\left(e^{\lambda(o-i)} - 1\right)\sum_{\ell=0}^{i} S_1(i,\ell)\left(\frac{\partial}{\partial\lambda}\right)^\ell$	$r_{i,o}\left((\nu+1)^o - (\nu+1)^i\right)\left(\frac{\partial}{\partial\nu}\right)^i$
$(0,2)$	$r_{0,2}\left(a^{\dagger 2} - 1\right)$	$r_{0,2}\left(e^{2\lambda} - 1\right)$	$r_{0,2}\left(\nu^2 + 2\nu\right)$
$(0,1)$	$r_{0,1}(a^\dagger - 1)$	$r_{0,1}\left(e^\lambda - 1\right)$	$r_{0,1}\nu$
$(1,2)$	$r_{1,2}\left(a^{\dagger 2} - a^\dagger\right)a$	$r_{1,2}\left(e^\lambda - 1\right)\frac{\partial}{\partial\lambda}$	$r_{1,2}(\nu^2 + \nu)\frac{\partial}{\partial\nu}$
$(1,0)$	$r_{1,0}\left(1 - a^\dagger\right)a$	$r_{1,0}\left(e^{-\lambda} - 1\right)\frac{\lambda}{\partial\lambda}$	$-r_{1,0}\nu\frac{\partial}{\partial\nu}$
$(2,1)$	$r_{2,1}\left(a^\dagger - a^{\dagger 2}\right)a^2$	$r_{2,1}\left(e^{-\lambda} - 1\right)\left(\frac{\partial^2}{\partial\lambda^2} - \frac{\partial}{\partial\lambda}\right)$	$-r_{2,1}(\nu^2 + \nu)\frac{\partial^2}{\partial\nu^2}$
$(2,0)$	$r_{2,0}\left(1 - a^{\dagger 2}\right)a^2$	$r_{2,0}\left(e^{-2\lambda} - 1\right)\left(\frac{\partial^2}{\partial\lambda^2} - \frac{\partial}{\partial\lambda}\right)$	$-r_{2,0}(\nu^2 + 2\nu)\frac{\partial^2}{\partial\nu^2}$

The fact that these linear operators foster a representation of the multi-species HW algebra is confirmed by noting that the following *canonical commutation relations** hold true:

$$[a_i, a_j^\dagger] = \delta_{i,j}, \quad [a_i, a_j] = 0 = [a_i^\dagger, a_j^\dagger]. \quad (43)$$

The following additional multi-index notations will prove convenient throughout this section:

$$\underline{n}! \equiv \prod_{i\in\mathbf{S}}(n_i)!, \quad \underline{x}^{\underline{y}} \equiv \prod_{i\in\mathbf{S}}\left(x_i^{y_i}\right), \quad e^{\underline{x}} \equiv \prod_{i\in\mathbf{S}}e^{x_i}$$

$$(\underline{m})_{\underline{n}} \equiv \prod_{i\in\mathbf{S}}\left((m_i)_{n_i}\right) := \prod_{i\in\mathbf{S}}\left(\frac{(m_i)!}{(m_i - n_i)!}\right) \quad (44)$$

$$\binom{\underline{m}}{\underline{n}} = \prod_{i\in\mathbf{S}}\binom{m_i}{n_i}, \quad \delta_{\underline{m},\underline{n}} \equiv \prod_{i\in\mathbf{S}}\delta_{m_i,n_i}.$$

An immediate consequence of (43) is that we can express any element of the multi-species HW algebra in terms of linear combinations of *normal-ordered* expressions of the form $\underline{a}^{\dagger\underline{r}}\underline{a}^{\underline{s}}$ (with $\underline{r},\underline{s}$ vectors of non-negative integers), which is a consequence of the well-known *normal ordering formula*

* More precisely, each of the relations in (43) is an equation on linear operators, whose validity may be verified by acting on generic basis vectors $|\underline{n}\rangle$.

(cf. e.g. [8])

$$\underline{a}^{\dagger\underline{m}}\underline{a}^{\underline{n}}\underline{a}^{\dagger\underline{r}}\underline{a}^{\underline{s}} = \sum_{\underline{k}\geq\underline{0}}\underline{k}!\binom{\underline{n}}{\underline{k}}\binom{\underline{r}}{\underline{k}}\underline{a}^{\dagger(\underline{m}+\underline{r}-\underline{k})}\underline{a}^{(\underline{n}+\underline{s}-\underline{k})}. \quad (45)$$

Here, we have made use of the standard convention $\binom{x}{y} = 0$ whenever $y > x$.

We will use the following normalization for the (formal) dual vectors $\langle\underline{m}|$:

$$\langle\underline{m}|\underline{n}\rangle := \underline{n}!\delta_{\underline{m},\underline{n}}. \quad (46)$$

The multi-species version $\langle\mathscr{R}|$ of the *reference dual vector* is defined as

$$\langle\mathscr{R}| := \sum_{\underline{m}\geq\underline{0}}\frac{1}{\underline{m}!}\langle\underline{m}|. \quad (47)$$

With the relations

$$|\underline{n}\rangle = \underline{a}^{\dagger\underline{n}}|\underline{0}\rangle, \quad \langle\underline{m}| = (\underline{a}^{\underline{m}}|\underline{0}\rangle)^\dagger = \langle\underline{0}|\underline{a}^{\underline{m}}, \quad (48)$$

which are compatible with $\langle\underline{m}|\underline{n}\rangle = \underline{n}!\delta_{\underline{m},\underline{n}}$, we may identify $\langle\mathscr{R}|$ as a *simultaneous left eigenvector* of the creation operators $a_i^\dagger$ ($i\in\mathfrak{S}$) of eigenvalue 1:

$$\langle\mathscr{R}|a_i^\dagger = \langle\underline{0}|e^{\underline{a}}a_i^\dagger = \langle\underline{0}|(a_i^\dagger + 1)e^{\underline{a}} = \langle\mathscr{R}|. \quad (49)$$

Here, we have made use of the following very useful general formula (which played a pivotal role in the derivation of several analytical combinatorics formulae, see Section 6):

Proposition 1. *For all real vectors $\underline{\alpha}, \underline{\beta}$, and for all Taylor-expandable functions $f(\underline{a}^{\dagger}, \underline{a})$, we have that*

$$e^{\underline{\alpha} \cdot \underline{a}} f(\underline{a}^{\dagger}, \underline{a}) e^{\underline{\beta} \cdot \underline{a}^{\dagger}} = e^{\underline{\alpha} \cdot \underline{\beta}} e^{\underline{\beta} \cdot \underline{a}^{\dagger}} f\left(\underline{a}^{\dagger} + \underline{\alpha}, \underline{a} + \underline{\beta}\right) e^{\underline{\alpha} \cdot \underline{a}}.$$
(50)

Proof. See Appendix C.4. □

Noticing that the multi-species number vectors are simultaneous eigenstates of the *number operators* $\hat{n}_i := a_i^{\dagger} a_i$ ($i \in \mathfrak{S}$),

$$\hat{n}_i |\underline{n}\rangle = (n_i)_1 |\underline{n}\rangle,$$
(51)

we have all preparations in place to state the multi-species generalization of the stochastic mechanics framework for chemical reaction systems:

Definition 9. *Consider a chemical reaction system with reactions of the form*

$$\underline{i} \cdot \underline{A} \xrightarrow{r_{\underline{i},\underline{o}}} \underline{o} \cdot \underline{A},$$
(52)
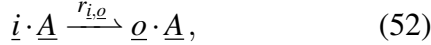
where $\underline{i}, \underline{o}$ are vectors of non-negative integers (encoding the numbers of particles of each type on input and output of the transition, respectively), and with the reaction rates* $r_{\underline{i},\underline{o}} \in \mathbb{R}_{\geq 0}$. Then together with an input state

$$|\Psi(0)\rangle \in Prob(\mathcal{N})$$
(53)

over the state space $\mathcal{N}$ of multi-species number vectors,

$$Prob(\mathcal{N}) \ni |\Psi(0)\rangle \equiv \sum_{\underline{n} \geq \underline{0}} \psi_{\underline{n}}(0) |\underline{n}\rangle,$$
(54)

this data specifies a continuous-time Markov chain with evolution operator

$$H := \sum_{\underline{i},\underline{o}} r_{\underline{i},\underline{o}} \left(\underline{a}^{\dagger \underline{o}} - \underline{a}^{\dagger \underline{i}}\right) \underline{a}^{\underline{i}}.$$
(55)

---

* We use this convention in order to be able to specify a compact formula for H, i.e. only the reactions for parameter pairs $(\underline{i}, \underline{o})$ for $r_{\underline{i},\underline{o}} > 0$ contribute to H.

The main general result on the dynamics of generic chemical reaction systems are the following two generalizations of the moment and factorial moment generating function evolution equations from the previous section. As in the single-species setting, knowing all Eigenvalues n_i of a given pure state $|\underline{n}\rangle$ completely characterizes the Eigenstate, such that the moments of the number operators $\hat{n}_i$ or the respective factorial moments are providing equivalent knowledge about a system state $|\Psi(t)\rangle$ as the probability distribution itself:

Theorem 5. *For a multi-species chemical reaction system as specified in Definition 9, let the exponential moment generating function $\mathcal{M}(t; \underline{\lambda})$ of the number operators $\hat{n}_i$ and the factorial moment generating function $\mathcal{F}(t; \underline{v})$ be defined as*

$$\mathcal{M}(t; \underline{\lambda}) := \left\langle \mathcal{R} \left| e^{\underline{\lambda} \cdot \hat{\underline{n}}} \right| \Psi(t) \right\rangle$$
$$\mathcal{F}(t; \underline{v}) := \sum_{\underline{k} = \underline{0}}^{\infty} \frac{\underline{v}^{\underline{k}}}{\underline{n}!} \left\langle \mathcal{R} \left| \underline{a}^{\dagger \underline{k}} \underline{a}^{\underline{k}} \right| \Psi(t) \right\rangle.$$
(56)

Then the generating functions satisfy the evolution equations

$$\frac{\partial}{\partial t} \mathcal{M}(t; \underline{\lambda}) = \mathbb{D}(\underline{\lambda}, \partial_{\underline{\lambda}}) \mathcal{M}(t; \underline{\lambda})$$
$$\mathbb{D}(\underline{\lambda}, \partial_{\underline{\lambda}}) = \sum_{\underline{i},\underline{o}} r_{\underline{i},\underline{o}} \left(e^{\underline{\lambda} \cdot (\underline{o} - \underline{i})} - 1\right) \cdot$$
$$\cdot \sum_{\underline{k} = \underline{0}}^{\underline{i}} \underline{s}_1(\underline{i}, \underline{k}) \left(\frac{\partial}{\partial \underline{\lambda}}\right)^{\underline{k}}$$
$$\underline{s}_1(\underline{i}, \underline{k}) := \prod_{j \in \mathbf{S}} s_1(i_j, k_j),$$
(57)

*with $s_1(i, j)$ denoting the unsigned Stirling numbers of the first kind, and*

$$\frac{\partial}{\partial t} \mathcal{F}(t; \underline{v}) = \mathbb{d}(\underline{v}, \partial_{\underline{v}}) \mathcal{F}(t; \underline{v})$$
$$\mathbb{d}(\underline{v}, \partial_{\underline{v}}) = \sum_{\underline{i},\underline{o}} r_{\underline{i},\underline{o}} \left((\underline{v} + \underline{1})^{\underline{o}} - (\underline{v} + \underline{1})^{\underline{i}}\right) \left(\frac{\partial}{\partial \underline{v}}\right)^{\underline{i}}.$$
(58)

Proof. See Appendix C.5. □

11

We summarize both these statements as well as explicit formulae for all chemical reactions of practical interest (i.e. for transitions with $1 \leq \sum_{j \in \mathfrak{S}} (i_j + o_j) \leq 3$) in Table 2.

As an aside and to conclude the discussion of this general framework, it may be worthwhile noting that while as in the single-species cases only a very limited number of reaction parameters $(\underline{i}, \underline{o})$ is realized in realistic chemical reaction systems, the multi-species formulae presented thus far in fact can describe *generic discrete transition systems*, whence *discrete graph rewriting systems* (compare [4, 6]). Interesting special cases would for example include *branching processes* such as

$$A_i \xrightarrow{r_{i,\underline{o}}} \underline{o} \cdot \underline{A}.$$
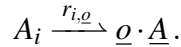
Since branching processes constitute a class of *semi-linear transition systems*, one could in principle envision to provide exact closed-form formulae for the respective time-dependent states $|\Psi(t)\rangle$ as well as for the generating functions $\mathscr{M}(t; \underline{\lambda})$ and $\mathscr{F}(t; \underline{v})$ along the lines presented in Section 6.

5 Reaction systems with first order moment closure

The stochastic mechanics framework in its "three-fold view" of chemical reaction systems in terms of time-dependent state $|\Psi(t)\rangle$, exponential moment generating function $\mathscr{M}(t; \underline{\lambda})$ and factorial moment generating function $\mathscr{F}(t; \underline{v})$, offers a number of very interesting insights into the combinatorics that govern chemical transitions. While in Section 6 techniques will be presented to derive explicit formulae for $|\Psi(t)\rangle$, in this section we will focus on some general properties of moment evolution equations.

In practical applications, amongst the simplest to analyze reaction systems are those that possess the property of *first order moment closure*, whence for

which the evolution equations for the first moments of the number operators $\hat{n}_i$ read

$$\frac{d}{dt} \langle \mathscr{R} | \hat{n}_i | \Psi(t) \rangle = \sum_{j \in \mathbf{S}} \mu_{ij} \langle \mathscr{R} | \hat{n}_j | \Psi(t) \rangle, \quad (59)$$

with (μ_{ij}) a matrix of coefficients. In other words, the special property encoded in (59) is that the time-derivatives of the expectation values of the number operators only depend on the expectation values of the number operators, and not on any higher moments thereof. It is straightforward to derive the class of reaction systems with this property from the evolution equation for the factorial moments (noting that $\hat{n}_i = a_i^\dagger a_i$, whence that the first moments and first factorial moments coincide). To this end, let us expand the factorial moment generating function $\mathscr{F}(t; \underline{v})$ more explicitly (compare (56)):

$$\begin{aligned} \mathscr{F}(t; \underline{v}) &= \sum_{\underline{n} \geq 0} \frac{\underline{v}^{\underline{n}}}{\underline{n}!} \langle \mathscr{R} | \underline{a}^{\dagger \underline{n}} \underline{a}^{\underline{n}} | \Psi(t) \rangle \\ &\equiv \sum_{\underline{n} \geq 0} \frac{\underline{v}^{\underline{n}}}{\underline{n}!} f_{\underline{n}}(t). \end{aligned} \quad (60)$$

Then by expanding the evolution equation for $\mathscr{F}(t; \underline{n})$ and collecting terms contributing to $\frac{d}{dt} f_{\underline{n}}(t)$, we obtain the following:

Theorem 6. *Consider a multi-species chemical reaction system as specified in Definition 9. Then the evolution equations for the factorial moments $f_{\underline{n}}(t)$ read*

$$\begin{aligned} \frac{d}{dt} f_{\underline{n}}(t) &= \sum_{\underline{i}, \underline{o}} r_{\underline{i}, \underline{o}} \sum_{\underline{k}} \phi_{\underline{k}}(\underline{n}; \underline{i}, \underline{o}) f_{\underline{n} + \underline{i} - \underline{k}}(t) \\ \phi_{\underline{k}}(\underline{n}; \underline{i}, \underline{o}) &= \left[(\underline{o})_{\underline{k}} - (\underline{i})_{\underline{k}} \right] \binom{\underline{n}}{\underline{k}}. \end{aligned} \quad (61)$$

Here, we have made use of the standard conventions $\binom{x}{y} = 0$ and $(x)_y = 0$ whenever $y > x$.

Specializing to $\underline{n} = \underline{\Delta}_\alpha$, i.e. to $f_{\underline{\Delta}_\alpha}$ denoting the first moment of the number vector $\hat{n}_\alpha$, one obtains the evolution equations for the first moments as

$$\frac{d}{dt} f_{\underline{\Delta}_\alpha}(t) = \sum_{\underline{i}, \underline{o}} r_{\underline{i}, \underline{o}} (o_\alpha - i_\alpha) f_{\underline{i}}(t). \quad (62)$$

Table 2 Contributions $H_{(\underline{i},\underline{o})}$ to the evolution operator H, $\mathbb{D}_{(\underline{i},\underline{o})}$ to the differential operator $\mathbb{D} \equiv \mathbb{D}(\underline{\lambda}, \partial_{\underline{\lambda}})$ and $\mathfrak{d}_{(\underline{i},\underline{o})}$ to $\mathfrak{d} \equiv \mathfrak{d}(\underline{v}, \partial_{\underline{v}})$ of individual multi-species chemical reactions $\underline{i}\cdot\underline{A} \xrightarrow{r_{i,o}} \underline{o}\cdot\underline{A}$

Parameters	$H_{(\underline{i},\underline{o})}/r_{i,o}$	$\mathbb{D}_{(\underline{i},\underline{o})}/r_{i,o}$	$\mathfrak{d}_{(\underline{i},\underline{o})}/r_{i,o}$
$(\underline{o},\underline{i})$	$\left(\underline{a}^{\dagger\,\underline{o}} - \underline{a}^{\dagger\,\underline{i}}\right)\underline{a}^{\underline{i}}$	$\left(e^{\underline{\lambda}\cdot(\underline{o}-\underline{i})}-1\right)\sum_{\underline{\ell}=\underline{0}}^{\underline{i}}S_1(\underline{i},\underline{\ell})\left(\frac{\partial}{\partial\underline{\lambda}}\right)^{\underline{\ell}}$	$\left((\underline{v}+\underline{1})^{\underline{o}}-(\underline{v}+\underline{1})^{\underline{i}}\right)\left(\frac{\partial}{\partial\underline{v}}\right)^{\underline{i}}$
$(\underline{\delta}_\alpha+\underline{\delta}_\beta,\underline{0})$	$\left(a_\alpha^\dagger + a_\beta^\dagger - 1\right)$	$\left(e^{\lambda_\alpha+\lambda_\beta}-1\right)$	$v_\alpha v_\beta + v_\alpha + v_\beta$
$(\underline{\delta}_\alpha,\underline{0})$	$a_\alpha^\dagger - 1$	$e^{\lambda_\alpha}-1$	v_α
$(\underline{0},\underline{\delta}_\gamma)$	$\left(1-a_\gamma^\dagger\right)a_\gamma$	$\left(e^{-\lambda_\gamma}-1\right)\frac{\partial}{\partial\lambda_\gamma}$	$-v_\gamma\frac{\partial}{\partial v_\gamma}$
$(\underline{\delta}_\alpha,\underline{\delta}_\beta)$	$\left(a_\alpha^\dagger - a_\beta^\dagger\right)a_\beta$	$\left(e^{\lambda_\alpha-\lambda_\beta}-1\right)\frac{\partial}{\partial\lambda_\beta}$	$(v_\alpha-v_\beta)\frac{\partial}{\partial v_\beta}$
$(\underline{\delta}_\alpha+\underline{\delta}_\beta,\underline{\delta}_\gamma)$	$\left(a_\alpha^\dagger a_\beta^\dagger - a_\gamma^\dagger\right)a_\gamma$	$\left(e^{\lambda_\alpha+\lambda_\beta-\lambda_\gamma}-1\right)\frac{\partial}{\partial\lambda_\gamma}$	$(v_\alpha v_\beta + v_\alpha + v_\beta - v_\gamma)\frac{\partial}{\partial v_\gamma}$
$(\underline{0},\underline{\delta}_\beta+\underline{\delta}_\gamma)$	$\left(1-a_\beta^\dagger a_\gamma^\dagger\right)a_\beta a_\gamma$	$\left(e^{-\lambda_\beta-\lambda_\gamma}-1\right)\left(\frac{\partial^2}{\partial\lambda_\beta\partial\lambda_\gamma}-\delta_{\beta,\gamma}\frac{\partial}{\partial\lambda_\gamma}\right)$	$-\left(v_\beta v_\gamma + v_\beta + v_\gamma\right)\frac{\partial^2}{\partial v_\beta\partial v_\gamma}$
$(\underline{\delta}_\alpha,\underline{\delta}_\beta+\underline{\delta}_\gamma)$	$\left(a_\alpha^\dagger - a_\beta^\dagger a_\gamma^\dagger\right)a_\beta a_\gamma$	$\left(e^{\lambda_\alpha-\lambda_\beta-\lambda_\gamma}-1\right)\left(\frac{\partial^2}{\partial\lambda_\beta\partial\lambda_\gamma}-\delta_{\beta,\gamma}\frac{\partial}{\partial\lambda_\gamma}\right)$	$(v_\alpha-v_\beta v_\gamma - v_\beta - v_\gamma)\frac{\partial^2}{\partial v_\beta\partial v_\gamma}$

Therefore, the only reaction systems for which we have first order moment closure *are* semi-linear *reaction systems, i.e. systems for which for all* $(\underline{i},\underline{o})$ *with rates* $r_{\underline{i},\underline{o}} > 0$ *we have that*

$$\sum_{j\in\mathbf{S}} i_j \leq 1 . \tag{63}$$

The fact that semi-linear reaction (and branching) systems are the only systems with first order moment closure already indicates that they are somewhat simpler in their dynamical structure than generic systems. In Section 6 we will demonstrate in fact that all of them can be solved analytically. In view of practical applications, the following no-go corollary is important to note:

Corollary 1. *The factorial moment evolution equation* (61) *entails that there are* no factorial moment closures *beyond semi-linear reaction systems. Because of the relationship between moments of the number operators and the factorial moments as presented in Theorem 2, the same statement holds true for the moments of the number operators.*

In other words, we will in particular not be able to rely on a simple matrix exponential in order to solve the evolution equations for generic reaction systems. Nonetheless, there still exist interesting cases in which one encounters a form of a higher-order factorial moment closure, namely when we consider catalytic reaction systems (or other systems in which the numbers of certain types of particles are conserved): suppose that the reaction system leaves the number n_i of a certain species of particles constrained in the range $0 \leq n_i \leq N_i$ when initialized on a state $|\Psi(0)\rangle$ that satisfies this constraint. Then the factorial moments of $\hat{n}_i$ of orders greater than N_i will vanish identically, which may in certain cases provide some interesting computation strategies. We refer the interested readers to the standard literature (such as the review paper [2]) for explicit examples and further technical details.

6 Analytical solution strategies for probability generating functions

For the material presented in this section, it will prove most convenient to work in the Bargmann-Fock representation of the Heisenberg-Weyl alge-

bra, whence in other words to focus on the dynamical evolution properties of the *probability generating function*

$$P(t;\underline{x}) = \sum_{\underline{n} \geq \underline{0}} p_{\underline{n}}(t) \underline{x}^{\underline{n}} \tag{64}$$

of the state

$$|\Psi(t)\rangle = \sum_{\underline{n} \geq \underline{0}} p_{\underline{n}}(t) |\underline{n}\rangle \tag{65}$$

of the chemical reaction system. We will employ two rather different strategies to make progress: the first one involves the use of a result on boson normal-ordering by some of the authors for the special case of reactions that are not binary (i.e. that consume at most one particle per elementary transition), while the second method, applicable precisely to said binary reactions, involves the use of *(Sobolev-) orthogonal polynomials*.

6.1 Solving non-binary reactions via semilinear boson normal-ordering

Our main technical tool will be a normal-ordering formula developed by some of the authors [7] (presented here in the multi-variate version as e.g. described in [10]):

Theorem 7. *Consider an N-species Heisenberg-Weyl algebra, whose generators read in the Bargmann-Fock representation $\hat{x}_i$ (multiplication by x_i) and ∂_{x_j} (partial derivation by x_j) for $1 \leq i, j \leq N$. Let $\mathscr{H}$ (i.e. the Hamiltonian $H \equiv H(\underline{a}^\dagger, \underline{a})$ according to (55) expressed in the Bargmann-Fock basis following (26)) be a* semi-linear *expression in the generators,*

$$\mathscr{H} = v(\underline{\hat{x}}) + \sum_{i=0}^{N} q_i(\underline{\hat{x}}) \, \partial_{x_i}, \tag{66}$$

with $q_i(\underline{\hat{x}})$ and $v(\underline{\hat{x}})$ some functions in the operators $\hat{x}_i$. Let $F(0;\underline{x})$ be an entire function in the indeterminates x_i. Define the formal power series

$$F(\lambda;\underline{x}) := e^{\lambda \mathscr{H}} F(0;\underline{x}) \tag{67}$$

with formal variable λ, which is the exponential generating function *for terms of the form*

$$\mathscr{H}^n F(0;\underline{x}),$$

thus describing the n-fold application of $\mathscr{H}$ to $F(0;\underline{x})$. Then $F(\lambda;\underline{x})$ may be expressed in closed form as follows[*]

$$F(\lambda;\underline{x}) = g(\lambda;\underline{x}) F\big(0; \underline{T}(\lambda;\underline{x})\big) \tag{68a}$$

$$\frac{\partial}{\partial \lambda} T_i(\lambda;\underline{x}) = q_i(\underline{T}(\lambda;\underline{x})), \quad T_i(0;\underline{x}) = x_i \tag{68b}$$

$$g(\lambda;\underline{x}) := \int_0^\lambda v(\underline{T}(\kappa;\underline{x})) d\kappa \tag{68c}$$

Moreover, we have the relations

$$\begin{aligned} \underline{T}(\lambda + \mu;\underline{x}) &= \underline{T}(\mu;\underline{T}(\lambda;\underline{x})) \\ g(\lambda + \mu;\underline{x}) &= g(\lambda;\underline{x}) g(\mu;\underline{T}(\lambda;\underline{x})), \end{aligned} \tag{69}$$

which in turn implies that $e^{\lambda \mathscr{H}}$ induces a one-parameter group *of transformations.*

Note that in the above formulation, we understand the notion of integration as a notion of *formal* integration, whence as the operation on formal power series induced by the action of formal integration on monomials in the formal variables,

$$\int dx_i \, x_i^{n_i} = \frac{x_i^{n_i+1}}{(n_i+1)}.$$

As might not be immediately evident to the non-combinatorists, this result in fact provides a closed-form solution to a vast class of normal-ordering problems. Recall that in the HW algebra every element may be presented as a linear combination in normal-ordered terms, i.e. in terms of the form $\underline{\hat{x}}^r \partial_{\underline{x}}^s$

---

[*] Equation (68c) is sometimes expressed (see e.g. [7]) in the alternative form

$$\frac{\partial}{\partial \lambda} \ln(g(\lambda;\underline{x})) = v(\underline{T}(\lambda;\underline{x})), \quad g(0;\underline{x}) = 1.$$

(for multi-indices $\underline{r}, \underline{s} \geq \underline{0}$). To illustrate the practical value of Equation (68a), consider for example computing the term $\mathscr{H}^2$ explicitly:

$$
\begin{aligned}
\mathscr{H}^2 &= \left(v(\hat{\underline{x}}) + \sum_{i=1}^{n} q_i(\hat{\underline{x}}) \partial_{x_i} \right)^{\circ 2} \\
&=: \left(v + \underline{q} \cdot \partial_{\underline{x}} \right)^2 : \\
&\quad + \sum_{i=1}^{n} q_i \, \partial_{x_i} + \sum_{i,j=1}^{n} q_i \left(\frac{\partial q_j}{\partial x_i} \right) \partial_{x_j} \,.
\end{aligned}
$$

Here, the notation $: \ldots :$ stands for the *forgetful normal-ordering* operation (c.f. e.g. [9]), defined as bringing the expression $\ldots$ into normal-ordered form *ignoring* the commutation relations

$$
[\hat{x}_i, \partial_{x_j}] = \delta_{i,j} \,.
$$

Thus, in the example above we find concretely

$$
: \left(v + \underline{q} \cdot \partial_{\underline{x}} \right)^2 : \; = v^2 + 2 v \underline{q} \cdot \partial_{\underline{x}} + \sum_{i,j=1}^{n} q_i q_j \partial_{x_i} \partial_{x_j} \,.
$$

As this example computation demonstrates, finding closed-form solutions to the normal-ordering problem is a highly intricate task, whence it is convenient that for the semi-linear operators Theorem 7 provides such a practicable solution.

Returning to the main topic of the present paper, we are now in a position to formulate one of our key results. The argument employs some standard notions and results from the theory of probability distributions and of probability generating functions (see e.g. the modern textbooks [31, 32]), presented for convenience in the following proposition:

Proposition 2. *(i) (*Multiplication theorem*) Let $X_1, \ldots, X_N$ be independent discrete random variables taking values in $\mathbb{Z}_{\geq 0}$, with probability generating functions $F_1(x), \ldots, F_N(x)$. Then $X_1 + \ldots + X_N$ is a discrete random variable with probability generating function (PGF)*

$$
F(x) = \prod_{i=1}^{N} F_i(x) \,. \tag{70}
$$

On the level of discrete probability distributions, this amounts to computing the convolution *of the distributions of the random variables X_i.*

(ii) Let M be a discrete random variable taking values in $\mathbb{Z}_{\geq 0}$ with PGF $F_M(x)$, and denote by $\{X_i\}_{i \in \mathbb{Z}_{\geq 0}}$ a family of independent identically distributed discrete random variables taking values in $\mathbb{Z}_{\geq 0}$ with common PGF $F_X(x)$. Then the compound distribution *of the discrete random variable*

$$
Y = X_1 + \ldots + X_M
$$

has the probability generating function

$$
F_Y(x) = F_M(F_X(x)) \,. \tag{71}
$$

Moreover, both concepts generalize in a straightforward fashion to multivariate generating functions.

With these preparations, we find:

Theorem 8. *Consider an N-species chemical reaction system specified according to Definition 9. Let $P(0; \underline{x})$ be the generating function of a probability distribution of a random variable taking values in N-tuples of non-negative integers, whence*

$$
P(0; \underline{x}) = \sum_{\underline{n} \geq \underline{0}} p_{\underline{n}}(0) \underline{x}^{\underline{n}}, \; P(0; \underline{1}) = \sum_{\underline{n} \geq \underline{0}} p_{\underline{n}}(0) = 1 \,,
\tag{72}
$$

and with all coefficients $p_{\underline{n}}(0)$ non-negative real numbers. Let $\mathscr{H}$ be the infinitesimal generator of the system (whence the operator H of Definition 9 expressed in the Bargmann-Fock basis). Then for each of the non-binary elementary chemical reactions, acting with $e^{\lambda \mathscr{H}}$ on the initial state's probability generating function $P(0; \underline{x})$ results in a new generating function $P(\lambda; \underline{x})$,

$$
P(\lambda; \underline{x}) = e^{\lambda \mathscr{H}} P(0; \underline{x}) = g(\lambda; \underline{x}) P\big(0; \underline{T}(\lambda; \underline{x})\big) \,.
\tag{73}
$$

The resulting $P(\lambda; \underline{x})$ is a well-posed probability distribution for all values of $\lambda \in \mathbb{R}_{\geq 0}$.

Proof. Referring to the concrete results of applying Theorem 7 as listed in Table 3, in all cases we have that $g(\lambda;\underline{x})$ as well as $\underline{T}(\lambda;\underline{x})$ are in fact themselves probability generating functions of certain discrete N-multivariate random variables. Therefore, by a combination of equations (70) and (71), the formula for $P(\lambda;\underline{x})$ as stated identifies $P(\lambda;\underline{x})$ as the PGF of the convolution of the distribution encoded in $g(\lambda;\underline{x})$ with the compound distribution encoded in $P(0;\underline{T}(\lambda;\underline{x}))$ whenever $\lambda \in \mathbb{R}_{\geq 0}$. $\qquad\square$

In this result, it is important to note that while for arbitrary generating functions $F(\lambda;\underline{x})$ such as those mentioned in Theorem 7 we find one-parameter *groups* of transformations*, specializing to generating functions that are PGFs we have to also comply with the requirement that all coefficients of the generating functions must be *non-negative* real numbers. Upon closer inspection of the concrete formulae provided for $P(\lambda;\underline{x})$, one can verify that $P(\lambda;\underline{x})$ with the formal parameter λ set to a real number value will yield a proper PGF if and only if λ is a non-negative real number. Therefore, truncating the parameter space to $\lambda \in \mathbb{R}_{\geq 0}$ entails that the one-parameter *group* described by Theorem 7 specializes to a *one-parameter semi-group* (of real parameter λ): denote by $\mathscr{E}(\lambda)$ the operator

$$\mathscr{E}(\lambda) := e^{\lambda\mathscr{H}}. \qquad (74)$$

Then we have for all $\lambda, \mu \in \mathbb{R}_{\geq 0}$

$$
\begin{aligned}
&\mathscr{E}(\lambda)P(\mu;\underline{x}) \\
&= \mathscr{E}(\lambda)\mathscr{E}(\mu)P(0;\underline{x}) \\
&= \mathscr{E}(\lambda)g(\mu,\underline{x})P(0;\underline{T}(\mu,\underline{x})) \\
&= g(\lambda,\underline{x})g(\mu,\underline{T}(\lambda,\underline{x}))P(0;\underline{T}(\mu;\underline{T}(\lambda;\underline{x}))) \\
&\overset{(69)}{=} g(\lambda+\mu,\underline{x})P(0;\underline{T}(\lambda+\mu;\underline{x})) \\
&= \mathscr{E}(\lambda+\mu)P(0;\underline{x}) = P(\lambda+\mu;\underline{x}).
\end{aligned}
\qquad (75)
$$

This finally allows us to make the following statement:

---

* Note that this statement refers to a statement on formal parameters, for which addition and subtraction are well-defined operations.

Corollary 2. *Consider an N-species chemical reaction system with a single non-binary chemical reaction. Then the action of the* evolution semi-group $\mathscr{E}(t)$ *on the initial state's probability generating function $P(0;\underline{x})$ is completely described in terms of the action of the evolution semi-group at the level of probability generating functions, whence the PGF of $|\Psi(t)\rangle$ is given for all $t \in \mathbb{R}_{\geq 0}$ by*

$$P(t;\underline{x}) = \mathscr{E}(t)P(0;\underline{x}) = g(t;\underline{x})P(0;\underline{T}(t;\underline{x})). \quad (76)$$

6.2 Comments on elementary non-binary reactions

While it would in principle be just as feasible to compute either the moment- or the factorial moment generating functions rather than the probability generating functions for the non-binary reactions as presented in Table 3 via application of Theorem 7, here since we have available the probability generating functions, we may alternatively make use of the well-known relationship

$$\mathscr{M}(t;\lambda) = P(t;e^{\lambda}) \qquad (77)$$

as well as of the *Stirling transform* (cf. Appendix B)

$$\mathscr{F}(t;\nu) = \mathscr{M}(t;\ln(\nu+1)) = P(t;\nu+1). \quad (78)$$

For practical applications, it is often also of interest to consider the *cumulant generating function* $\mathscr{C}(t;\lambda)$, defined as

$$\mathscr{C}(t;\lambda) := \ln(\mathscr{M}(t;\lambda)) = \ln(P(t;e^{\lambda})). \quad (79)$$

We present in Figure 2 some concrete examples of discrete distributions for the non-binary elementary reactions of one species, and also examples for binary reactions as computed in Section 6.4ff. Reaction rates and time instances were chosen such as to allow for a qualitative comparison of the six elementary types of reaction systems. Note in particular that the probability distributions for systems consisting of an individual reaction $0A \xrightarrow{\gamma} 2A$ or

Table 3 Closed-form results for the time-dependent probability generating functions $P(t;\underline{x})$ for reaction systems of N species with a single non-binary elementary reaction; here, $S_1,\ldots,S_N$ denote the N different species, while $\underline{\Delta}_i$ ($i \in \{1,\ldots,N\}$) denotes the N-vector with coordinates $(\underline{\Delta}_i) = \delta_{i,j}$.

reaction	$\mathscr{H} = \underline{q}(\underline{x}) \cdot \partial_{\underline{x}} + v(\underline{x})$	$P(t;\underline{x}) = g(t;\underline{x})P(0;\underline{T}(t;\underline{x}))$	comments
$0 \xrightarrow{\alpha} S_i$	$\alpha(\hat{x}_i - 1)$	$Pois(\alpha t;x_i) \cdot P(0;\underline{x})$	$Pois(\mu;x) := e^{\mu(x-1)}$ (Poisson distribution, $0 \le \mu < \infty$)
$0 \xrightarrow{\alpha} S_i + S_j$	$\alpha(\hat{x}_i \hat{x}_j - 1)$	$\left(e^{\alpha t(x_i x_j - 1)}\right) \cdot P(0;\underline{x})$	
$S_i \xrightarrow{\alpha} \emptyset$	$\alpha(1 - \hat{x}_i)\,\frac{\partial}{\partial x_i}$	$P(0;\underline{x} + (-x_i + Bern(e^{-\alpha t};x_i))\,\underline{\Delta}_i)$	$Bern(\mu;x) := (1-\mu) + x\mu$ (Bernoulli distribution, $0 \le \mu \le 1$)
$S_i \xrightarrow{\alpha} S_j$ $(i \ne j)$	$\alpha(\hat{x}_j - \hat{x}_i)\,\frac{\partial}{\partial x_i}$	$P(0;\underline{x} + (-x_i + (x_j(1 - e^{-\alpha t}) + x_i e^{-\alpha t})\,\underline{\Delta}_i)$	
$S_i \xrightarrow{\alpha} 2S_i$	$\alpha(\hat{x}_i^2 - \hat{x}_i)\,\frac{\partial}{\partial x_i}$	$P(0;\underline{x} + (-x_i + Geom(e^{-\alpha t};x_i))\,\underline{\Delta}_i)$	$Geom(\mu;x) := \frac{x\mu}{1-x(1-\mu)}$ (Geometric distribution, $0 < \mu \le 1$)
$S_i \xrightarrow{\alpha} S_i + S_j$ $(i \ne j)$	$\alpha(\hat{x}_i \hat{x}_j - \hat{x}_i)\,\frac{\partial}{\partial x_i}$	$P(0;\underline{x} + (-x_i + x_i Pois(\alpha t;x_j)))\,\underline{\Delta}_i)$	
$S_i \xrightarrow{\alpha} S_j + S_k$ $(i \ne j \ne k)$	$\alpha(\hat{x}_j \hat{x}_k - \hat{x}_i)\,\frac{\partial}{\partial x_i}$	$P(0;\underline{x} + (-x_i + x_j x_k(1 - e^{-\alpha t}) + x_i e^{-\alpha t})\,\underline{\Delta}_i)$	

$2A \xrightarrow{\kappa} 0A$, respectively, can evidently only evolve taking non-zero values on the even integer particle numbers when initialized on a pure state $|\Psi(0)\rangle = |100\rangle$. Reaction systems consisting of more than one reaction will be treated in the next section, where we will provide an exact result on the probability generating functions for systems consisting of all four types of non-binary chemical reactions (cf. Theorem 9).

6.3 On compositionality of elementary non-binary reactions

Even though the results presented in Table 3 constitute a complete list of evolution formulae for *individual* non-binary chemical reactions, this does not yet provide a full answer to the question of which form the evolution of reaction systems composed of *multiple* non-binary chemical reactions would take. While the techniques of Theorem 7 remain applicable for such a generic non-binary reaction system, one would a priori have to compute the functions $g(\lambda;\underline{x})$ and $\underline{T}(\lambda;\underline{x})$ anew for each possible such reaction system. It would thus be desirable to find a technique that allows to simplify this computation via providing a form of *composition law* for solutions to chemical reaction systems assembled from elementary reactions. As a first step in this direction, we present a result for a reaction system composed of all four possible one-species non-binary elementary reactions:

Theorem 9. *Consider a* 1*-species reaction system with reactions*

$$\emptyset \xrightarrow{\beta} S, \quad \emptyset \xrightarrow{\gamma} 2S$$
$$S \xrightarrow{\tau} \emptyset, \quad S \xrightarrow{\alpha} 2S,$$
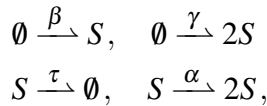
with parameters $\alpha, \beta, \gamma, \tau \in \mathbb{R}_{\geq 0}$. *The reactions describe birth* (B), *pair creation* (C), *death* (D) *and autocatalysis* (A) *transitions. Then the probability generating function* $P(t;x)$ *given* $P(0;x)$ *reads*

$$P(t;x) = g_{BDA}(t;x)g_{CDA}(t;x)P(0;T_{DA}(t;x)), \tag{80}$$

where the concrete formulae differ slightly for the various possible choices of parameters. Let us for convenience introduce the notations

$$\begin{aligned} T_D(t;x) &:= Bern(e^{-t\tau};x) \\ T_A(t;x) &:= Geom(e^{-t\alpha};x) \end{aligned} \tag{81}$$

and the following formulae for some standard probability generating functions (i.e. the alternative geometric and 2*-aerated Poisson distributions' PGFs):*

$$\begin{aligned} Geom_2(p;x) &:= \frac{p}{1-x(1-p)} \quad (0 < p \leq 1) \\ A_2Pois(p,x) &:= e^{p(x^2-1)} \quad (0 \leq p). \end{aligned} \tag{82}$$

Case $\alpha = \tau = 0$:

$$\begin{aligned} T_{DA}^{\alpha=0,\tau=0}(t;x) &= x \\ g_{BDA}^{\alpha=0,\tau=0}(t;x) &= Pois(\beta t;x) \\ g_{CDA}^{\alpha=0,\tau=0}(t;x) &= A_2Pois(\gamma t;x). \end{aligned} \tag{83}$$

Case $\alpha = \tau \neq 0$:

$$\begin{aligned} T_{DA}^{\alpha=\tau>0}(t;x) &= T_D(s(t);T_A(s(t);x)) \\ g_{BDA}^{\alpha=\tau>0}(t;x) &= \left(Geom_2\left(e^{-\alpha s(t)};x\right)\right)^{\beta/\alpha} \\ g_{CDA}^{\alpha=\tau>0}(t;x) &= \left(Geom_2\left(e^{-\alpha s(t)};x\right)\right)^{2\gamma/\alpha} \\ &\quad \cdot e^{\gamma t(x-1)^2 Geom_2(e^{-\alpha s(t)};x)} \end{aligned} \tag{84}$$

$$s(t) := \frac{1}{\alpha}\ln(1+\alpha t).$$

Case $\alpha \neq \tau$ and $\alpha > 0, \tau \geq 0$:

$$T_{DA}^{\alpha \neq \tau}(t;x) = T_D(s_D(t); T_A(s_A(t);x))$$

$$g_{BDA}(t;x) = \left(Geom_2\left(e^{-\alpha s_A(t)}; x \right) \right)^{\beta/\alpha}$$

$$g_{CDA}(t;x) = \left(Geom_2\left(e^{-\alpha s_A(t)}; x \right) \right)^{\gamma(\alpha+\tau)/\alpha^2}$$

$$\cdot e^{f_{CDA}(t;x) Geom_2(e^{-\alpha s_A(t)};x)}$$

$$f_{CDA}(t;x) = \frac{\gamma}{\alpha+\tau}\left((x-1)^2 \left(e^{t(\alpha-\tau)} - 1 \right) \right.$$

$$\left. + (x-1)(1+\tfrac{\tau}{\alpha})\left(1 - e^{t(\alpha-\tau)} \right) \right)$$

$$\tau s_D(t) := \ln\left(\frac{\alpha - \tau e^{(\tau-\alpha)t}}{\alpha - \tau} \right)$$

$$\alpha s_A(t) := \ln\left(\frac{\alpha e^{(\alpha-\tau)t} - \tau}{\alpha - \tau} \right).$$

$$(85)$$

Case $\alpha = 0$ and $\tau > 0$:

$$T_{DA}^{\alpha=0,\tau>0}(t;x) = T_D(t;x)$$

$$g_{BDA}^{\alpha=0,\tau>0}(t;x) = Pois(\tfrac{\beta}{\tau}\left(1 - e^{-t\tau} \right);x)$$

$$g_{CDA}^{\alpha=0,\tau>0}(t;x) = Pois(\tfrac{\gamma}{\tau}\left(1 - e^{-t\tau} \right)^2;x) \cdot$$

$$\cdot A_2 Pois(\tfrac{\gamma}{2\tau}\left(1 - e^{-2t\tau} \right);x).$$

$$(86)$$

Proof. Via applying Theorem 7, one obtains formulae for $P(t;x)$ in the various cases of interest that may then be brought into the forms as presented by an elementary, if somewhat tedious "factorization" of the formulae into compounds and convolutions of probability generating functions. $\square$

It might be worthwhile to consider the following auxiliary formula, which expresses the fact that taking a power of the $Geom_2(p;x)$ PGF with a positive

real exponent $\mu \in \mathbb{R}_{\geq 0}$ results in a well-posed PGF[*]:

$$(Geom_2(p;x))^\mu = \left(\frac{p}{1 - x(1-p)} \right)^\mu$$

$$= \delta_{\mu,0} + \delta_{\mu>0} p^\mu \sum_{n \geq 0} \binom{n+\mu-1}{n}(1-p)^n x^n.$$

$$(87)$$

For the permissible paramter range of the $Geom_2(p;x)$ ($0 < p \leq 1$), we thus observe that all coefficients of x^n for all $n \geq 0$ are non-negative real numbers, and evidently $Geom_2(p;1)^\mu = 1$.

Since admittedly the results of Theorem 9 are somewhat hard to interpret from the presentation in terms of generating functions, we exhibit in Figure 3 a set of illustrative examples of parameter choices and their respective effects on the first three moments and on the probability distributions. Moreover, in Figure 4 we provide ternary data plots for the effects of relative parameter choices for reaction systems of three of the four possible semi-linear reactions for illustration.

A particularly interesting effect concerns the example of the reaction system

$$1A \xrightarrow{\alpha=1/3} 2A, \quad 0A \xrightarrow{\gamma=1/3} 2A, \quad 1A \xrightarrow{\tau=1/3} 0A,$$

$$(88)$$
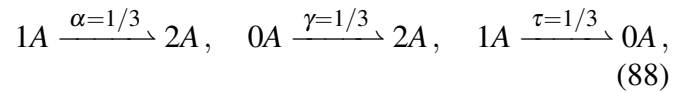
for which we depict the evolution of the first three cumulants in Figure 3a as well as the evolution of the probability distribution in Figure 3b. Notably, the mean (i.e. the first cumulant) remains fixed at the initial value (in this case $c_1(0) = 100$), and throughout the evolution the probability distribution effectively undergoes a progressive broadening of the distribution around this mean value.

---

[*] For the case $0 < \mu < 1$, we find

$$\binom{k+\mu-1}{k} = \frac{\Gamma(\mu+k)}{k!\Gamma(\mu)} = \frac{1}{k!}(\mu)^{(k)},$$

with $(x)^{(k)}$ the *rising factorial*, i.e. $(x)^{(0)} = 1$ and $x^{(k)} = x(x+1)\ldots(x+k-1)$ for $k > 0$.

In Figures 3c and 3d, we depict cumulant and probability distribution evolutions for a variant of a birth-death system, namely a system with also a pair creation reaction,

$$0A \xrightarrow{\beta=1/5} 1A, \quad 0A \xrightarrow{\gamma=3/5} 2A, \quad 1A \xrightarrow{\tau=1/5} 0A.$$
(89)
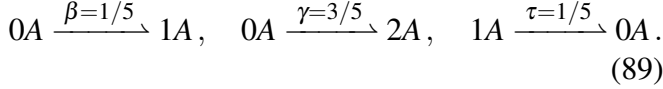
It is instructive to consider the exact formula for the probability generating function of this system (cf. (86)), which for an initial state $|\Psi(0)\rangle = |M\rangle$ of precisely M particles (i.e. for $P(0;x) = x^M$) reads

$$P(t;x) = e^{\frac{\gamma}{2\tau}\left(1-e^{-2t\tau}\right)(x^2-1)}$$
$$\cdot e^{\left[\frac{\gamma}{\tau}\left(1-e^{-t\tau}\right)^2 + \frac{\beta}{\tau}\left(1-e^{-t\tau}\right)\right](x-1)} \quad (90)$$
$$\cdot \left(Bern(e^{-t\tau};x)\right)^M.$$

Taking the limit $t \to \infty$ and noting that

$$\lim_{t\to\infty} Bern(e^{-t\tau};x) = 1,$$

we find the limit distribution

$$\lim_{t\to\infty} P(t;x) = e^{\frac{\gamma}{2\tau}(x^2-1)+\frac{\beta+\gamma}{\tau}(x-1)}. \quad (91)$$

We present in Figure 4 a number of ternary data plots the evolution of first and second cumulants for this reaction system for illustration.

As a final example, we depict in Figures 3e and 3f the evolution of a reaction system with all four types of non-binary one-species reactions. Depending on the choices of the four reaction rates involved, we find that the system evolves in a fashion either similar to the reaction system (88) or to the reaction system (89) (where in the example depicted in Figures 3e and 3f parameters are chosen in the latter regime).

To conclude our treatment of non-binary chemical reaction systems, it is worthwhile noting that there exists previous work in the mathematical chemistry literature [11] on closed-form exact solutions to probability generating functions for multi-species non-binary reactions except for the autocatalytic reaction types. Since our formulae cover an arbitrary single-species non-binary reaction system, it appears worthwhile to consider the generalization of our Theorem 9 to the multi-species cases. We plan to present such results in future work.

6.4 Closed-form solutions for probability generating functions of single-species binary reactions

In the 1960's, McQuarrie [2, 12] suggested to solve the problem of finding the probability generating functions for single-species binary chemical reaction systems and for a deterministic initial state $|\Psi(0)\rangle = |M\rangle$, whence for $P(0;x) = x^M$, via a separation of variables Ansatz of the form

$$P(t;x) = \sum_{n\geq 0} A_{M;n} T_n(t) X_n(x). \quad (92)$$

Here, the functions $X_n(x)$ are *polynomials*, with $degree(X_n) = n$. The time-dependent functions $T_n(t)$ read

$$T_n(t) = e^{\lambda_n t}, \quad (93)$$

while the constant coefficients $A_{M;n} \in \mathbb{R}$ are determined by the initial condition $P(0;x) = x^M$,

$$x^M = \sum_{n\geq 0} A_{M;n} X_n(x) = \sum_{n=0}^{M} A_{M;n} X_n(x). \quad (94)$$

In order for McQuarrie's Ansatz to be consistent, it is thus strictly necessary that the polynomials $X_n(x)$ form a *complete system of eigenfunctions* of the Hamiltonian $\mathscr{H} \equiv \mathscr{H}(\hat{x}, \partial_x)$ in the Bargmann-Fock basis (compare (55) for the one-species case upon invoking the one-to-one correspondence described in (26)):

$$\mathscr{H} X_n(x) = \lambda_n X_n(x). \quad (95)$$

In other words, it must be the case that precisely one degree n polynomial $X_n(x)$ is available for each

non-negative integer value n.

McQuarrie suggests the following approach: the functions $X_n(x)$ should be found as certain *orthogonal polynomials of the Jacobi type*. Let us briefly review some of the necessary background material on this mathematical structure.

6.5 Classical Jacobi polynomials and orthogonality

The Jacobi-type second-order differential operator (cf. the standard reference book [33], chapter 18, p. 445, Table 18.8.1)

$$D_{Jac}^{\alpha,\beta} := (1-z^2)\frac{\partial^2}{\partial z^2} + q^{(\alpha,\beta)}(z)\frac{\partial}{\partial z}$$
$$q^{(\alpha,\beta)}(z) = (\beta - \alpha - (\alpha+\beta+2)z) \tag{96}$$

with *real parameters*

$$\alpha, \beta > -1 \tag{97}$$

is known to possess a *system of orthogonal polynomials* $\{P_n^{(\alpha,\beta)}(z)\}_{n\geq 0}$, the so-called *classical Jacobi polynomials*, as its complete basis of eigenfunctions. More explicitly (cf. e.g. [33], Equation 18.5.8),

$$P_n^{(\alpha,\beta)}(z)$$
$$= 2^{-n}\sum_{\ell=0}^{n}\binom{n+\alpha}{\ell}\binom{n+\beta}{n-\ell}(z-1)^{n-\ell}(z+1)^\ell, \tag{98}$$

with eigenequation

$$D_{Jac}^{\alpha,\beta}P_n^{(\alpha,\beta)}(z) = -n(n+\alpha+\beta+1)P_n^{(\alpha,\beta)}(z). \tag{99}$$

The orthogonality property is found by defining for each admissible choice of parameters $\alpha, \beta > -1$ an *inner product* $\Phi_{\alpha,\beta}$ on the space of polynomials,

$$\Phi_{\alpha,\beta}(p(z),q(z)) := \int_{-1}^{+1} dz\, w_{\alpha,\beta}(z)p(z)q(z)$$
$$w_{\alpha,\beta}(z) := (1-z)^\alpha(1+z)^\beta. \tag{100}$$

It is one of the classical results of the theory of orthogonal polynomials that for all $n \geq 0$

$$P_n^{(\alpha,\beta)}(z) \in L_{\alpha,\beta}^2([-1,1]). \tag{101}$$

Here, $L_{\alpha,\beta}^2([-1,1])$ denotes the space of functions on the interval $[-1,1]$ which are *square-integrable* w.r.t. the integral against the weight function $w_{\alpha,\beta}(z)$. Moreover, orthogonality manifests itself as

$$\Phi_{\alpha,\beta}\left(P_m^{(\alpha,\beta)}(z), P_n^{(\alpha,\beta)}(z)\right) = \delta_{m,n}\phi_{\alpha,\beta}(n), \tag{102}$$

with $\phi_{\alpha,\beta}(n) \in \mathbb{R}_{>0}$ some non-zero real numbers.

6.6 The case of binary reactions and a technical problem

As we shall see momentarily, the type of differential operators we are interested in when studying chemical reaction systems, while superficially of the structure of the Jacobi-type differential operators, is precisely *not* fulfilling the parameter constraint (97). The problem presents itself as follows:

Lemma 2. *Consider a chemical reaction system of one species of particles and the reactions*

$$A \xrightarrow{r_d} \emptyset, \quad 2A \xrightarrow{r_k} \emptyset, \quad 2A \xrightarrow{r_\ell} A, \tag{103}$$
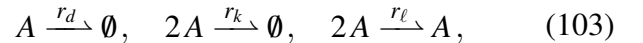
where $r_k + r_\ell > 0$ (i.e. the system involves at least one binary reaction), and where we assume for simplicity and without loss of generality [*]

$$r_k + r_\ell = 1.$$

The corresponding Hamiltonian reads in the Bargmann-Fock basis

$$\mathscr{H} = r_d(1-\hat{x})\frac{\partial}{\partial x}$$
$$r_k(1-\hat{x}^2)\left(\frac{\partial}{\partial x}\right)^2 + r_\ell(\hat{x}-\hat{x}^2)\left(\frac{\partial}{\partial x}\right)^2. \tag{104}$$

[*] Note that generic values of r_k and r_ℓ may be achieved via appropriately rescaling t. See Proposition 6 for the precise details.

21

Then one can bring the Jacobi-type differential operator $D_{Jac}^{\alpha,\beta}$ as presented in (96) via the affine transformation

$$x(z) = \frac{r_\ell}{2} + \left(1 - \frac{r_\ell}{2}\right) z \qquad (105)$$

into the form of $\mathcal{H}$, which entails that

$$(\mathcal{H} + n(n + r_d - 1)) P_n^{(-1;r_d-1)} \left(\frac{x - \frac{r_\ell}{2}}{1 - \frac{r_\ell}{2}}\right) = 0.$$
$$(106)$$

Unfortunately, we thus encounter for *all* of the interesting reaction rate choices yielding binary chemical reaction systems the "illegal" parameter ranges, since we are interested specifically in the cases

(i) $\alpha = -1$, $\beta = r_d - 1 > -1$
(ii) $\alpha = -1$, $\beta = -1$ (i.e. $r_d = 0$). $\qquad (107)$

We encounter the following technical problems for these two types of parameter choices:

(*i*) *Case $\alpha = -1, \beta > -1$:* With respect to the inner product $\Phi_{-1,\beta}$ as defined in (100) (with the variable named x from hereon), we find that the degree zero polynomial

$$P_0^{(-1,\beta)}(x) = 1$$

is *not* square-integrable,

$$P_0^{(-1,\beta)}(x) \notin L_{-1,\beta}^2([-1,1]).$$

Therefore, the set $\{P_n^{(-1,\beta)}(x)\}_{n\geq 0}$ fails to constitute a complete set of eigenfunctions of the operator $\mathcal{H}$.

(*ii*) *Case $\alpha = -1, \beta = -1$:* In addition to the failure of $P_0^{(-1,-1)}(x)$ to be square-integrable w.r.t. the inner product $\Phi_{-1,-1}$, we moreover find from the explicit formula (98) for the Jacobi polynomials that

$$P_1^{(-1,-1)}(x) = 0. \qquad (108)$$

In other words, in this case we lack the degree zero *and* the degree one eigenfunctions, and thus again do not obtain a complete system of eigenfunctions.

As it stands, one would thus have to conclude that McQuarrie's approach to binary chemical reaction systems is ill-conceived.

A resolution of the problem of incompleteness of the eigenbasis has been provided by Kwon and Littlejohn [13–16] (see also their survey article [34]). They observe first that any constant function and any linear function are solutions to the Jacobi type differential equation, and that moreover the family $\{P_n^{(\alpha,\beta)}(x)\}_{n\geq 2}$ even at the problematic parameter values $\alpha = -1$ and $\beta \geq -1$ remains an orthogonal family of eigenfunctions. Their key insight is then to consider an *alternative notion of orthogonality*, so-called *Sobolev-orthogonality*, which admits to construct a family $\{\tilde{P}_n^{(\alpha,\beta)}\}_{n\geq 0}$ that is a complete system of eigenfunctions and orthogonal w.r.t. a Sobolev-type inner product.

We quote from [16] the relevant material (cf. Propositions 4.2 and 4.3 of loc cit):

Proposition 3. *Let the* Sobolev inner products $\Phi_{A,B}^{(-1,-1)}$ *and* $\Phi_C^{(-1,\beta)}$ *(for $\beta > -1$) be defined as follows (for $p, q \in \mathbb{R}[x]$ some polynomials and $\beta \in \mathbb{R}_{\geq -1}$)*

$$\Phi_{A,B}^{(-1,-1)}(p,q) := Ap(1)q(1) + Bp(-1)q(-1)$$
$$+ \int_{-1}^{+1} dx\, p'(x)q'(x)$$

$$\Phi_C^{(-1,\beta)}(p,q) := Cp(1)q(1) \qquad (\beta > -1)$$
$$+ \int_{-1}^{+1} dx\,(x+1)^{\beta+1} p'(x)q'(x).$$
$$(109)$$

Here, $p'(x)$ and $q'(x)$ denote the first derivatives of the polynomials, and $A, B, C \in \mathbb{R}$ are some parameters, which have to satisfy the following conditions:

for $\Phi_{A,B}^{(-1,-1)}$, A and B must verify

$$A + B > 0, \quad A(\gamma+1)^2 + B(\gamma-1)^2 + 2 \neq 0$$
$$A(\gamma+1) + B(\gamma-1) = 0$$
$$\gamma := (B-A)/(A+B),$$
$$(110)$$

while for $\Phi_C^{(-1,\beta)}$ the parameter C must verify

$$C > 0. \qquad (111)$$

Then the Sobolev-Jacobi orthogonal polynomials $\tilde{P}_n^{(\alpha,\beta)}(x)$ *are defined for* $\alpha = -1, \beta = -1$ *as*

$$\tilde{P}_0^{(-1,-1)}(x) := 1$$
$$\tilde{P}_1^{(-1,-1)}(x) := x + \gamma$$
$$\tilde{P}_{n\geq 2}^{(-1,-1)}(x) := \binom{2n-2}{n}^{-1}. \qquad (112)$$
$$\cdot \sum_{k=1}^{n-1} \binom{n-1}{k}\binom{n-1}{n-k}(x-1)^{n-k}(x+1)^k,$$

where $\gamma = (B-A)/(A+B)$, *while for the parameters* $\alpha = -1, \beta > -1$ *one defines*

$$\tilde{P}_0^{(-1,\beta)}(x) := 1$$
$$\tilde{P}_1^{(-1,\beta)}(x) := x - 1$$
$$\tilde{P}_{n\geq 2}^{(-1,\beta)}(x) := \binom{2n+\beta-1}{n}^{-1} \qquad (113)$$
$$\cdot \sum_{k=0}^{n} \binom{n-1}{k}\binom{n+\beta}{n-k}(x-1)^{n-k}(x+1)^k.$$

Then the $\tilde{P}_n^{(\alpha,\beta)}(x)$ form a complete orthogonal system of polynomial eigenfunctions of the Jacobi differential operator at parameters for the aforementioned parameter ranges.

The readers may have noticed that indeed the Sobolev-Jacobi polynomials $\tilde{P}_n^{(\alpha,\beta)}(x)$ as defined above coincide for $n \geq 2$ with the monic versions of the classical Jacobi polynomials (i.e. a different normalization choice where in each polynomial of

degree n the coefficient of x^n is normalized to be 1).

In other words, while the strategy of using systems of *classical* orthogonal polynomial eigenfunctions of the Jacobi-operator as originally suggested by McQuarrie fails, one can "repair" the Ansatz by working with the Sobolev-orthogonal polynomials of Kwon and Littlejohn instead, which are orthogonal over the *weighted Hilbert spaces* defined via the inner products $\Phi_{A,B}^{(-1,-1)}$ and $\Phi_C^{(-1,\beta)}$, respectively. This application of their framework to the study of binary chemical reactions appears to be new.

6.7 Intermezzo: how to fix the free parameters in the inner products

In a remarkable series of papers [35–40], N. Gurappa and P.K. Panigrahi developed a novel method to construct polynomial eigenfunctions of second and higher order differential operators. If such a differential operator is given* as a polynomial in x and $\frac{d}{dx}$, they observe that one can always uniquely express the equation for the eigenfunctions $y(x)$ of the operator in the form

$$\left(F(D) + P(x, \tfrac{d}{dx})\right) y(x) = 0. \qquad (114)$$

Here, $D = x\frac{d}{dx}$ is the *Euler operator* (which the readers may recognize as the number operator in the Bargmann-Fock basis). For instance, given a generic second order differential operator (see [35–40] for more general cases), the term $F(D)$ reads

$$F(D) = a_2 D^2 + a_1 D + a_0, \qquad (115)$$

with $a_i \in \mathbb{R}$ some constant coefficients. Crucially, $F(D)$ is a *diagonal operator* on the space of monomials, due to

$$Dx^n = nx^n. \qquad (116)$$

---

* In this subsection, in order to be compatible in notation with the material presented in as a polynomial in [35–40], we will temporarily employ the notational conventions $x \equiv \hat{x}$ and $\frac{d}{dx} \equiv \partial_x$.

The operator $P \equiv P(x, \frac{d}{dx})$ on the other hand contains all terms not expressible via polynomials in D, whence P contains all "off-diagonal" contributions to the differential operator. The authors of [35–40] then construct the polynomial eigenfunctions via the Ansatz

$$F(D)x^\lambda = 0 \ \Rightarrow \ y_\lambda(x) = c_\lambda \hat{G}_\lambda x^\lambda$$

$$\hat{G}_\lambda := \sum_{m \geq 0} (-1)^m \left[\frac{1}{F(D)} P(x, \tfrac{d}{dx}) \right]^m. \quad (117)$$

Here, $\lambda \in \mathbb{Z}$ is an integer parameter and $c_\lambda \in \mathbb{R}$ a normalization constant. The function $y_\lambda(x)$ must in addition satisfy the consistency condition

$$y_\lambda(x) - c_\lambda x^\lambda = 0. \quad (118)$$

Specializing the approach of [35–40] to the case of the Jacobi-type differential operator $D_{Jac}^{\alpha,\beta}$ for parameters $\alpha = -1$ and $\beta \geq -1$, we first compute the decomposition of the operator into "diagonal" and "off-diagonal" parts. After some elementary manipulations, we find the following results:

Proposition 4. *For $\beta \geq -1$, the* Sobolev-Jacobi *polynomials $\tilde{P}_n^{(-1,\beta)}(x)$ have the presentation*

$$\tilde{P}_n^{(-1,\beta)}(x) = \hat{G}_{\beta,n} x^n$$

$$\hat{G}_{\beta,n} = \sum_{m \geq 0} \left[\frac{1}{(D-n)(D+n+\beta)} \cdot \right. \quad (119)$$

$$\left. \cdot \left(\frac{d^2}{dx^2} + (\beta+1)\frac{d}{dx} \right) \right]^m.$$

More explicitly, since by virtue of this definition

$$\tilde{P}_1^{(-1,\beta)}(x) = x - 1 + \delta_{\beta,-1}, \quad (120)$$

the polynomials $\tilde{P}_n^{(-1,\beta)}(x)$ are identified as Sobolev-Jacobi polynomials w.r.t. the inner products $\Phi_{A,A}^{(-1,-1)}$ and $\Phi_C^{(-1,\beta)}$, respectively. Moreover, for the case $\beta = -1$ one may simplify the expression for $\tilde{P}_n^{(-1,-1)}(x)$ into the form

$$\tilde{P}_n^{(-1,-1)}(x) = e^{-\frac{1}{2}\frac{1}{D+n-1}\frac{d^2}{dx^2}} x^n. \quad (121)$$

Proof. See Appendix D.1. □

It is via these novel techniques that one may hope to tackle the cases of multi-species binary chemical reaction systems, which we plan to elaborate upon in future work. For the present paper, we merely note that this approach suggests a particular choice for the parameters A and B of the inner product $\Phi_{A,B}^{(-1,-1)}$, which we will adopt from hereon in all of the ensuing computations: the "canonical" choice of parameters is according to Proposition 4 to set $A = B$, which yields indeed

$$\gamma = (B-A)/(A+B) = 0.$$

For later convenience, we further choose to fix the remaining free parameters to $A = 1$ and $C = 1$.

6.8 Solving binary chemical reaction systems via Sobolev-orthogonal polynomials

It remains to provide a formula for computing the linear combination coefficients $A_{M;n}$ in the separation of variables Ansatz (92) for the time-dependent probability generating function $P(t;x)$. Let us then assume as before that we are given the PGF of the initial state of the form

$$P(0;x) = x^M,$$

whence the PGF of a pure state with $M \geq 0$ particles. We will follow the ideas of McQuarrie, adapted to the generalized Jacobi-polynomial case. In order to clearly distinguish the coefficients $A_{M,n}$ from those in the present Ansatz, we choose to introduce the notations

$$A_\beta^{M,n} \text{ and } B_\beta^{M,n}$$

for the coefficients in the ensuing formulae. Adapting the Ansatz presented in Lemma 2 according to

$$P(t;x) =$$

$$\sum_{n=0}^{M} A_{(r_d-1)}^{M,n} e^{-n(n+r_d-1)t} \tilde{P}_n^{(-1,r_d-1)} \left(\frac{x - \frac{r_\ell}{2}}{1 - \frac{r_\ell}{2}} \right), \quad (122)$$

it remains to solve the following equation:

$$x^M = \sum_{n=0}^{M} A^{M,n}_{(r_d-1)} \tilde{P}_n^{(-1,r_d-1)} \left(\frac{x - \frac{r_\ell}{2}}{1 - \frac{r_\ell}{2}} \right). \qquad (123)$$

For the purpose of simplicity of the presentation, it is more economical to focus instead on the equation

$$x^M = \sum_{n=0}^{M} B^{M,n}_\beta \tilde{P}_n^{(-1,\beta)}(x), \qquad (124)$$

which then allows to extract the coefficients $A^{M,n}_{(r_d-1)}$ via performing a suitable change of variables in (123) (see Appendix D.2 for further details):

$$A^{M,n}_{(r_d-1)} =$$
$$\sum_{P=n}^{M} \binom{M}{P} (\tfrac{r_\ell}{2})^{M-P} (1 - \tfrac{r_\ell}{2})^P B^{P,n}_{(r_d-1)}. \qquad (125)$$

According to Proposition 3 and relying in addition on the results if the previous section, for both of the cases $\beta = -1$ and $\beta > -1$ we have appropriate inner products $\Phi_{1,1}^{(-1,-1)}$ and $\Phi_1^{(-1,\beta)}$ available with respect to which the Sobolev-Jacobi polynomials are orthogonal. Therefore, one can compute the coefficients $B^{M,n}_\beta$ as follows:

- *Case $\beta = -1$:*

$$B^{M,n}_{-1} = \frac{\Phi_{1,1}^{(-1,-1)}\left(x^M, \tilde{P}_n^{(-1,-1)}(x)\right)}{\Phi_{1,1}^{(-1,-1)}\left(\tilde{P}_n^{(-1,-1)}(x), \tilde{P}_n^{(-1,-1)}(x)\right)}. \qquad (126)$$

- *Case $\beta > -1$:*

$$B^{M,n}_\beta = \frac{\Phi_1^{(-1,\beta)}\left(x^M, \tilde{P}_n^{(-1,\beta)}(x)\right)}{\Phi_1^{(-1,\beta)}\left(\tilde{P}_n^{(-1,\beta)}(x), \tilde{P}_n^{(-1,\beta)}(x)\right)}. \qquad (127)$$

We may yet again rely on the results of Kwon and Littlejohn for the values of the denominators

(cf. Propositions 4.2 and 4.3 in [16]):

$$\Phi_{1,1}^{(-1,-1)}\left(\tilde{P}_n^{(-1,-1)}(x), \tilde{P}_n^{(-1,-1)}(x)\right)$$
$$= \begin{cases} 2 & \text{if } n = 0 \\ 4 & \text{if } n = 1 \\ n^2 K_{n-1} & \text{else.} \end{cases} \qquad (128)$$
$$K_{n-1} = \frac{2^{2n-1}((n-1)!)^4}{(2n-2)!(2n-1)!} \quad (n \geq 2)$$

$$\Phi_1^{(-1,\beta)}\left(\tilde{P}_n^{(-1,\beta)}(x), \tilde{P}_n^{(-1,\beta)}(x)\right)$$
$$= \begin{cases} 1 & \text{if } n = 0 \\ n^2 \kappa_{n-1}(0, \beta+1) & \text{else.} \end{cases}$$
$$\kappa_{n-1}(0, \beta+1) = \frac{2^{2n+\beta}((n-1)!(n+\beta+1)!)^2}{(2n+\beta-1)!(2n+\beta)!}$$
$$(n \geq 1). \qquad (129)$$

However, the computation of the numerators of the formulae for $B^{M,n}_\beta$ appears to not have been performed in the literature before, and thus necessitates a considerable amount of additional work. Referring to Appendix D.2 for some of the technical details of this derivation, suffice it to report here the final results for the coefficients $B^{M,n}_\beta$:

Proposition 5. *For the case $\beta = -1$, the coefficients $B^{M,n}_\beta$ read:*

$$B^{M,n}_{-1} = \begin{cases} \binom{2n}{n} \frac{m!\left(\frac{M+n}{2}\right)!}{(M+n)!\left(\frac{M-n}{2}\right)!}, & \text{if } n \leq M \wedge M+n \in 2\mathbb{Z} \\ 0, & \text{else.} \end{cases} \qquad (130)$$

For the case $\beta > -1$, we find

$$B^{M,n}_\beta = \begin{cases} \delta_{n,0} & \text{if } M = 0 \\ 1 & \text{if } n = 0 \\ b^{M,n}_\beta & \text{if } M, n \geq 1 \wedge n \leq M \\ 0 & \text{else} \end{cases}, \qquad (131)$$

where

$$b_\beta^{M,n} = \frac{M(2n+\beta)!}{2^{n-1}n!(n+\beta+1)!} \cdot$$
$$\cdot \sum_{k=0}^{n-1} \left[\binom{n-1}{k} (-1)^{n-1-k} \right.$$
$$\left. {}_2F_1(1-M, n-k; \beta+n+2; 2) \right].$$

(132)

Proof. See Appendix D.2. □

While we thus do not find a closed-form solution for the coefficients $B_\beta^{M,n}$ for the case $\beta > -1$, the result may be easily and efficiently implemented with the help of a modern computer algebra software such as e.g. MATHEMATICA, MAPLE or SAGE.

In summary, we find the following analytic solutions to binary chemical reaction systems of the type related to Sobolev-Jacobi polynomials:

Proposition 6. *Consider a chemical reaction system with Hamiltonian (for arbitrary $r_d, r_k, r_\ell \in \mathbb{R}_{\geq 0}$)*

$$\mathscr{H} = r_d (1-\hat{x}) \frac{\partial}{\partial x}$$
$$+ r_k (1-\hat{x}^2) \left(\frac{\partial}{\partial x}\right)^2 + r_\ell (\hat{x}-\hat{x}^2) \left(\frac{\partial}{\partial x}\right)^2$$

(133)

Let $P_M(0;x) := x^M$ denote the PGF of an input state that contains exactly $M \geq 0$ particles. Then the analytic solution to the PGF $P_M(t;x)$ for $t \geq 0$ reads for the case $r_k = r_\ell = 0$

$$P_M(t;x) = Bern(e^{-r_d t}; x),$$

(134)

and for the case $r_k + r_\ell > 0$

$$P_M(t;x) = \sum_{n=0}^{M} \left[A_{(\bar{r}_d-1)}^{M,n} e^{-n(n+\bar{r}_d-1)\sigma t} \cdot \right.$$
$$\left. \cdot \tilde{P}_n^{(-1,\bar{r}_d-1)} \left(\frac{x-\frac{\bar{r}_\ell}{2}}{1-\frac{\bar{r}_\ell}{2}} \right) \right],$$

(135)

where

$$\sigma := r_k + r_\ell, \quad \bar{r}_d = \frac{r_d}{\sigma}, \quad \bar{r}_\ell = \frac{r_\ell}{\sigma}$$
$$A_{(\bar{r}_d-1)}^{M,n} = \sum_{P=n}^{M} \binom{M}{P} (\frac{\bar{r}_\ell}{2})^{M-P} (1-\frac{\bar{r}_\ell}{2})^P B_{(\bar{r}_d-1)}^{P,n}.$$

(136)

Proof. For the case $r_k = r_\ell = 0$, we simply recover our previous result for the pure decay reaction. For the case $r_k + r_\ell > 0$, one may bring $\mathscr{H}$ into the form compatible with the assumptions of Lemma 2 via

$$\mathscr{H} = \sigma \tilde{\mathscr{H}},$$

(137)

where $\sigma = r_k + r_\ell$. It then remains to apply the results presented above in order to derive the formula for $P_M(t;x)$. □

We present in Figure 2 examples of the time-evolution of the resulting probability distributions for the two elementary types of binary chemical reactions for illustration.

Conclusion and outlook

Our goal in this work consisted in developing a framework within which the combinatorial properties and stochastic dynamics of chemical reaction networks may be discovered and analyzed. To this end, it proved fruitful to consider first a slightly more general theoretical setting, namely that of continuous-time Markov chains (CTMCs) [17] over discrete state spaces, and to construct a suitable *stochastic mechanics* framework for such theories (Section 2). It is interesting to note that while some concepts well-known from the study of quantum mechanical theories might provide a certain level of intuition in understanding the stochastic mechanics setup, the analogy is in fact not a particularly close one (see Appendix A). We then specialized our framework to the main focus of interest in this paper, namely to chemical reaction systems (Sections 3 and 4). Our approach highlights a rather fruitful synergy between the ideas of Dellbrück [1],

Doi [3] and the techniques based on generating functions (for probability distributions and their various types of moments) known from the combinatorics literature. We provide a concise derivation of three different types of *evolution equations* for analyzing chemical reaction systems, namely for the time-dependent states, for the exponential moment generating functions and for the factorial moment generating functions (Tables 1 and 2). On the level of general results, our framework then permits to derive a precise statement on which reaction systems possess the special dynamical feature of first-order moment closure (Section 5).

The second main contribution of our paper consists in a set of exact, closed-form results for the time-evolution of probability distributions of chemical reaction systems (Section 6). For chemical reactions of the non-binary type, we identify the problem of finding a solution to the evolution equations of the chemical reaction system as a special case of a so-called semi-linear normal-ordering problem (Section 6.3). Combining a powerful result from the combinatorics literature [7, 10] (cf. Theorem 7) with some elementary notions on constructions of convolutions and compound probability distributions, we are able to derive not only exact solutions for individual elementary non-binary chemical reactions (Table 3), but also for generic single-species non-binary reaction systems (Section 6.3). While there exists a large body of work in the standard chemistry literature (see [2] and in particular [11]) on such non-binary reactions, our approach itself and some of our exact results appear to be new. It would be interesting to also consider multi-species non-binary reaction systems (such as those already treated in [11] for reaction systems not involving autocatalytic reactions) in our novel framework, but in this paper we focussed on the single-species systems for brevity. Some examples for the dynamics of non-binary reaction systems are illustrated in Figures 3 and 4.

Finally, it has been an open problem in the

literature to analytically describe the evolution of chemical reaction systems involving binary chemical reactions. We highlight in Sections 6.4ff how the traditionally employed Ansatz due to McQuarrie [12] is ill-conceived due to an inconsistency in his choice of basis of eigenfunctions, and how the Ansatz may be repaired via modifying the relevant eigenbasis to so-called *Sobolev-Jacobi orthogonal polynomials*, which were originally introduced by Kwon and Littlejohn in the 1990s [13–16, 34]. We present in Section 6.8 an analytic solution to the time-evolution of the probability distributions for binary chemical reaction systems of one species (Proposition 6), which is illustrated for individual binary reactions in Figures 2e and 2f. We also present an interesting alternative definition of the Sobolev-Jacobi polynomials in terms of a remarkable framework due to Gurappa and Panigrahi [35–40] (Section 6.7), which provides intriguing perspectives in terms of extending our analytical results from single- to multi-species binary chemical reaction systems.

Two main avenues of future research suggest themselves as a result of our work: firstly, as already realized by one of the authors in a long-standing research project [4–6], chemical reactions may alternatively be considered as a particular class of stochastic transition systems, so-called discrete transition systems. More precisely, the pure states of a chemical system are fully characterized in terms of the respective numbers of particles of each of the chemical species in the pure state. In contrast, it is possible to formulate a full stochastic mechanics framework (based on the concept of the so-called *rule algebras* [4, 5]) for more general types of stochastic transition systems, whose state spaces consist of certain isomorphism classes of graphs, and whose transitions consist of *graph rewriting rules*. Without going into much further detail on these systems here, it is worthwhile to mention that the results presented in this paper on the three types of evolution equations for the

moments and factorial moments of generic multi-species chemical reaction systems (Table 2) may not only be generalized to the setting of generic transition systems, but in some cases conversely even arise as the evolution equations for certain observables in the generic systems. These results will be presented in [6].

Finally, while we are well aware that of course the examples presented do not reflect the complexity of reaction systems typically encountered in applications, we hope that our techniques may provide a starting point for the development of novel approximation techniques and algorithms for computing the time-dependent states of realistically complex reaction systems. We believe that it is of the utmost importance in this endeavor to take into account the combinatorics of the infinitesimal generators of chemical reaction systems, such as highlighted for example in the semi-linear normal-ordering techniques, and that it will prove paramount to understand in further detail why there appears to exist a clear structural divide between non-binary and binary reactions in terms of their evolution semi-groups. Referring to classification schemes for ODEs of second order such as in the work of Zhang [41], already the appearance of Sobolev-Jacobi orthogonal polynomials as eigenbasis appears to be a new development, and it would be rather interesting to understand in which way these polynomials interact with the infinitesimal generators of generic reaction systems in order to potentially provide a general closed-form solution for at least the single-species cases.

Acknowledgements

References

1 Max Delbrück. Statistical fluctuations in autocatalytic reactions. *The Journal of Chemical Physics*, 8(1):120–124, 1940.

2 Donald A McQuarrie. Stochastic approach to chemical kinetics. *Journal of applied probability*, 4(3):413–478, 1967.

3 Masao Doi. Second quantization representation for classical many-particle system. *Journal of Physics A: Mathematical and General*, 9(9):1465, 1976.

4 Nicolas Behr, Vincent Danos, and Ilias Garnier. Stochastic mechanics of graph rewriting. *Proceedings of the 31st Annual ACM-IEEE Symposium on Logic in Computer Science (LICS 2016)*, pages 46–55, 2016.

5 Nicolas Behr, Vincent Danos, Ilias Garnier, and Tobias Heindel. The algebras of graph rewriting. *arXiv:1612.06240*, 2016.

6 Nicolas Behr, Vincent Danos, and Ilias Garnier. Combinatorial conversion and disassociator dynamics for stochastic rewriting systems (in preparation).

7 P Blasiak, A Horzela, KA Penson, GHE Duchamp, and AI Solomon. Boson normal ordering via substitutions and Sheffer-Type Polynomials. *Physics Letters A*, 338(2):108–116, 2005.

8 Pawel Blasiak. Combinatorics of boson normal ordering and some applications. *arXiv preprint quant-ph/0507206*, 2005.

9 Pawel Blasiak, Andrzej Horzela, Karol A Penson, Allan I Solomon, and Gerard HE Duchamp. Combinatorics and boson normal ordering: A gentle introduction. *American Journal of Physics*, 75(7):639–646, 2007.

10 Pawel Blasiak and Philippe Flajolet. Combinatorial models of creation-annihilation. *Séminaire Lotharingien de Combinatoire*, 65(B65c):1–78, 2011.

11 Tobias Jahnke and Wilhelm Huisinga. Solving the chemical master equation for monomolecular reaction systems analytically. *Journal of mathematical biology*, 54(1):1–26, 2007.

12 Donald A McQuarrie, CJ Jachimowski, and ME Russell. Kinetics of small systems. ii. *The Journal of Chemical Physics*, 40(10):2914–2921, 1964.

13 KH Kwon, LL Littlejohn, and BH Yoo. Character-

izations of orthogonal polynomials satisfying differential equations. *SIAM Journal on Mathematical Analysis*, 25(3):976–990, 1994.

14 Kil H Kwon, LL Littlejohn, and BH Yoo. New characterizations of classical orthogonal polynomials. *Indagationes Mathematicae*, 7(2):199–213, 1996.

15 Kil H Kwon and Lance L Littlejohn. Classification of classical orthogonal polynomials. *J. Korean Math. Soc*, 34(4):973–1008, 1997.

16 Kil H Kwon and LL Littlejohn. Sobolev orthogonal polynomials and second-order differential equations. *The Rocky Mountain journal of mathematics*, pages 547–594, 1998.

17 James R. Norris. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998.

18 John C Baez and Jacob Biamonte. Quantum techniques for stochastic mechanic. *arXiv:1209.3632*, 2012.

19 Masao Doi. Stochastic theory of diffusion-controlled reaction. *Journal of Physics A: Mathematical and General*, 9(9):1479, 1976.

20 L Peliti. Path integral approach to birth-death processes on a lattice. *Journal de Physique*, 46(9):1469–1483, 1985.

21 Ludwig Gauckler and Harry Yserentant. Regularity and approximability of the solutions to the chemical master equation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 48(6):1757–1775, 2014.

22 Gerd Edzard Harry Reuter. Denumerable markov processes and the associated contraction semigroups onl. *Acta Mathematica*, 97(1-4):1, 1957.

23 Evgeniĭ Borisovich Dynkin. Markov processes. In *Markov Processes*, pages 77–104. Springer, 1965.

24 Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*, volume 194. Springer Science & Business Media, 1999.

25 Anthony F Bartholomay. Enzymatic reaction-rate theory: A stochastic approach. *Annals of the New York Academy of Sciences*, 96(1):897–912, 1962.

26 Kenji Ishida. The stochastic model for unimolecular gas reaction. *Bulletin of the Chemical Society of Japan*, 33(8):1030–1036, 1960.

27 Donald A McQuarrie. Kinetics of small systems. i. *The journal of chemical physics*, 38(2):433–436, 1963.

28 K Singer. Application of the theory of stochastic processes to the study of irreproducible chemical reactions and nucleation processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 92–106, 1953.

29 Allan I Solomon and J Katriel. Multi-mode q-coherent states. *Journal of Physics A: Mathematical and General*, 26(20):5443, 1993.

30 Jacob Katriel and Gérard Duchamp. Ordering relations for q-boson operators, continued fraction techniques and the q-cbh enigma. *Journal of Physics A: Mathematical and General*, 28(24):7209, 1995.

31 Allan Gut. *Probability: a graduate course*, volume 75. Springer Science & Business Media, 2012.

32 Achim Klenke. *Probability theory: a comprehensive course*. Springer Science & Business Media, 2013.

33 Frank WJ Olver. *NIST handbook of mathematical functions hardback and CD-ROM*. Cambridge University Press, 2010.

34 WN Everitt, KH Kwon, LL Littlejohn, and R Wellman. Orthogonal polynomial solutions of linear ordinary differential equations. *Journal of computational and applied mathematics*, 133(1):85–109, 2001.

35 N Gurappa, C Nagaraja Kumar, and Prasanta K Panigrahi. New exactly and conditionally exactly solvable n-body problems in one dimension. *Modern Physics Letters A*, 11(21):1737–1744, 1996.

36 N Gurappa and Prasanta K Panigrahi. Equivalence of the sutherland model to free particles on a circle. *arXiv preprint hep-th/9908127*, 1999.

37 N Gurappa, Prasanta K Panigrahi, T Shreecharan, and S Sree Ranjani. A novel method to solve familiar differential equations and its applications. In *Frontiers of Fundamental Physics 4*, pages 269–277. Springer, 2001.

38 N Gurappa, Prasanta K Panigrahi, and T Shreecharan. Linear differential equations and orthogonal polynomials: A novel approach. *arXiv preprint math-ph/0203015*, 2002.

39 PK Panigrahi, T Shreecharan, J Banerji, and V Sundaram. Coherent states: A general approach. *arXiv preprint quant-ph/0302171*, 2003.

40 N Gurappa and Prasanta K Panigrahi. On polynomial solutions of the heun equation. *Journal of Physics A: Mathematical and General*, 37(46):L605, 2004.

41 Yao-Zhong Zhang. Exact polynomial solutions of second order differential equations and their applications. *Journal of Physics A: Mathematical and Theoretical*, 45(6):065206, 2012.

42 Pawel Blasiak, Gerard HE Duchamp, Allan I Solomon, Andrzej Horzela, Karol A Penson, et al. Combinatorial algebra for second-quantized quantum theory. *Advances in Theoretical and Mathematical Physics*, 14(4):1209–1243, 2010.

43 Mira Bernstein and Neil JA Sloane. Some canonical sequences of integers. *Linear Algebra and its Applications*, 226:57–72, 1995.

44 Brian C Hall. *Lie groups, Lie algebras, and representations: an elementary introduction*, volume 222. Springer, 2015.
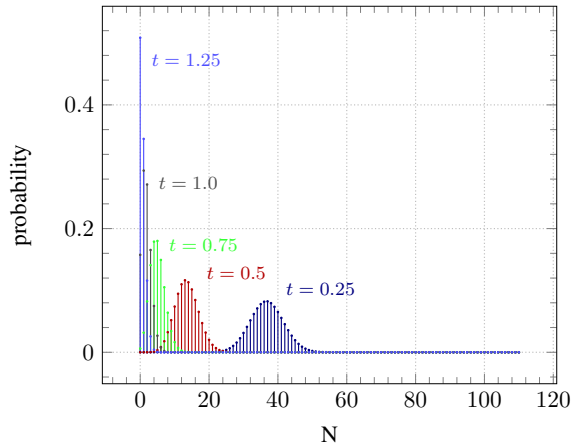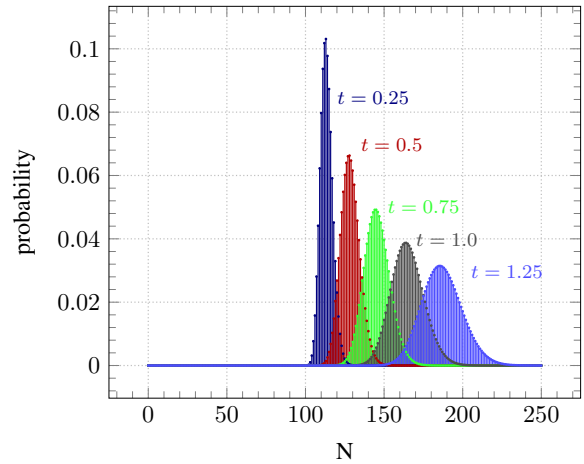
a) birth reaction $0A \xrightarrow{\beta=50} 1A$
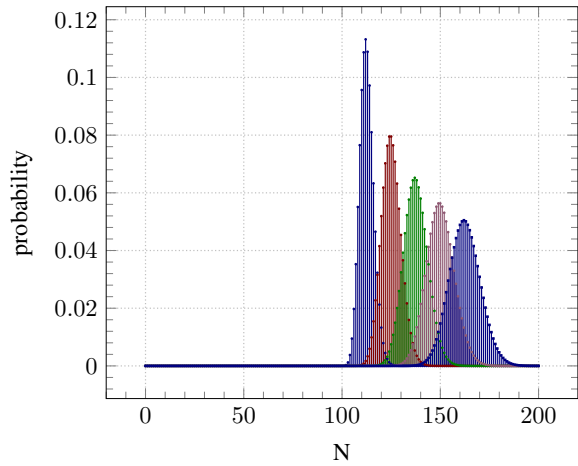
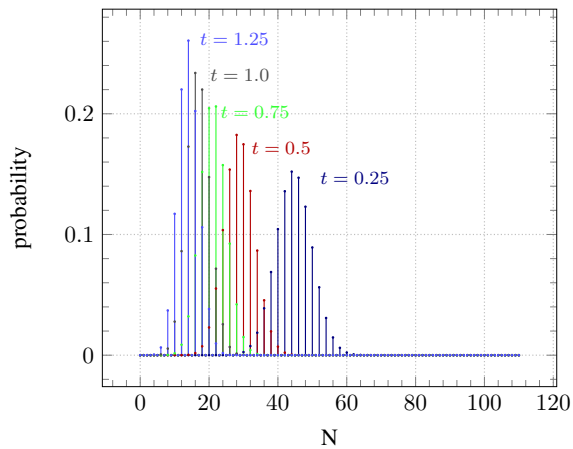**b)** pair creation reaction $0A \xrightarrow{\gamma=25} 2A$
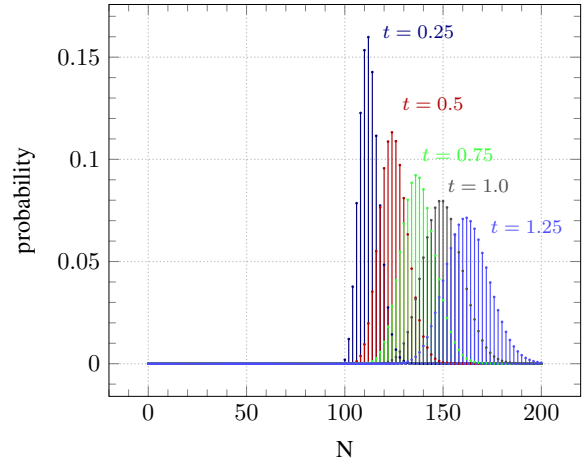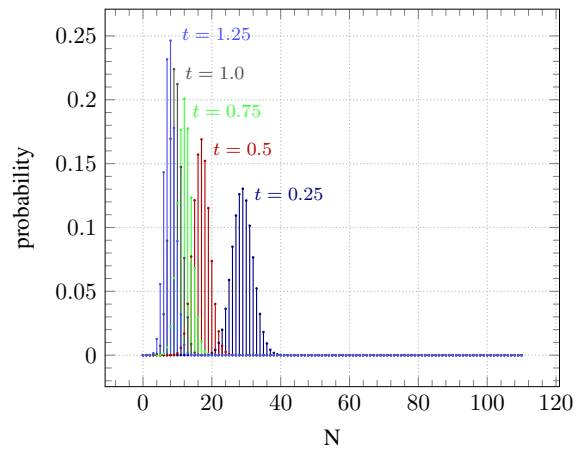

c) decay reaction $1A \xrightarrow{\tau=4} 0A$

d) autocatalysis reaction $1A \xrightarrow{\alpha=\frac{1}{2}} 2A$

e) pair annihilation reaction $2A \xrightarrow{\kappa=\frac{1}{40}} 0A$
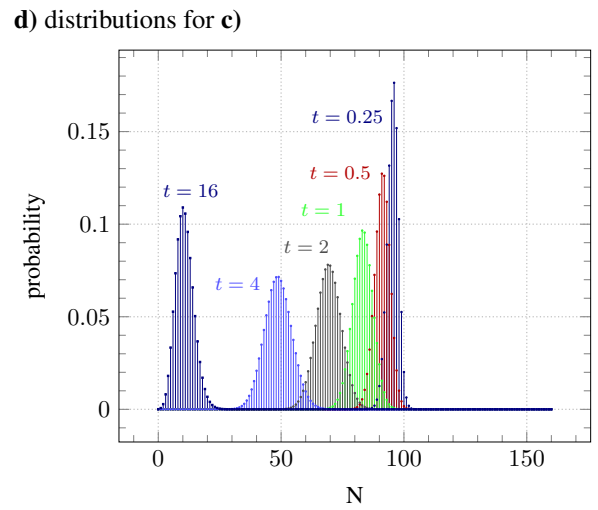

f) catalytic decay reaction $2A \xrightarrow{\lambda=\frac{1}{10}} 1A$

Fig. 2 Discrete probability distributions for initial state $|\Psi(0)\rangle = |100\rangle$ and for individual elementary reactions.

a) $1A \xrightarrow{\alpha=1/3} 2A,\ 0A \xrightarrow{\gamma=1/3} 2A,\ 1A \xrightarrow{\tau=1/3} 0A$

b) distributions for **a)**

c) $0A \xrightarrow{\beta=1/5} 1A,\ 0A \xrightarrow{\gamma=3/5} 2A,\ 1A \xrightarrow{\tau=1/5} 0A$

d) distributions for **c)**

e) $0A \xrightarrow{0.4} 1A,\ 0A \xrightarrow{0.2} 2A,\ 1A \xrightarrow{0.3} 0A,\ 1A \xrightarrow{0.1} 2A$

f) distributions for **e)**
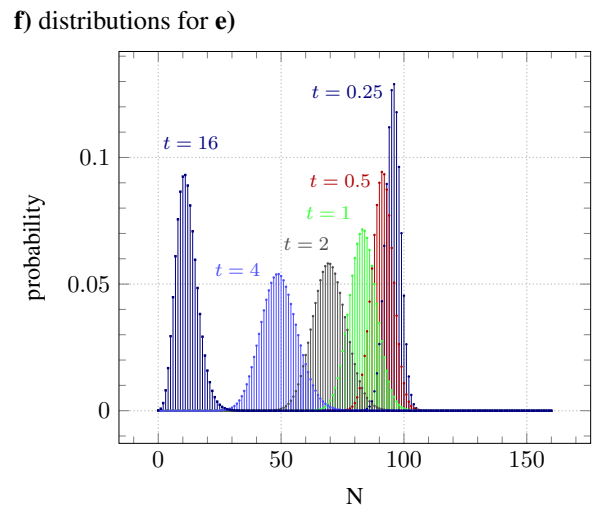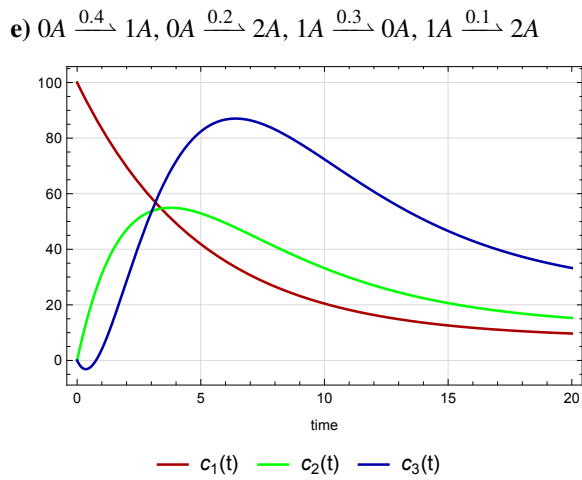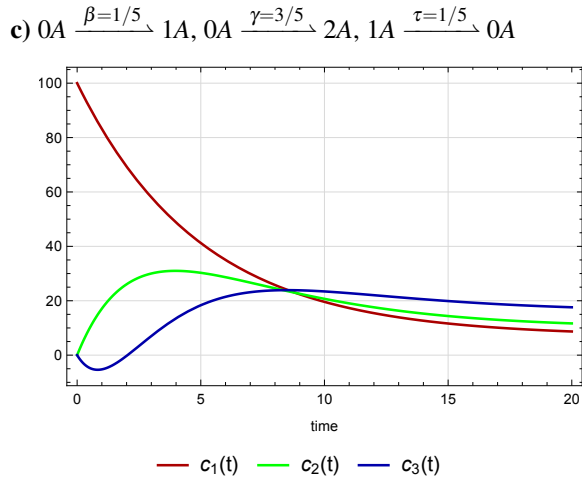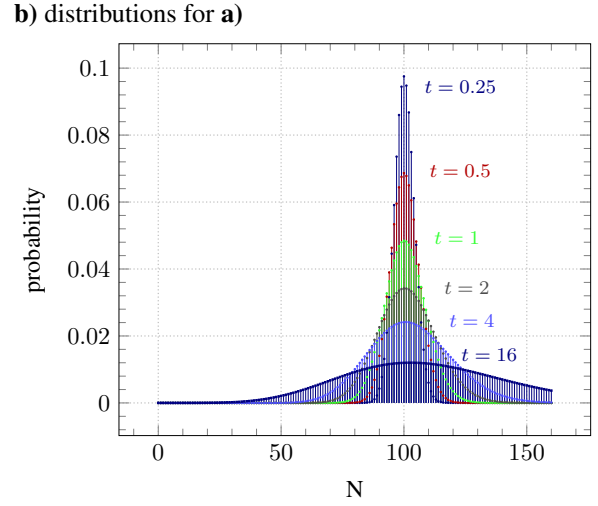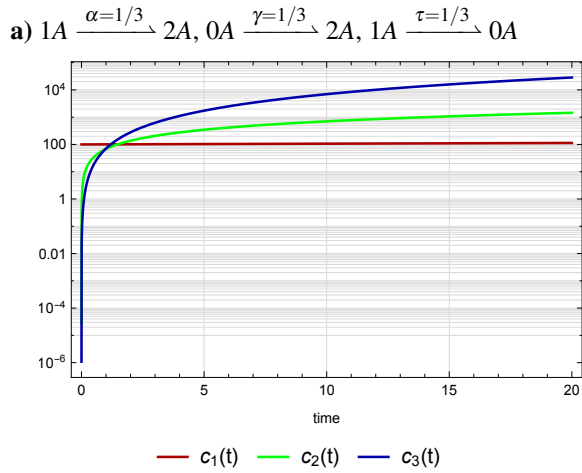
Fig. 3 First three cumulants $c_i(t)$ and discrete probability distributions for initial state $|\Psi(0)\rangle = |100\rangle$ and systems of non-binary elementary reactions.

a) Mean number of particles at time $t = 1$

cumulant $c_1(t = 1.)$
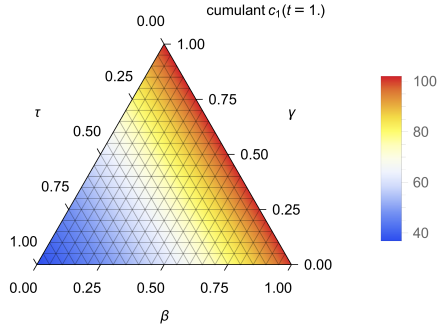
b) Variance of number of particles at time $t = 1$

cumulant $c_2(t = 1.)$
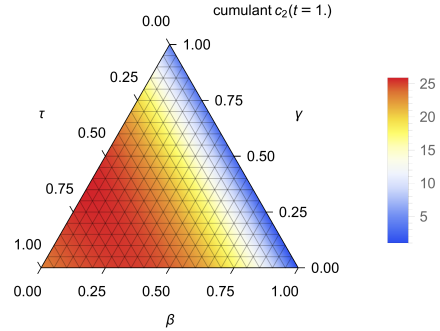
c) Mean number of particles at time $t = 4$

cumulant $c_1(t = 4.)$
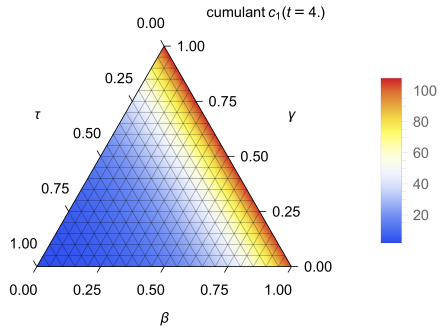
d) Variance of number of particles at time $t = 4$

cumulant $c_2(t = 4.)$
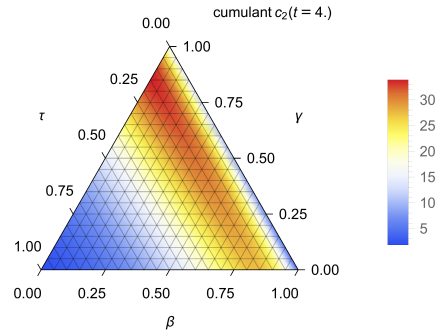
e) Mean number of particles at time $t = 16$

cumulant $c_1(t = 16.)$
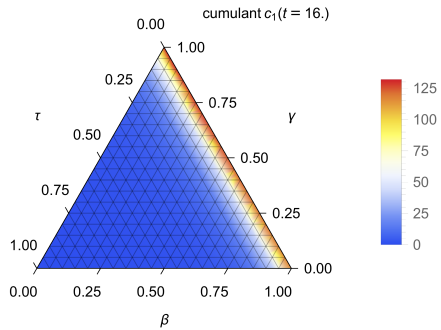
f) Variance of number of particles at time $t = 16$

cumulant $c_2(t = 16.)$
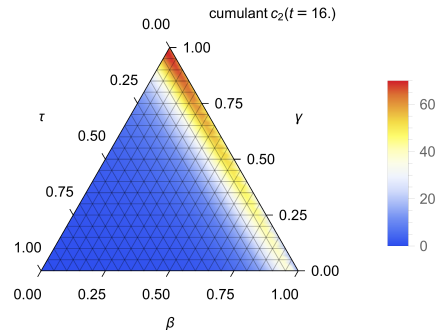
Fig. 4 Ternary parameter dependence plot for a reaction system composed of birth, pair creation and decay reactions, for initial state $|\Psi(0)\rangle = |100\rangle$.

A Comparison of stochastic and quantum mechanics formalism

Since this constitutes a rather crucial aspect of our formalism, we review here the key differences between stochastic and quantum mechanics. We will focus on the scenarios most directly comparable, namely the cases where the space of pure states is some denumerable space of configurations (for simplicity considered in 0 spacial dimensions), as e.g. in the case of a chemical reaction system vs. the case of a quantum mechanical harmonic oscillator. Our aim is not so much to perform an in-depth analysis, but rather allow the interested readers to sharpen their intuition about the mathematical and conceptual properties of stochastic mechanics. Finally, for the experts, we note that in the following comparison we only consider the case of *time-independent* infinitesimal generators H and Hamiltonians $\mathcal{H}$ (where the latter is not to be confused with the expression of the infinitesimal generator H in the Bargmann-Fock basis as used in the main text). We refer the interested readers to the excellent review [18] for references and further technical details.

A.1 Explanation of conventions in definitions of stochastic mechanics

In quantum mechanics, one would tend to normalize the basis and dual vectors differently, namely instead of $\langle m|n\rangle = \delta_{m,n}n!$ one would rather define the inner product such as to have orthonormal basis vectors. However, the present choice is motivated by the following combinatorial interpretation:

Lemma 3. *From the canonical commutation relation $[a,a^\dagger] = \mathbb{1}$, it follows by applying the Hermitean conjugation $\dagger$ to this equation that*

$$(a)^\dagger = a^\dagger, \quad \left(a^\dagger\right)^\dagger = a.$$

Fixing the normalization of $\langle 0|$ to

$$\langle 0|0\rangle := 1 \tag{138}$$

and using the defining relations of the canonical representation of the HW algebra to infer that

$$|n\rangle = a^{\dagger n}|0\rangle, \quad \langle m| = (|m\rangle)^\dagger = \langle 0|a^m, \tag{139}$$

this implies the stochastic mechanics choice of inner product,

$$\langle m|n\rangle = \delta_{m,n}n!.$$

Proof. The claim follows from a direct application of (142). It follows from the definition of the canonical representation of the HW algebra that

$$a^m|n\rangle = a^m a^{\dagger n}|0\rangle = \sum_{k=0}^{min(m,n)} k!\binom{m}{k}\binom{n}{k}a^{\dagger(n-k)}a^{(m-k)}|0\rangle = \Theta(n-m)(n)_m|n-m\rangle,$$

where $\Theta(x) = \delta_{x\geq 0}$ is the Heaviside step function, and where $(n)_m := m!/(n-m)!$ is a *falling factorial*. Together with

$$(a^r|0\rangle)^\dagger = 0 = \langle n|a^{\dagger n},$$

the claim follows then from

$$\langle m|n\rangle = \langle 0|a^m a^{\dagger n}|0\rangle = (n)_m\langle 0|a^{\dagger(n-m)}|0\rangle = n!\delta_{m,n}.$$

$\square$

Another consequence of the definition of the dual number vectors $\langle m|$ is that the dual reference vector $\langle\mathscr{R}|$ is in fact analogous to a *coherent state* in quantum mechanics:

Lemma 4. *The dual reference vector $\langle \mathscr{R}|$ as defined above may equivalently be defined as*[*]

$$\langle \mathscr{R}| = \langle 0| e^a, \tag{140}$$

which implies that it is a left Eigenvector with Eigenvalue 1 of the creation operator $a^\dagger$,

$$\langle \mathscr{R}| a^\dagger = \langle \mathscr{R}|. \tag{141}$$

Proof. By direct computation:

$$\langle \mathscr{R}| a^\dagger = \langle 0| e^a a^\dagger = \sum_{m=0}^{\infty} \frac{1}{m!} \langle 0| a^m a^\dagger = \sum_{m=1}^{\infty} \frac{(m)_1}{m!} \langle 0| a^{m-1} = \langle \mathscr{R}|.$$

$\square$

A.2 Comparison of stochastic and quantum mechanics

Notably, one of the crucial differences concerns the nature of the ambient spaces the two types of theory are formulated in: for quantum mechanics, since the space of ℓ^2-summable sequences with complex coefficients may be equipped with the structure of a *Hilbert space* whose dual is isomorphic to the space itself, and in contrast to the stochastic mechanics case there exists a rather precise construction known as the *Stone-von Neumann theorem* that uniquely relates Hermitean operators $\mathscr{H}$ to unitary evolution semigroups $\hat{U}(t)$,

$$\hat{U}(t) = e^{\frac{t}{i\hbar} \mathscr{H}}.$$

Moreover, the existence of a proper sesqui-linear inner product $\langle,\rangle$ renders expressions such as the expectation value

$$\langle \hat{O} \rangle_{|\Phi\rangle} = \langle \Phi| \hat{O} |\Phi\rangle$$

of a quantum mechanical observable $\hat{O}$ well-posed.

In stark contrast, no such mathematical elegance is available in the general case in the setting of stochastic mechanics in general. While in the case of a *finite* state space one has the rather similar relationship

$$\mathscr{E}(t) = e^{tH}$$

between infinitesimal generator H and evolution semi-group $\mathscr{E}(t)$, one in general is confronted with *denumerable* aka countably infinite-dimensional spaces and *unbounded* infinitesimal generators H. In this general setting, while the evolution equation

$$\frac{d}{dt} \mathscr{E}(t) = H \mathscr{E}(t)$$

remains valid, one needs to employ the intricate functional analysis techniques of the *Hille-Yosida theory* in order to interpret the precise relationship between H and $\mathscr{E}(t)$ further. Note also that the interpretation of observables and expectation values is far more direct than in the quantum mechanical case, since they amount to the standard respective notions of probability theory. For example,

$$\mathbb{E}_{|\Psi\rangle}(O) = \langle \mathscr{R}| O |\Psi\rangle = \sum_s \psi_s \langle \mathscr{R}| O |s\rangle$$
$$= \sum_s \psi_s \omega_s \langle \mathscr{R}|s\rangle = \sum_s \psi_s \omega_s,$$

which amounts to the conventional definition of the expectation value of an observable. It is therefore rather remarkable that the evolution equations for the expectation values of observables take such analogous looking forms (bearing in mind the entirely different meanings of expectation values in stochastic and quantum mechanics of course). In fact, these formal similarities motivated the moniker "stochastic mechanics" for this particular flavor of stochastic theories.

[*] In comparison to the standard definitions of (dual) coherent states,

$$|z\rangle := e^{-\frac{|z|^2}{2}} e^{z a^\dagger} |0\rangle, \quad \langle z| := e^{-\frac{|z|^2}{2}} \langle 0| e^{\bar{z} a} \quad (z \in \mathbb{C}),$$

the dual reference vector $\langle \mathscr{R}|$ is proportional to a dual coherent state with $z = 1$.

Table 4 Comparison of mathematical concepts of stochastic and of quantum mechanics.

concept	stochastic mechanics	quantum mechanics
pure states	$\lvert s\rangle \in S$	$\lvert s\rangle \in S$
state space	$Prob(S) := \left\{ \lvert\Psi\rangle = \sum_s \psi_s \lvert s\rangle \,\middle|\, \psi_s \in \mathbb{R}_{\geq 0} \text{ and } \sum_s \psi_s = 1 \right\}$ (probability distributions)	$Wave(S) := \left\{ \lvert\Psi\rangle = \sum_s \psi_s \lvert s\rangle \,\middle|\, \psi_s \in \mathbb{C} \text{ and } \sum_s \lvert\psi_s\rvert^2 = 1 \right\}$ (wave functions)
comments	$Prob(S)$ is *not* a Hilbert space!	$\ell^2_{\mathbb{C}}(S)$ is a Hilbert space, with *dual space* $\cong \ell^2_{\mathbb{C}}(S)$
ambient space	$\ell^1_{\mathbb{R}}(S)$	$\ell^2_{\mathbb{C}}(S)$
time-dependent state	$\lvert\Psi(t)\rangle = \mathscr{E}(t)\lvert\Psi(0)\rangle$	$\lvert\Phi(t)\rangle = \hat{U}(t)\lvert\Phi(0)\rangle$
evolution equations	$\frac{d}{dt}\lvert\Psi(t)\rangle = H\lvert\Psi(t)\rangle$ (Master equation) $\frac{d}{dt}\mathscr{E}(t) = H\mathscr{E}(t)$ $\mathscr{E}(t)$ – evolution semi-group $\mathscr{E}(s+t) = \mathscr{E}(s)\mathscr{E}(t) = \mathscr{E}(t)\mathscr{E}(s)$ $\mathscr{E}(0) = 1$	$i\hbar\frac{d}{dt}\lvert\Phi(t)\rangle = \mathscr{H}\lvert\Phi(t)\rangle$ (Schrödinger equation) $i\hbar\frac{d}{dt}\hat{U}(t) = \mathscr{H}\hat{U}(t)$ $\hat{U}(t)$ – unitary evolution operator aka *propagator* $\hat{U}(s+t) = \hat{U}(s)\hat{U}(t) = \hat{U}(t)\hat{U}(s)$ $\hat{U}(0) = 1$
properties of H vs. $\mathscr{H}$	$\langle\!\langle\mathscr{D}\rvert\, H = 0$	$\mathscr{H}^\dagger = \mathscr{H}$
structure of $\mathscr{E}(t)$ vs. $\hat{U}(t)$	$\mathscr{E}(t) \neq e^{tH}$	$\hat{U}(t) = e^{\frac{i}{\hbar}t\mathscr{H}}$ (Stone–von-Neumann theorem)
observables	O – diagonal linear operators: $O\lvert s\rangle = \omega_s\lvert s\rangle$ for all $\lvert s\rangle \in S$	$\hat{\mathscr{O}}$ – Hermitean operators: $\hat{\mathscr{O}}^\dagger = \hat{\mathscr{O}}$
expectation values of observables	$\mathbb{E}_{\lvert\Psi\rangle}(O) \equiv \langle O\rangle_{\lvert\Psi\rangle}$ $:= \langle\!\langle\mathscr{D}\rvert O\lvert\Psi\rangle$	$\langle\hat{\mathscr{O}}\rangle_{\lvert\Phi\rangle} := \langle\Phi\rvert O\lvert\Phi\rangle$
evolution of observables	$\frac{d}{dt}\langle O\rangle(t) = \langle[O,H]\rangle(t)$ $= \langle\!\langle\mathscr{D}\rvert[O,H]\lvert\Psi(t)\rangle$	$\frac{d}{dt}\langle\hat{O}\rangle(t) = \langle[O,H]\rangle(t)$ $= \langle\Phi(t)\rvert[O,H]\lvert\Phi(t)\rangle$

35

A.3 Normal-ordering in the HW algebra

The combinatorics encoded in the HW algebra is conveniently encoded in the well-known *normal ordering formula*:

$$a^{\dagger m} a^n a^{\dagger r} a^s = \sum_{k=0}^{min(n,r)} k! \binom{n}{k} \binom{r}{k} a^{\dagger (m+r-k)} a^{(n+s-k)} \,. \tag{142}$$

This formula may be derived in a number of ways, including inductive application of $[a, a^{\dagger}] = 1$, using the Bergman-Fock representation $a^{\dagger} \equiv z$, $a \equiv \partial_z$, or more intuitively via using diagram algebraic reasoning [5, 42]. It plays a crucial role in our framework, since it allows to represent any operator constructed from an arbitrary word in a and $a^{\dagger}$ in terms of its *normal-ordered form*, either by inductively applying (142) or via more sophisticated techniques from the combinatorics literature.

B Definitions and identities involving Stirling numbers

Denote by $s_1(k, \ell)$ the (unsigned) Stirling numbers of the first kind, and by $S_2(\ell, k)$ the Stirling numbers of the second kind, respectively:

$$s_1(k, \ell) := \delta_{k,0} \delta_{0,\ell} + \Theta(k-1)(-1)^{k-\ell} \begin{bmatrix} k \\ \ell \end{bmatrix}, \quad \begin{bmatrix} k \\ \ell \end{bmatrix} := \sum_{1 \le i_1 < \dots < i_{\ell-1} < k} \frac{(k-1)!}{i_1 \cdots i_{\ell-1}}$$

$$S_2(\ell, k) = \frac{1}{k!} \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} j^\ell \,. \tag{143}$$

Some useful identities involving the two kinds of Stirling numbers are ($\forall n > 0$)

$$s_1(n,0) = s_1(0,n) = 0, \ s_1(n,1) = (-1)^{n-1}(n-1)! \ s_1(n,n) = 1, \ s_1(n,n-1) = -\binom{n}{2}$$

$$S_2(0,k) = \delta_{k,0}, \ S_2(\ell,1) = \Theta(\ell-1), \ S_2(\ell,\ell) = 1\,.$$

and for all $n > k$ we have that $s_1(k, n) = S_2(k, n) = 0$. We also note the well-known orthogonality relations

$$\sum_{n=\ell}^{k} s_1(k,n) S_2(n,\ell) = \sum_{n=0}^{k} s_1(k,n) S_2(n,\ell) = \delta_{k,\ell}$$

$$\sum_{n=k}^{\ell} S_2(\ell,n) s_1(n,k) = \sum_{n=0}^{\ell} S_2(\ell,n) s_1(n,k) = \delta_{k,\ell}\,, \tag{144}$$

where equality of the respective two variants follows from $s_1(k, \ell) = S_2(k, \ell) = 0$ for $\ell > k$. The orthogonality relations are straightforward to derive from the exponential generating functions for the Stirling numbers,

$$\sum_{\ell=0}^{\infty} S_2(\ell,k) \frac{x^\ell}{\ell!} = \frac{(e^x-1)^k}{k!}, \quad \sum_{\ell=0}^{\infty} s_1(\ell,k) \frac{x^\ell}{\ell!} = \frac{(\ln(1+x))^k}{k!}\,. \tag{145}$$

The essential step in the derivation of the orthogonality relations consists in noticing that the functions $f(x) = e^x - 1$ and $g(x) = \ln(1 + x)$ are compositional inverses of one another, i.e. $f(g(x)) = g(f(x)) = x$.

Of particular importance in the present paper is the concept of *Stirling transforms* [43]. Given two sequences $(a_n)_{n \in \mathbb{Z}_{\ge 0}}$ and $(b_k)_{k \in \mathbb{Z}_{\ge 0}}$ that are related by

$$b_n = \sum_{k=0}^{n} S_2(n,k) a_k\,, \quad a_k = \sum_{n=0}^{k} S_1(k,n) b_n\,, \tag{146}$$

then their exponential generating functions $\mathscr{A}(x) = \sum_{k=0}^{\infty} a_k \frac{x^k}{k!}$ and $\mathscr{B}(x) = \sum_{k=0}^{\infty} b_k \frac{x^k}{k!}$ are related via

$$\mathscr{B}(x) = \mathscr{A}(e^x - 1)\,, \quad \mathscr{A}(x) = \mathscr{B}(\ln(x+1))\,. \tag{147}$$

C Proofs

C.1 Proof of the EGF evolution equation in its most general form

In preparation of the proof, let us state the following useful corollary of the Baker-Campbell-Hausdorff formula:

Proposition 7 (Cf. [44], Prop. 3.35)**.** *For two linear operators A, B and a formal variable λ, we obtain the following formal power-series result:*

$$e^{\lambda A} B e^{-\lambda A} = e^{\lambda \, ad_A} B, \tag{148}$$

where

$$ad_A B := [A, B] = AB - BA, \; ad_A^0 B := B, \; ad_A^q B := \underbrace{ad_A \circ \ldots \circ ad_A}_{q > 0 \text{ times}} B. \tag{149}$$

Proof. Noting that at this point we *do not* make any statement whatsoever about the convergence of the power-series involved in the above formula, we simply proceed completely analogously to [44] (p. 74, Ex. 14) and expand both exponentials into their Taylor series, followed by collecting terms:

$$
\begin{aligned}
e^{\lambda A} B e^{-\lambda A} &= \sum_{m,n \geq 0} \frac{\lambda^{m+n}(-1)^n}{m! n!} A^m B A^n \\
&= \sum_{p \geq 0} \lambda^p \sum_{m,n \geq 0} \delta_{m+n,p} \frac{(-1)^n}{m! n!} A^m B A^n \\
&= \sum_{p \geq 0} \frac{\lambda^p}{p!} \sum_{k=0}^{p} \binom{p}{k} A^k B (-A)^{p-k}.
\end{aligned}
$$

Finally, using the standard auxiliary formula

$$ad_X^p Y = \sum_{k=0}^{p} \binom{p}{k} X^k Y (-X)^{p-k},$$

we obtain the statement. $\qquad\square$

The following material is a variation on material by the first-named author as presented in [4, 6]:

Theorem 1. *For a CTMC with space of pure states S, initial state $|\Psi(0)\rangle \in Prob(S)$, evolution semi-group $\mathscr{E}(t)$ and infinitesimal generator H, the exponential moment-generating function $\mathscr{M}(t; \underline{\lambda})$ for a given (finite) set of observables $\{O_1, \ldots, O_n\}$ ($O_i \in \mathscr{O}(S)$) fulfills the evolution equation*

$$\frac{d}{dt} \mathscr{M}(t; \underline{\lambda}) = \sum_{q \geq 1} \frac{1}{q!} \langle \mathscr{R} | \left(ad_{(\underline{\lambda} \cdot \underline{O})}^{\circ q} H \right) e^{\underline{\lambda} \cdot \underline{O}} |\Psi(t)\rangle. \tag{17}$$

Here, for two linear operators A and B the notation $ad_A B$ stands for the adjoint action *aka* commutator,

$$ad_A B := [A, B] = AB - BA,$$

while $ad_A^{\circ q} B$ denotes the q-fold commutator ($q > 1$),

$$
\begin{aligned}
ad_A^{\circ q} B &:= \underbrace{(ad_A \circ \ldots \circ ad_A)}_{q \text{ times}} (B) \\
&= \underbrace{[A, [A, \ldots, [A, B] \ldots]]}_{q \text{ times}}.
\end{aligned}
$$

37

Proof. Recall first the following crucial properties of the evolution semi-group $\mathscr{E}(t)$ and of the infinitesimal generator H:

$$\frac{d}{dt}\mathscr{E}(t) = H\mathscr{E}(t), \quad \langle\mathscr{R}|H = 0.$$

Then the claim of the theorem follows from a judicious application of these properties and of Proposition 7:

$$
\begin{aligned}
\frac{d}{dt}\mathscr{M}(t;\underline{\lambda}) &= \frac{d}{dt}\langle\mathscr{R}|e^{\underline{\lambda}\cdot\underline{O}}\mathscr{E}(t)|\Psi(0)\rangle \\
&= \langle\mathscr{R}|e^{\underline{\lambda}\cdot\underline{O}}H\mathscr{E}(t)|\Psi(0)\rangle \\
&= \langle\mathscr{R}|\left(e^{\underline{\lambda}\cdot\underline{O}}He^{-\underline{\lambda}\cdot\underline{O}}\right)e^{\underline{\lambda}\cdot\underline{O}}\mathscr{E}(t)|\Psi(0)\rangle \\
&= \langle\mathscr{R}|\left(e^{ad_{\underline{\lambda}\cdot\underline{O}}}H\right)e^{\underline{\lambda}\cdot\underline{O}}|\Psi(t)\rangle \\
&= \langle\mathscr{R}|He^{\underline{\lambda}\cdot\underline{O}}|\Psi(t)\rangle + \sum_{q\geq 1}\frac{1}{q!}\langle\mathscr{R}|\left(ad_{\underline{\lambda}\cdot\underline{O}}^{\circ q}H\right)e^{\underline{\lambda}\cdot\underline{O}}|\Psi(t)\rangle,
\end{aligned}
$$

from which the claim follows by noting that the first term vanishes due to $\langle\mathscr{R}|H = 0$. □

C.2 Proof of the evolution equation Theorem 2

We need the following auxiliary formula:

$$
\begin{aligned}
ad_{\lambda\hat{n}}a^{\dagger o}a^i = \lambda[a^\dagger a, a^{\dagger o}a^i] &= \lambda\left[\sum_{k=0}^{\min(1,o)}k!\binom{1}{k}\binom{o}{k} - \sum_{k=0}^{\min(i,1)}k!\binom{1}{k}\binom{i}{k}\right]a^{\dagger(o+1-k)}a^{(i+1-k)} \\
&= \lambda(o-i)a^{\dagger o}a^i.
\end{aligned}
$$

Consequently,

$$ad_{\lambda\hat{n}}^{\circ q}a^{\dagger o}a^i = \lambda^q(o-i)^q a^{\dagger o}a^i.$$

The claim of the theorem then follows by a direct computation:

$$
\begin{aligned}
\frac{d}{dt}\mathscr{M}(t;\lambda) = \left\langle e^{\lambda\hat{n}}H\right\rangle(t) &= \left\langle\left(e^{\lambda\hat{n}}He^{-\lambda\hat{n}}\right)e^{\lambda\hat{n}}\right\rangle(t) \overset{(*)}{=} \sum_{q=1}^{\infty}\frac{1}{q!}\left\langle\left(ad_{\lambda\hat{n}}^{\circ q}\hat{H}\right)e^{\lambda\hat{n}}\right\rangle(t) \\
&= \sum_{(i,o)}r_{i,o}\left(e^{\lambda(o-i)}-1\right)\left\langle a^{\dagger o}a^i e^{\lambda\hat{n}}\right\rangle(t) \overset{(**)}{=} \mathbb{D}(\lambda,\partial_\lambda)\mathscr{M}(t;\lambda).
\end{aligned}
$$

Here, in the step marked $(*)$ we have used that $\langle\mathscr{R}|H = 0$ as well as the auxiliary relation

$$[a^\dagger a, H] = [a^\dagger a, \hat{H}],$$

while in the step marked $(**)$ we have used that $\langle\mathscr{R}|a^\dagger = \langle\mathscr{R}|$, and consequently

$$\left\langle a^{\dagger o}a^i e^{\lambda\hat{n}}\right\rangle(t) = \left\langle a^i e^{\lambda\hat{n}}\right\rangle(t) = \left\langle a^{\dagger i}a^i e^{\lambda\hat{n}}\right\rangle(t) = \sum_{\ell=0}^{i}s_1(i,\ell)\left\langle\hat{n}^\ell e^{\lambda\hat{n}}\right\rangle(t) = \left[\sum_{\ell=0}^{i}s_1(i,\ell)\frac{\partial^\ell}{\partial\lambda^\ell}\right]\mathscr{M}(t;\lambda).$$

C.3 Proof of the factorial moment EGF evolution Theorem 3

The proof is a straightforward application of a Stirling transform followed by a change of variables (cf. Appendix B). According to Theorem 2, the exponential generating function $\mathscr{M}(t;\lambda)$ of the moments of the number operator $\hat{n} = a^\dagger a$ and the EGF $\mathscr{F}(t;\nu)$ of the factorial moments $a^{\dagger k}a^k$ are related via the Stirling transform

$$\mathscr{M}(t;\lambda) = \mathscr{F}(t;e^\lambda - 1).$$

Therefore,

$$\frac{d}{dt}\mathscr{M}(t;\lambda) = \frac{d}{dt}\mathscr{F}(t;e^{\lambda}-1) = \mathbb{D}(\lambda,\partial_{\lambda})\mathscr{M}(t;\lambda) = \mathbb{D}(\lambda,\partial_{\lambda})\mathscr{F}(t;e^{\lambda}-1).$$

Performing a change of variables

$$v := e^{\lambda} - 1 \Rightarrow \lambda = \ln(v+1),$$

which in particular entails that

$$\frac{\partial^n}{\partial\lambda^n} = \sum_{k=0}^{n} S_2(n,k)(v+1)^k \frac{\partial^k}{\partial v^k},$$

the claim of the theorem follows from the formula (38) for $\mathbb{D}(\lambda,\partial_{\lambda})$ (together with the orthogonality relation of the Stirling numbers as presented in Appendix B):

$$\mathbb{d}(v,\partial_v) = \left[\mathbb{D}(\lambda,\partial_{\lambda})\right]_{\lambda\mapsto v:=e^{\lambda}-1} = \sum_{(i,o)} r_{i,o}\left((v+1)^{(o-i)}-1\right)\sum_{\ell=0}^{i} S_1(i,\ell)\frac{\partial^{\ell}}{\partial\lambda^{\ell}}$$

$$= \sum_{(i,o)} r_{i,o}\left((v+1)^{(o-i)}-1\right)\underbrace{\sum_{\ell=0}^{i} S_1(i,\ell)\sum_{k=0}^{\ell} S_2(\ell,k)(v+1)^k\frac{\partial^k}{\partial v^k}}_{\stackrel{(144)}{=}(v+1)^i\frac{\partial^i}{\partial v^i}}.$$

This concludes the proof.

C.4   Proof of the multi-species HW algebra auxiliary formula

In order to prove that

$$e^{\underline{\alpha}\cdot\underline{a}}f(\underline{a}^{\dagger},\underline{a})e^{\underline{\beta}\cdot\underline{a}^{\dagger}} = e^{\underline{\alpha}\cdot\underline{\beta}}e^{\underline{\beta}\cdot\underline{a}^{\dagger}}f\left(\underline{a}^{\dagger}+\underline{\alpha},\underline{a}+\underline{\beta}\right)e^{\underline{\alpha}\cdot\underline{a}},$$

it suffices to prove the equation for an elementary normal ordered term $f = \underline{a}^{\dagger\underline{r}}\underline{a}^{\underline{s}}$, namely via using the formula for the adjoint action of linear operators on one another,

$$e^A B e^{-A} = e^{ad_A}B.$$

We begin with the proof for a single species:

$$e^{\alpha a}a^{\dagger r} = \left(e^{\alpha a}a^{\dagger r}e^{-\alpha a}\right)e^{\alpha a} = \left(\sum_{n=0}^{\infty}\frac{\alpha^n}{n!}ad_a\left(a^{\dagger r}\right)\right)e^{\alpha a}.$$

It follows from the general normal ordering formula that

$$ad_a\left(a^{\dagger r}\right) = [a,a^{\dagger r}] = \Theta(r)ra^{\dagger(r-1)},$$

and by induction

$$ad_a^{\circ n}\left(a^{\dagger r}\right) = \Theta(r-n)(r)_b a^{\dagger(r-n)}.$$

This implies that

$$e^{\alpha a}a^{\dagger r} = \left(\sum_{n=0}^{r}\binom{r}{n}\alpha^n a^{\dagger(r-n)}\right)e^{\alpha a} = (a^{\dagger}+\alpha)^r e^{\alpha a}.$$

It follows thus (with the second formula obtained by Hermitean conjugation)

$$e^{\underline{\alpha}\cdot\underline{a}}\underline{a}^{\dagger\underline{r}} = \left(\underline{a}^{\dagger}+\underline{\alpha}\right)^{\underline{r}}e^{\underline{\alpha}\cdot\underline{a}}, \quad \underline{a}^{\underline{s}}e^{\underline{\beta}\cdot\underline{a}^{\dagger}} = e^{\underline{\beta}\cdot\underline{a}^{\dagger}}\left(\underline{a}+\underline{\beta}\right)^{\underline{s}}.$$

The claim of the proposition then follows by repeated application of these two identities.

C.5 Proofs of the multi-species generating function evolution equation theorem

To derive the evolution equation for the exponential moment generating function $\mathcal{M}(t;\underline{\lambda})$, we first use the normal ordering formula (45) to demonstrate that

$$[\hat{n}_i,\underline{a}^{\dagger\,\underline{r}}\underline{a}^{\underline{s}}] = [a_i^\dagger a_i, a_i^{\dagger\,r_i} a_i^{s_i}]\underline{a}^{\dagger\,(\underline{r}-r_i\underline{\delta}_i)}\underline{a}^{(\underline{s}-s_i\underline{\delta}_i)} = (r_i - s_i)\underline{a}^{\dagger\,\underline{r}}\underline{a}^{\underline{s}} \quad \Rightarrow \quad ad_{\underline{\lambda}\cdot\hat{n}}^{\circ q}\left(\underline{a}^{\dagger\,\underline{r}}\underline{a}^{\underline{s}}\right) = [\underline{\lambda}\cdot(\underline{r}-\underline{s})]^q\,\underline{a}^{\dagger\,\underline{r}}\underline{a}^{\underline{s}}.$$

The remaining part of the proof is then entirely analogous to the single-species case.

In order to prove the validity of the evolution equation for the multi-species factorial moment generating function $\mathcal{F}(t;\underline{v})$, observe first that the canonical commutation relations (43) entail that for all $i,j \in \mathbf{S}$ (with $\mathbf{S}$ the set of species) and for all $k,\ell \geq 0$,

$$\left[a_i^{\dagger\,k} a_i^k, a_j^{\dagger\,\ell} a_j^\ell\right] = 0. \tag{150}$$

Therefore, the orthogonality relations (144) for the Stirling numbers $s_1(k,n)$ and $S_2(n,k)$ generalize to the following multi-species variants:

$$\sum_{\underline{n}=\underline{\ell}}^{\underline{k}} \underline{s}_1(\underline{k},\underline{n})\underline{S}_2(\underline{n},\underline{\ell}) = \sum_{\underline{n}=\underline{0}}^{\underline{k}} \underline{s}_1(\underline{k},\underline{n})\underline{S}_2(\underline{n},\underline{\ell}) = \delta_{\underline{k},\underline{\ell}}, \quad \sum_{\underline{n}=\underline{k}}^{\underline{\ell}} \underline{S}_2(\underline{\ell},\underline{n})\underline{s}_1(\underline{n},\underline{k}) = \sum_{\underline{n}=\underline{0}}^{\underline{\ell}} \underline{S}_2(\underline{\ell},\underline{n})\underline{s}_1(\underline{n},\underline{k}) = \delta_{\underline{k},\underline{\ell}}$$

$$\underline{s}_1(\underline{x},\underline{y}) := \prod_{i\in\mathbf{S}}(s_1(x_i,y_i)), \quad \underline{S}_2(\underline{x},\underline{y}) := \prod_{i\in\mathbf{S}}(S_2(x_i,y_i)). \tag{151}$$

For precisely the same reason, the relationships between the moments of the number operators $\hat{n}_i$ ($i \in \mathbf{S}$) and the factorial moments are entirely analogous to the single species case presented in Theorem 2:

$$(\hat{\underline{n}})^{\underline{k}} = \sum_{\underline{\ell}=\underline{0}}^{\underline{k}} \underline{S}_2(\underline{k},\underline{\ell})\underline{a}^{\dagger\,\underline{\ell}}\underline{a}^{\underline{\ell}}, \quad \underline{a}^{\dagger\,\underline{\ell}}\underline{a}^{\underline{\ell}} = \sum_{\underline{k}=\underline{0}}^{\underline{\ell}} \underline{s}_1(\underline{\ell},\underline{k})(\hat{\underline{n}})^{\underline{k}}. \tag{152}$$

Therefore, it is the case also for the multi-species variant of the generating functions that $\mathcal{F}(t;\underline{nu})$ is a (multi-species) *Stirling transform* of $\mathcal{M}(t;\underline{\lambda})$:

$$\mathcal{M}(t;\underline{\lambda}) = \mathcal{F}(t;\underline{\eta}), \quad \eta_i := e^{\lambda_i} - 1. \tag{153}$$

Performing the *change of variables*

$$v_i := e^{\lambda_i} - 1 \Rightarrow \lambda_i = \ln(v_i + 1), \quad \frac{\partial^{\underline{p}}}{\partial\underline{\lambda}^{\underline{p}}} = \prod_{i\in\mathbf{S}}\left(\sum_{k_i=0}^{p_i} S_2(p_i,k_i)(v_i+1)^{k_i}\frac{\partial^{k_i}}{\partial v_i^{k_i}}\right), \tag{154}$$

the derivation of the evolution equation for $\mathcal{F}(t;\underline{v})$ is then entirely analogous to the single-species case (where the multi-species orthogonality relations (151) are to be used).

D Details of the derivation of the Sobolev-Jacobi polynomial formulae

D.1 On the Gurappa-Panigrahi -type formulae for the Sobolev-Jacobi polynomials

Consider the eigenfunction equation for the Jacobi-type second-order differential operator,

$$(D_{Jac}^{\alpha,\beta} + n(n+\alpha+\beta+1))f(x) = 0$$
$$D_{Jac}^{\alpha,\beta} := (1-x^2)\frac{d^2}{dx^2} + (\beta - \alpha - (\alpha+\beta+2)x)\frac{d}{dx}. \tag{155}$$

Employing the auxiliary relation

$$x^2\frac{d^2}{dx^2} = \left(x\frac{d}{dx}\right)^2 - \left(x\frac{d}{dx}\right) = D^2 - D \quad (D = x\frac{d}{dx}), \tag{156}$$

we may express the eigenfunction equation for $D_{Jac}^{\alpha,\beta}$ in the alternative form

$$
\begin{aligned}
(F_n(D) + P(x, \tfrac{d}{dx})) f_n(x) &= 0 \\
F_n(D) &= -D^2 - (\alpha + \beta + 1)D + n(n + \alpha + \beta + 1) = -(D - n)(D + n + \alpha + \beta + 1) \\
P(x, \tfrac{d}{dx}) &= \tfrac{d^2}{dx^2} + (\beta - \alpha)\tfrac{d}{dx}.
\end{aligned}
\tag{157}
$$

Specializing to $\alpha = -1$ and employing the Ansatz (117)

$$
F_\lambda(D)x^\lambda = 0 \ \Rightarrow \ y_\lambda(x) = c_\lambda \hat{G}_\lambda x^\lambda, \quad \hat{G}_\lambda := \sum_{m \geq 0} (-1)^m \left[\frac{1}{F(D)} P(x, \tfrac{d}{dx}) \right]^m
\tag{158}
$$

of Gurappa and Panigrahi, we may recover the formula for the Sobolev-Jacobi polynomials as presented in (119),

$$
\tilde{P}_n^{(-1,\beta)}(x) = \hat{G}_{\beta,n} x^n, \quad \hat{G}_{\beta,n} = \sum_{m \geq 0} \left[\frac{1}{(D - n)(D + n + \beta)} \left(\tfrac{d^2}{dx^2} + (\beta + 1)\tfrac{d}{dx} \right) \right]^m.
\tag{159}
$$

Finally, in order to prove the alternative formula (121) for the case $\alpha = -1$, we have to employ the relation

$$
\begin{aligned}
\frac{1}{(D - n)(D + n - 1)} \frac{d^2}{dx^2} x^n &= \frac{1}{(D + n - 1)} \frac{d^2}{dx^2} \frac{1}{(D - n - 2)} x^n = -\frac{1}{2(D + n - 1)} \frac{d^2}{dx^2} x^n \\
\Rightarrow \quad \left(\frac{1}{(D - n)(D + n - 1)} \frac{d^2}{dx^2} \right)^m x^n &= \frac{1}{m!} \left(-\frac{1}{2(D + n - 1)} \frac{d^2}{dx^2} \right)^m x^n,
\end{aligned}
\tag{160}
$$

where the latter statement may be derived from the former by induction on m. This concludes the proof of the alternative formula

$$
\tilde{P}_n^{(-1,-1)}(x) = e^{-\frac{1}{2} \frac{1}{D + n - 1} \frac{d^2}{dx^2}} x^n.
\tag{161}
$$

D.2 Decomposition coefficients

The linear combination coefficient formula (123),

$$
x^M = \sum_{n=0}^{M} A_\beta^{M,n} \tilde{P}_n^{(-1,\beta)} \left(\frac{x - \frac{r_\ell}{2}}{1 - \frac{r_\ell}{2}} \right),
\tag{162}
$$

may be transformed into a more tractable form as follows: via the change of variables

$$
y(x) = \frac{x - \frac{r_\ell}{2}}{1 - \frac{r_\ell}{2}}, \quad \Rightarrow \quad x(y) = \frac{r_\ell}{2} + y \left(1 - \frac{r_\ell}{2} \right),
$$

which entails in particular that

$$
x^M = \sum_{P=0}^{M} \binom{M}{P} \left(\frac{r_\ell}{2} \right)^{M-P} \left(1 - \frac{r_\ell}{2} \right)^P y^P.
$$

Using the decomposition formula

$$
y^P = \sum_{n=0}^{P} B_\beta^{P,n} \tilde{P}_n^{(-1,\beta)}(y),
$$

one may obtain an expression of the coefficients $A_\beta^{M,n}$ in terms of the coefficients $B_\beta^{M,n}$ as follows:

$$
\begin{aligned}
x^M = \sum_{n=0}^{M} A_\beta^{M,n} \tilde{P}_n^{(-1,\beta)}(y) &= \sum_{P=0}^{M} \binom{M}{P} \left(\frac{r_\ell}{2} \right)^{M-P} \left(1 - \frac{r_\ell}{2} \right)^P \sum_{n=0}^{P} B_\beta^{P,n} \tilde{P}_n^{(-1,\beta)}(y) \\
&= \sum_{n=0}^{M} \left(\sum_{P=n}^{M} \binom{M}{P} \left(\frac{r_\ell}{2} \right)^{M-P} \left(1 - \frac{r_\ell}{2} \right)^P B_\beta^{P,n} \right) \tilde{P}_n^{(-1,\beta)}(y) \\
\Rightarrow \quad A_\beta^{M,n} &= \sum_{P=n}^{M} \binom{M}{P} \left(\frac{r_\ell}{2} \right)^{M-P} \left(1 - \frac{r_\ell}{2} \right)^P B_\beta^{P,n}.
\end{aligned}
$$

41

In order to compute the decomposition coefficients $B^{M,n}_\beta$ for the two cases of interest, the following auxiliary formula proves quintessential:

Lemma 5. *Let $A \in \mathbb{Z}_{\geq 0}$ and $B \in \mathbb{R}_{\geq 0}$ be two parameters of the integral expression*

$$\mathscr{I}_{A,B} := \int_{-1}^{+1} dx\, (x-1)^A (x+1)^B. \tag{163}$$

Then the explicit formula for $\mathscr{I}_{A,B}$ reads

$$\mathscr{I}_{A,B} = \frac{(-1)^A A!\, \Gamma(B+1) 2^{A+B+1}}{\Gamma(A+B+2)}. \tag{164}$$

Proof. Via partial integration and induction – since by assumption $A \in \mathbb{Z}_{\geq 0}$ is an integer, we may use partial integration in order to reduce the exponent of the term $(x-1)^A$ to zero, with a typical step of the reduction reading

$$\mathscr{I}_{P,Q} = \int_{-1}^{+1} dx\, (x-1)^P (x+1)^Q = \frac{1}{(Q+1)} \int_{-1}^{+1} dx\, (x-1)^P \left(\frac{d}{dx}(x+1)^{Q+1} \right)$$

$$= \left[\frac{(x-1)^P (x+1)^{Q+1}}{(Q+1)} \right]_{-1}^{+1} - \Theta(P-1) \frac{P}{(Q+1)} \int_{-1}^{+1} (x-1)^{P-1}(x+1)^{Q+1}$$

$$\Leftrightarrow \quad \mathscr{I}_{P,Q} = \delta_{P,0} \frac{2^{Q+1}\Gamma(Q+1)}{\Gamma(Q+2)} - \Theta(P-1)\frac{P}{(Q+1)} \mathscr{I}_{P-1,Q+1}.$$

$\square$

Considering now the definitions of the Sobolev-type inner products $\Phi^{(-1,-1)}_{1,1}$ and $\Phi^{(-1,\beta)}_1$, respectively, we first need to compute the derivatives of the Sobolev-Jacobi polynomials $\tilde{P}^{(-1,\beta)}_n(x)$:

$$\frac{d}{dx}\tilde{P}^{(-1,\beta)}_n(x) := \delta_{n,1} + \Theta(n-2) \binom{2n-2}{n}^{-1} \sum_{k=1}^{n-1} \binom{n-1}{k}\binom{n-1}{n-k} \left[(n-k)(x-1)^{n-1-k}(x+1)^k + k(x-1)^{n-k}(x+1)^{k-1} \right]$$

With these preparations, and noting that

$$\tilde{P}^{(-1,-1)}_{n\geq 2}(-1) = \tilde{P}^{(-1,-1)}_{n\geq 2}(1) = 0,$$

the formulae for the coefficients may be derived. Using MATHEMATICA to simplify the resulting summation formulae, the coefficients $B^{M,n}_{-1}$ for the Sobolev-Jacobi polynomials $\tilde{P}^{(-1,-1)}_n(x)$ and $M > 0, n > 0$ (with $M+n \in 2\mathbb{Z}$) are computed as

$$B^{M,n}_{-1} = \frac{M!\,\Gamma(n+\frac{1}{2})}{2^{M-n} n!\, \Gamma(\frac{M+n+1}{2})\Gamma(1+\frac{M-n}{2})}. \tag{165}$$

The final form as presented in (130) may be obtained via using the identities

$$\Gamma(p+\tfrac{1}{2}) = \frac{(2p-1)!!}{2^p}\sqrt{\pi}, \quad (2p-1)!! = \frac{(2p)!}{2^p p!}. \tag{166}$$

The computations for the case $\beta > -1$ are entirely analogous, and are thus omitted here for brevity.