Moment Semantics for Reversible Rule-Based Systems

Vincent Danos[1]([⊠]), Tobias Heindel[2], Ricardo Honorato-Zimmer[2], and Sandro Stucki[3]

[1] Département d'Informatique, École Normale Supérieure, Paris, France
vincent.danos@gmail.com
[2] School of Informatics, University of Edinburgh, Edinburgh, UK
[3] Programming Methods Laboratory, EPFL, Lausanne, Switzerland

Abstract. We develop a notion of stochastic rewriting over marked graphs – i.e. directed multigraphs with degree constraints. The approach is based on double-pushout (DPO) graph rewriting. Marked graphs are expressive enough to internalize the 'no-dangling-edge' condition inherent in DPO rewriting. Our main result is that the linear span of marked graph occurrence-counting functions – or *motif functions* – form an algebra which is closed under the infinitesimal generator of (the Markov chain associated with) any such rewriting system. This gives a general procedure to derive the moment semantics of any such rewriting system, as a countable (and recursively enumerable) system of differential equations indexed by motif functions. The differential system describes the time evolution of moments (of any order) of these motif functions under the rewriting system. We illustrate the semantics using the example of preferential attachment networks; a well-studied complex system, which meshes well with our notion of marked graph rewriting. We show how in this case our procedure obtains a finite description of all moments of degree counts for a fixed degree.

Keywords: Stochastic processes · Moment semantics · Reversible Computing · Graph rewriting · Rule-based systems

1 Introduction

To explain the purpose of this paper, we start with a simple case of stochastic Petri net (PN) using the following pair of reactions:

$$A \xrightarrow{k_0} 2A \qquad\qquad (\rho_0)$$

$$A \xrightarrow{k_1} \varnothing \qquad\qquad (\rho_1)$$

The PN has a single species (or place) A and two reactions (or transitions) modelling the birth and death of cells. Given a mother cell, reaction ρ_0 will produce one daughter cell. Reaction ρ_1 models the death of cells. The firing rate of the reactions is given by the law of *mass action* and depends on the number N of cells present in the system as well as the rate constants k_0, k_1 of the reactions:

$$\theta_0 = k_0 N, \qquad\qquad \theta_1 = k_1 N.$$

The physical interpretation of the law of mass action states roughly that the propensity of a reaction is proportional to the concentration of the reactants, that is, the occupants of the left-hand side (LHS) of the reaction. We get a more computational interpretation by treating reactions such as ρ_0 and ρ_1 as *rewrite rules* over the state space of the PN. In order to apply a rule, we first need to *pattern-match* its LHS against the current state of the PN. The PN itself can then be seen as a labelled transition system, where transitions are rule applications, that is, they are identified by a rule together with an associated match of the LHS in the current state. If we assign the same constant transition rate to all applications of a given rule, the overall firing rate of that rule is exactly the product of said rate constant and the number of matches of the LHS in the current state. Returning to our birth and death model, the activity of the rule ρ_0 is k_0 times the number of ways one can match a single cell in a population of N, which is just $k_0 N$. These transitions and associated rates define a *continuous-time Markov chain* (CTMC), which provides the stochastic semantics of the PN. The CTMC is often expressed as a so-called *master equation* (ME), a system of differential equations describing the time evolution of the probability of finding the PN in a particular state [27].

As far as PNs go, the birth and death model is simple. Yet, because ρ_0 creates cells, its state space is countably infinite, and so is the number of equations in the ME. Nonetheless, the average evolution of the number of cells can be compactly described by the single rate equation (RE)

$$\tfrac{d}{dt}\,\mathbf{E}(N) \;=\; k_0\,\mathbf{E}(N) - k_1\,\mathbf{E}(N).$$

Indeed, the average occurrence count of any species in a given PN can always be approximated by a finite set of REs, providing us with a differential semantics for PNs [17,27].

1.1   Our Goal

In this paper, we wish to investigate similar REs but for models of dynamic networks that are more richly structured than PNs. We are looking for the following two ingredients:

1. a simple formal language that is flexible enough to capture a broad class of network dynamics
2. a method to generate REs for motif functions for any model of the above class

Regarding point 1, we propose a notion of *stochastic marked graph rewriting* which follows the general guidelines of the theory of graph transformation systems (GTS) [6,7,15,16,25,29]. Stochastic processes are modelled as rewrite rules over directed multigraphs with *marks* allowing for pre- and post-conditions on node degrees. Marked graphs double as a simple query language for identifying subgraphs subject to degree constraints. This provides a formal modeling framework in which we develop the method of point 2. We show how to generate (in general) countable systems of ordinary differential equations (ODEs) describing the mean evolution of marked graph motifs counts, or any higher-order statistics thereof. In fact, these ODEs completely describe the dynamics of the moments of marked graph observables. We therefore refer to them as the *moment semantics* of the rewrite system.

1.2   Preferential Attachment

We can elaborate on our basic birth and death model to illustrate these ideas. In the following, unless stipulated otherwise, graph is short for directed multigraph.

We start by endowing the model with a network structure. While the PN model allows us to track the evolution of a population over time, it does not capture mother-daughter relationships among cells. We now extend the PN model to a simple GTS that will do exactly that. The state of the PN is replaced by a directed multigraph with cells as nodes and edges pointing from daughters to mothers; the reactions of the PN are replaced by graph rewrite rules. This extension allows us to track "genealogical patterns" such as the number of sibling relationships. The updated rules of the birth and death model are

$$A \quad \xrightarrow{k_0} \quad A \leftarrow A' \qquad\qquad (\alpha_0)$$

$$A \quad \xrightarrow{k_1} \quad \varnothing \qquad\qquad (\alpha_1)$$

The birth rule α_0 introduces a new node A' (the daughter indicated by a prime) and a new edge $A \leftarrow A'$ representing the mother-daughter relationship; the death rule α_1 is identical to its counterpart ρ_1 from the PN model.

We can express sibling relationships through the motif

$$A' \rightarrow A \leftarrow A''$$

Tracking the number of such motifs amounts to counting the number of subgraphs in the state of our system that are isomorphic to the sibling graph.

So far, the sole purpose of edges in our model is to track relationships. This does not do justice to the expressive power of our rewrite formalism. In particular, there is no reason edges should not also influence the dynamics of the model. Let us add a third rule to illustrate this principle.

$$A' \rightarrow A \quad \xrightarrow{k_2} \quad A' \rightarrow A \leftarrow A'' \qquad\qquad (\alpha_2)$$
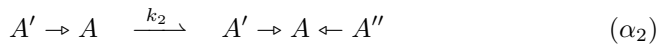
On their own, the rules α_0, α_1 model the evolution of a culture of rather uniform cells: any cell can divide or die at any time. Rule α_2, on the other hand, reflects

the fact that some cells may be more prolific than others: if a cell already has a daughter, it is likely to divide again. This positive feedback, known as *preferential attachment* (PA) or the *Matthew effect*, appears in many real-world complex systems and has been extensively studied [3,4,11,12].

At a first glance, the preferential attachment rule α_2 looks rather innocuous. It seems to be just a special case of the birth rule α_0. However, having a closer look at the right-hand side (RHS) of α_2, we realise that this rule directly creates siblings. Hence, we should expect a high k_2/k_0 ratio to increase the occurrence of siblings dramatically.

Just as for the simpler PN system, we can employ an RE to describe the evolution of the average number of sibling relationships over time. The RE consists in the following system of ordinary differential equations (ODE), with S, N and E counting siblings, cells (single nodes) and mother-daughter relationships (single edges), respectively:

$$\tfrac{d}{dt}\,\mathbf{E}(S) = 2(k_0 + k_2)\,\mathbf{E}(E) + 2k_2\,\mathbf{E}(S) - 3k_1\,\mathbf{E}(S)$$
$$\tfrac{d}{dt}\,\mathbf{E}(N) = k_0\,\mathbf{E}(N) + k_2\,\mathbf{E}(E) -\ \ k_1\,\mathbf{E}(N)$$
$$\tfrac{d}{dt}\,\mathbf{E}(E) = k_0\,\mathbf{E}(N) + k_2\,\mathbf{E}(E) - 2k_1\,\mathbf{E}(E)$$

It is easy enough to convince ourselves that (in the absence of parallel edges) this system of ODEs does indeed describe the evolution of the preferential attachment process: the equations for N and E follow the law of mass action (modulo symmetry factors); the equation for S essentially says that one needs to create daughters in order to create siblings. Note also the positive feedback of E on itself: thanks to the positive dependency of S on E, a high k_2/k_0 ratio will indeed lead to an explosion in siblings. With these intuitions in mind, a clever modeller could certainly have come up with these equations. Yet, this manual process is error-prone and does not scale well. The combinatorics involved are non-trivial: as we will see later, more complex models can involve hundreds or thousands of equations. This prompts the need for tools to automate the derivation of REs from GTS similar to those available for PNs.

Consider now a more complex motif (which can be expressed directly using marked graphs). Define $N_i(G)$ as the number of nodes in the graph G that have in-degree *exactly* i (i.e. mothers with exactly i daughters). Note that N_i does indeed more than counting subgraphs: contrary to what one might think, N_i does *not* count the number of subgraphs in G that consist of a central node with i incident edges (i.e. i-stars) as that would also cover all the nodes in G with in-degree larger than i. In particular, N_2 is *not* the sibling pattern. Instead, we think of N_i as counting the number of matches of the single-node graph A in G, subject to the condition that the matching node have exactly i incident edges. For reasons that will become clear later, we call such pairs of graphs and associated degree conditions *marked graphs*. Perhaps surprisingly, we can write REs even for marked graph observables. The RE system for N_i, $i \in \mathbb{N}$ (which is derived explicitly in the next Section) is given by the following system of ODEs:
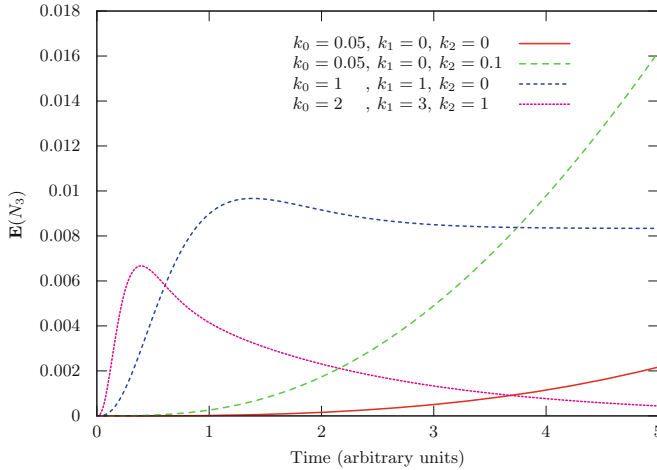
Fig. 1. Mean number of vertices with in-degree 3 (N_3) for various k_0, k_1, k_2

$$\tfrac{d}{dt}\,\mathbf{E}(N_i) = (k_0 + k_2(i-1))\,\mathbf{E}(N_{i-1}) - (k_0 + k_1(i+1) + k_2 i)\,\mathbf{E}(N_i) \quad \text{for } i \geq 1$$
$$\tfrac{d}{dt}\,\mathbf{E}(N_0) = k_0\,\mathbf{E}(N) + k_2\,\mathbf{E}(E) - (k_0 + k_1)\,\mathbf{E}(N_0)$$

Combined with the equations above for E and N, we have a set of equations which allows for a complete determination of the mean number of nodes of a given degree (with the size of the system being linear in the degree).

Fig. 1 shows the solutions of the RE for the case of $i = 3$ and various combinations of rates k_0, k_1, k_2. Note that the mean number of degree-three nodes is unbounded for low death rates, decays for high death rates, but reaches an equilibrium if the rates are suitably balanced. Yet it is unclear from these results alone *(i)* if N_3 converges in variance, and *(ii)* if its distribution around the mean is skewed. To answer *(i)* and *(ii)* we need to look at the *higher-order moments* of N_i. Luckily, we can derive ODEs not just for the mean but for arbitrary higher-order moments of marked graph observables. Due to the combinatorics involved, the resulting system of ODEs consists of 2097 equations! We therefore confine ourselves to presenting a plot (Fig. 2) summarising its solutions for the choice $k_0 = 1$, $k_1 = 2$ and $k_2 = 2$ of rates,[1] which shows that, despite the mean reaching an equilibrium, the skew and variance of N_i diverge over time. Clearly, the manual derivation of such an RE is beyond hope. We therefore developed a small, proof-of-concept tool for generating the REs in this example. The source code of our tool is freely available [1], and demonstrates that our construction can indeed be automated.

---

[1] It took approx. half a minute to generate the 2097 equations and another 33 minutes to solve them using GNU/Octave on a Intel Core i7 CPU.
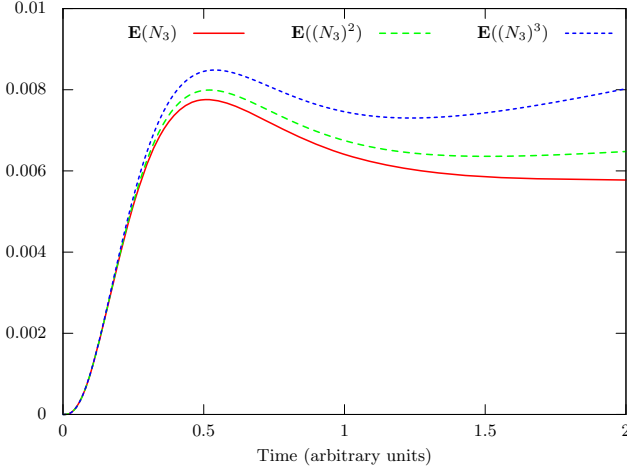
Fig. 2. The first three moments for the number of vertices with in-degree 3 (N_3) when $k_0 = 1, k_1 = 2, k_2 = 2$. The mean $\mathbf{E}(N_3)$ converges while the other moments may not.

1.3   A Sketch of the Solution

Let us briefly sketch our solution of the problem. Given a *marked graph observable* $[\tilde{G}](X)$, meaning a function counting the occurrences of the marked graph $\tilde{G}$ in the state X, we generate an ODE which describes the rate at which the mean occurrence count $\mathbf{E}([\tilde{G}](X(t)))$ changes over time. Careful inspection reveals that terms in the ODE are derived from the set of *minimal gluings* (MG) of the pattern $\tilde{G}$ with the LHS and RHS of the extant rules describing $X(t)$. (Note that the construction can thus be made incremental in the set of rules considered.) In particular, each term in $[\tilde{G}]$'s ODE depends on the current state *only* via expressions of the form $\mathbf{E}([\tilde{H}])$ for $\tilde{H}$ a pattern defining a new observable. This key property is referred to in the main part as *jump-closure* of graph observables. Each fresh observable $[\tilde{H}]$ can then be submitted to the same treatment.

To obtain higher-order moments, we exploit the commutative algebra structure of the linear space of pattern observables. We can compute $\mathbf{E}([\tilde{G}][\tilde{H}])$, i.e. covariances, etc. by expressing the product $[\tilde{G}][\tilde{H}]$ as a linear combination of motifs corresponding to the (minimal) gluings of $\tilde{G}$ and $\tilde{H}$. Though finite, the number of terms in the resulting expression is subject to the potentially high combinatorics of repeated pattern gluing.

As the generation of moment semantics is a symbolic procedure we can pursue it in principle to any order. This means that the order of the approximation is no longer limited to the humanly computable, and can be pushed further to acquire more accurate dynamics.

1.4   Related Work

Rate equations, and more generally mean field theories (MFT), are ubiquitous in the study of complex dynamics. Examples include, walkers on biopolymers [18,35]; models of epidemic spreading [21]; and of the evolution of social networks [13]. These examples witness both the power and universality of MFTs, and the fact that they are pursued in a seemingly ad hoc, and case-by-case fashion.

Conversely, various tools have been developed to automate the generation and solution of the ME and REs for the case of PNs [19,30,36]. But they suffer the limited expressivity of PNs as discussed above.

This paper follows ideas on applying the methods of abstract interpretation to the differential semantics of *site graph* rewriting [9,20,22]. From the GTS side, the theory of site graph rewriting had long been thought to be a lucky anomaly until a recent series of work showed that most of its ingredients could be made sense of, and given a much larger basis of applications, through the use of algebraic graph-rewriting techniques [2,23,24]. These latter investigations motivated us to try to address these questions at a higher level of generality [8]. Another more remote influence is Lynch's finite-model theoretic approach to rate equations [32].

1.5   Outline

The paper is organised as follows: §2 introduces the algebraic blueprint to build moment semantics and the notion of *jump-closure* of an algebra of observables; preferential attachment networks are used as an illustration; §3 introduces *marked graphs* and develops a formal stochastic graph transformation (GTS) framework based on *double-pushout* (DPO) rewriting of marked graphs; marks on graphs serve as simple *application conditions* and give rise to an algebra of *marked graph observables*, which are shown to be jump-closed with respect to the Markov chains generated by marked rewrite rules in §4; this is sufficient to derive the associated moment semantics.

2   The Blueprint of Moment Semantics

In this section we establish sufficient conditions for the existence of moment semantics for observables of suitable stochastic processes. These conditions provide the foundation on which we develop the moment semantics of concrete class of graphical rule-based systems in §3. We start with the necessary probabilistic preliminaries.

2.1   Markov Chains, Master Equation

Let S be an at most countable set, and write $\mathbb{R}^S$ for the vector space of real sequences indexed by S. We think of S as the state space of some process and

of elements of $\mathbb{R}^S$ as real-valued *observables* on that process. Let X_t, $t \geq 0$ be a continuous-time Markov chain (CTMC) with state-space S and *row-finite* rate matrix Q; we say an $S \times S$ matrix is row-finite if it has only finitely many non-zero coefficients per row. Write $p_x(t)$, or simply p_x, for the probability of X_t being at x, and let $p(t)$, or simply p, be the vector in $\mathbb{R}^S$ with coordinates p_x. Note that Q is a linear operator on $\mathbb{R}^S$. The evolution of p can be described by the *forward equation*, also known as the *master equation* (ME), which is the following linear ordinary differential equation (ODE) with values in $\mathbb{R}^S$ [27,33]:

$$\tfrac{d}{dt}p^T = p^T Q \tag{1}$$

In explicit coordinate form:

$$\tfrac{d}{dt}p_x = \sum_y p_y q_{yx} - p_x \sum_y q_{xy}.$$

The ME has unique (minimal non-negative) solutions at all times and for all initial conditions [33]. One caveat is that for *explosive* X_t, which have non-zero probability to complete countably many jumps in finite time, the resulting p is a sub-probability.

Example 1 (CTMC of the PA example). For the case of preferential attachment, we take the state space S to be a countable set of finite directed graphs G with finite node and edge sets V_G, E_G. We associate rate matrices Q_i, $i = 0, 1, 2$ to each of the rules α_i, where the rate q^i_{GH} of transitioning from a graph G to a graph H via the rule α_i, is given by the number of ways in which the LHS of α_i occurs as an isomorphic subgraph in G in such a way that replacing said subgraph with the RHS of α_i in G produces H:

$$q^0_{GH} = |\{n \in V_G \mid H = G^+(n)\}|$$
$$q^1_{GH} = |\{n \in V_G \mid H = G^-(n)\}|$$
$$q^2_{GH} = |\{e \in E_G \mid H = G^+(t(e))\}|$$

where $t(e)$ denotes the target node of the edge e, and $G^+(n)$, $G^-(n)$ are the graphs obtained, respectively, by adding a new node to G and connecting it to n, and by removing the node n and its edges from G. The overall rate matrix Q is given by $Q = \sum_i k_i Q_i$.

2.2 ODEs of Means

The rate matrix Q defines a linear transformation on $\mathbb{R}^S$ as follows:

$$(Qf)(x) = \sum_y q_{xy}(f(y) - f(x)). \tag{2}$$

Since the sum above is finite, Qf is indeed a well-defined element of $\mathbb{R}^S$. We call Qf the *jump* of f relative to Q. Intuitively, Qf is the expected rate of change in f given that the process sits at x. This interpretation of Q as a linear operator

on observables (rather than probabilities) is similar to predicate transformer semantics [5].

Given an observable f, we write $\mathbf{E}_p(f) := p^T f$ for the expected value of f according to p. From the (1) we can derive formally the following:

$$\tfrac{d}{dt}\,\mathbf{E}_p(f) \;=\; \tfrac{d}{dt} p^T f \;=\; p^T Q f \;=\; \mathbf{E}_p(Qf), \tag{3}$$

giving us an equation for the rate of change of the mean of $f(X_t)$. As such this equation is not very useful, $\mathbb{R}^S$ being an even larger index set than S. Indeed, unless S is finite, $\mathbb{R}^S$ does not even have a countable basis. But, suppose we are given a linear subspace $\mathscr{A}$ of $\mathbb{R}^S$ which *(i)* has a countable basis $\mathscr{B}$, and *(ii)* is *jump-closed* in the sense that $Q\mathscr{B} \subseteq \mathscr{A}$. Jump-closure means that for g in $\mathscr{B}$, one can write its jump Qg as:

$$Qg = \sum_{h \in \mathscr{B}} a_{g,h} h,$$

with finitely many non-zero coefficients $a_{g,h}$. By substituting this expression in (3), we get a linear ODE indexed by $\mathscr{B}$:

$$\tfrac{d}{dt}\,\mathbf{E}_p(g) = \sum_{h \in \mathscr{B}} a_{g,h}\,\mathbf{E}_p(h). \tag{4}$$

Note that the dependence in the probability distribution $p(t)$ of X_t has vanished! Thanks to jump-closure, (4) completely bypasses the probabilistic behaviour of the model, and predicts directly the mean evolution of the processes g_t for g in $\mathscr{B}$. The mean of any other observable f in $\mathscr{A}$ can then be expressed as a linear combination of the solutions of (4).

The vector space $\mathbb{R}^S$ can be equipped with the product topology. A linear map on $\mathbb{R}^S$ is continuous for that choice of topology iff it is row-finite. By a result of Shkarin [34, Theorem 2.3], any row-finite linear system of differential equations over $\mathbb{R}^S$ has solutions defined at all times and for all initial conditions. As $\mathscr{B}$ is countable, Shkarin's Theorem guarantees all-time (but, in general, non-unique) solutions for (4).

Example 2 (REs for Preferential Attachment). We will illustrate this idea by deriving the REs for the fixed-degree node-counting observable N_i from §1. We start by computing the jump of N_i with respect to Q_0 as given in Example 1. Expanding the definition of Q_0 in (2) and simplifying a bit, we get

$$(Q_0 N_i)(G) = \sum_{n \in V_G} (N_i(G^+(n)) - N_i(G)).$$

It is easy to verify that, for all $i \geq 1$, the difference under the sum is equal to -1 if n has in-degree i, to 1 if n has in-degree $(i-1)$, and to 0 otherwise. Hence the above simplifies to

$$Q_0 N_i = N_{i-1} - N_i, \qquad\qquad \text{for } i \geq 1.$$

Proceeding similarly for $Q_1 N_i$ and $Q_2 N_i$, one obtains

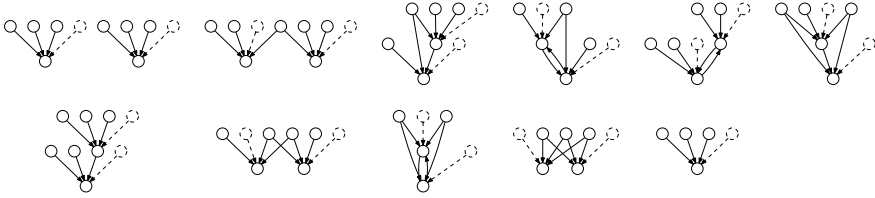$$QN_i = (k_0 + k_2(i-1))N_{i-1} - (k_0 + k_1(i+1) + k_2 i)N_i$$

Fig. 3. The eleven graph patterns $(S_3 + S_3)$, $G_1, \ldots, G_9$ and S_3 corresponding to base observables of the vector space containing $([S_3])^2$. Dotted nodes and edges represent negative existence conditions, i.e. ths absence of an additional edge; later to be represented by marks (see next Section).

for $i \geq 1$, suggesting $\{N_i\}_{i \in \mathbb{N}}$ as a candidate for the basis $\mathscr{B}$. However, the jump of N_0 depends on two additional observables, namely $N(G) = |V_G|$ and $E(G) = |E_G|$, counting the number of nodes and edges, respectively, in G:

$$QN_0 = k_0 N + k_2 E - (k_0 + k_1)N_0.$$

Expressing N as the infinite sum $N = \sum_{i=0}^{\infty} N_i$ does not solve the problem as a linear combination may only involve a finite number of vectors. Similarly, it is not clear how one would express the observable E in terms of N_i. Fortunately, the jumps of N and E do not involve any additional observables:

$$QN = k_0 N + k_2 E - k_1 N$$
$$QE = k_0 N + k_2 E - 2k_1 E$$

and we conclude that $\mathscr{B} = \cup_i \mathscr{B}_i$ with $\mathscr{B}_i = \{N, E\} \cup \{N_k\}_{k=0}^i$ form a jump-closed basis which indexes the RE for the N_i motifs. In this favourable case, each finite $\mathscr{B}_i$ spans a finite-dimensional subspace which is already jump-closed. Hence, solutions of the RE exist and are unique for all times and initial conditions.

2.3   Higher-Order Moments

So far, we have only considered the mean evolution of observables f in $\mathscr{A}$. One might also be interested in higher-order statistics of f, such as its standard deviation or skewness. Suppose then that $\mathscr{A} = \text{lin}(\mathscr{B})$ is a subspace of $\mathbb{R}^S$ as above, and additionally, that $\mathscr{A}$ is closed under (pointwise) product. One can write the powers of any f in $\mathscr{A}$, as a linear combination

$$f^n = \sum_{h \in \mathscr{B}} b_{f,h,n} h, \qquad\qquad n \geq 1,$$

where only finitely many of the coefficients $b_{f,h,k}$ are non-zero. Thus, solutions of (4) already describe *arbitrary* moments of observables. (Considerations on the existence and uniqueness of solutions remain the same.)

Example 3 (Higher moments of N_i). Suppose we want to compute the variance of N_i with respect to p, that is,

$$\mathbf{V}_p(N_i) \;=\; \mathbf{E}_p\big((N_i - \mathbf{E}_p(N_i))^2\big) \;=\; \mathbf{E}_p\big((N_i)^2\big) - \mathbf{E}_p(N_i)^2.$$

We already have an RE describing the mean of N_i, and we need one for the second moment of N_i, $\mathbf{E}_p((N_i)^2)$. To apply the above idea, it is enough to find a sub-algebra $\mathscr{A}$ which contains $\mathscr{B}$ and has a countable basis; the linear span of $\mathscr{B}$ alone does not work as we cannot even express the square N^2 of its "simplest" observable N as a linear combination in $\mathscr{B}_i$.

At this point it is worth making the following observation: all of the observables we have considered so far count the occurrences of degree-constrained graph patterns, that is, given a graph G, they count the number of subgraphs in G that are isomorphic to a fixed graph, subject to conditions fixing the degrees of some nodes. For example, the observable $N_i(G)$ counts the number of occurrences of the single-node graph in G that have in-degree i, while the observable $E(G)$ counts the number of single-edge graphs with no additional conditions. Later, we call such graph-counting functions *marked graph observables* and prove that they span a sub-algebra of $\mathbb{R}^S$. (A proof of this fact along with a more detailed account of marked graph observables is given in §3.3.)

To illustrate this point, consider first the simple node-counting observable N. Writing A for the single-node pattern (with no extra conditions), $A + A$ for its disjoint union, the two-node pattern, and $[P]$ for the observable counting the pattern P, we can express the second moment of N as the sum

$$\mathbf{E}_p(N^2) = \mathbf{E}_p\big(([A])^2\big) = \mathbf{E}_p([A + A] + [A])$$
$$= \mathbf{E}_p([A + A]) + \mathbf{E}_p([A]) = \mathbf{E}_p([A + A]) + \mathbf{E}_p(N).$$

Intuitively, we can write the product of any two pattern observables $[P_1]$, $[P_2]$ as a sum over all possible overlaps of P_1 and P_2 (including the trivial one). In the case of the pattern A, there are just two such overlaps, namely the trivial one, $A + A$, and the complete overlap A. We make this intuition precise in §3.3 by introducing the notion of *minimal gluings*.

A more complex example is the pattern S_k, which we define to be the *k-star*, that is, the graph consisting in a *hub* node n with k neighbours $m_1, \ldots, m_k$, each connected to the hub through a *spoke* (m_i, n). Furthermore, we impose the condition on S_k that the in-degree of its hub be exactly k. The motifs N_k and $[S_k]$ are related through the equation $[S_k] = k!N_k$, where the factor $k!$ is due to the internal symmetries of S_k. Hence it is enough to express $\mathbf{E}_p(([S_k])^2)$, to complete the example.

To find this expression, we proceed as in the simple case above and compute all overlaps of S_3 with itself (Fig. 3). The square $[S_k]^2$ can then be expressed as

$$[S_3]^2 = [S_3 + S_3] + 9[G_1] + 36[G_2] + 9[G_4] + 36[G_5]$$
$$+ 6[G_6] + 18[G_7] + 6[G_9] + 36[G_3] + 18[G_8] + 6[S_3].$$

where the patterns $(S_3 + S_3)$, $G_1, \ldots, G_9$ and S_3 are those shown in Fig. 3 (in that order). The combinatorial integer coefficients are the multiplicities of non-isomorphic overlaps resulting in the same graph; e.g. we can obtain G_1 in $9 = 3 \times 3$ different ways, as each copy of S_3 must decide independently which of its three peripheral nodes to share with the other copy.

For more complex patterns, such as S_k, the manual enumeration of overlaps becomes difficult due to the combinatorics involved. They are therefore best automated. To this end, we developed a small, proof-of-concept tool [1] for generating ODEs of higher-order moments of S_k. In fact, the equation and diagrams in this example and those of the previous section have been automatically generated.

We can summarise the above discussion.

Suppose given a triple $(S, Q, \mathscr{A})$, with S an at most countable set, Q a row-finite $S \times S$ rate matrix, and $\mathscr{A}$ a linear sub-space of $\mathbb{R}^S$ with a countable basis $\mathscr{B}$ such that $Q\mathscr{B} \subseteq \mathscr{A}$ (jump-closure). We can define a linear ODE system indexed by $\mathscr{B}$:

$$\tfrac{d}{dt}\, \mathbf{E}_p(g) = \sum_{h \in \mathscr{B}} a_{g,h}\, \mathbf{E}_p(h),$$

with finitely many non-zero coefficients $a_{g,h}$. The ODE can be described more concisely as $\frac{d}{dt} X = Q_{\mathscr{B}}^T X$ with $Q_{\mathscr{B}}$ the restriction of Q to $\mathscr{B}$ (which exists by jump closure), and $X(t) \in \mathbb{R}^{\mathscr{B}}$. One sees this to be just the master equation (1) restricted to $\mathscr{B}$.

If in addition $\mathscr{A}$ is a sub-algebra of $\mathbb{R}^S$, by linear combinations, we can derive from the above equations for any moment formed over $\mathscr{B}$.

It remains now to build an interesting example of this situation. In the next section we develop our concrete graphical framework of stochastic rule-based systems and associated observables, for which moment semantics can be built.

3   Reversible Stochastic Graph Rewriting

We turn now to the GTS framework for which we will derive moment semantics as outlined in §2. We build on a well-known approach from algebraic graph rewriting, namely the *double pushout* (DPO) approach [7,28]. The reasons for this choice are twofold: first, we profit from a solid body of preexisting work, and second, it allows for an "axiomatic" presentation abstracting over the details of the graph-like structures that are being rewritten. Indeed, while we only treat the case of directed multigraphs (graphs with an arbitrary number of directed edges between any two nodes), the theory generalizes to DPO rewriting in other adhesive categories [29] with negative application conditions [14,26].

We start with preliminaries on directed multigraphs, followed by a brief summary of the DPO approach and its stochastic semantics [25]. Next, we introduce *marked graphs* as a means to add simple application conditions to graph rewrite rules. We establish two key properties of marks: firstly, we show that marks are sufficient to internalize the reversibility conditions inherent in DPO rewriting into rewrite rules; secondly, marked graphs give rise naturally to a class of observables over graphs with the algebraic structure outlined in §2. As we will see in §4, such observables are jump-closed with respect to CTMCs defined by any finite set of marked rewrite rules.

3.1   Graph-Theoretic Preliminaries and DPO Rewriting

A *finite directed multigraph* (henceforth simply *graph*) G consists of a finite set of *nodes* V_G, a finite set of *edges* E_G, and *source* and *target* maps $s_G, t_G \colon E_G \to V_G$. A *morphism* $f \colon G \to H$ between graphs G and H is a pair of maps $f_V \colon V_G \to V_H$, $f_E \colon E_G \to E_H$ on edges and nodes, such that for every edge e in E_G,

$$s_H(f_E(e)) = f_V(s_G(e)) \qquad \text{and} \qquad t_H(f_E(e)) = f_V(t_G(e)).$$

The graphs G and H are called the *domain* and *codomain* of f, respectively. Given a pair $f \colon F \to G$, $g \colon G \to H$ of morphisms, their *composition* $(g \circ f) \colon F \to H$ is defined as $(g \circ f) = (g_V \circ f_V, g_E \circ f_E)$. A graph morphism $f \colon G \to H$ is a *monomorphism*, or simply a *mono*, if f_V and f_E are injective; it is a *graph inclusion* if both f_V and f_E are inclusion maps, in which case G is a *subgraph* of H and we write $G \subseteq H$. Every morphism $f \colon G \to H$ defines a subgraph $f(G) \subseteq H$ called the *direct image* (or just the *image*) of f in H, such that $V_{f(G)} = f_V(V_G)$ and $E_{f(G)} = f_E(E_G)$. Finally, a graph morphism $f \colon G \to H$ is an *isomorphism*, or simply an *iso*, if f_V and f_E are bijections. Given an iso $f \colon G \to H$, we say that G is *isomorphic* to H and write $G \simeq H$.

Graph morphisms provide us with a notion of *pattern matching* on graphs. We restrict pattern matching to monos: a *match* of G in H is a mono $f \colon G \to H$. This restriction ensures that $f(G) \simeq G$ if f is a match, that is, matches of G in H identify subgraphs in H that are isomorphic to G. It is easy to verify that the composition of two matches is again a match. We write $[G; H]$ for the set of matches of G in H.

The main ingredient for graph rewriting are rewrite rules. A *rule* $\alpha \colon L \rightharpoonup R$ with *left-hand side* (LHS) L and *right-hand side* (RHS) R is a pair $\alpha_1 \colon K \to L$, $\alpha_2 \colon K \to R$ where both α_1 and α_2 are monos. Given a rule $\alpha \colon L \rightharpoonup R$, we define its *inverse* $\alpha^\dagger \colon R \rightharpoonup L$ as $(\alpha_1, \alpha_2)^\dagger = (\alpha_2, \alpha_1)$.

By combining matches and rules, we obtain *derivations*, the basic rewrite steps of a GTS. We first describe them informally. Fig. 4 shows a derivation with a match $f \colon L \to G$ on the left and a rule $\alpha \colon L \rightharpoonup R$ on top. The match f identifies the subgraph in G that is to be modified, while the rule α describes how to carry out the modification. In order to obtain the *comatch* $g \colon R \to H$ on the right, one starts by removing nodes and edges from $f(L)$ which do not have a preimage under $f \circ \alpha_1$ (colored red in the figure). This operation is allowed only if it leaves no edges dangling in G, that is, a node may be removed only if all its incident edges are also removed. To complete the derivation, one extends the resulting match $h \colon K \to D$ by adjoining to D the nodes and edges in R that do not have a preimage under α_2 (colored green in the figure). The two monos β_1 and β_2 witness, respectively, the deletions from G and additions to D, and form the *corule* of the derivation.

Derivations constructed in this way have the defining property of being *double pushouts* (DPO): they consist of a pair of *pushout squares* (PO) of graph morphisms [7, 28].

There are certain points worth noting. Firstly, not every pair $f \colon L \to G$, $\alpha \colon L \rightharpoonup R$ of compatible matches and rules gives rise to a derivation. The reason
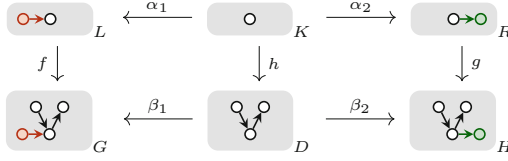
Fig. 4. A simple derivation or rewrite step: one starts with a match f for the left-hand side L of the rule in G; then, one constructs D by deleting nodes in G which correspond to nodes in L with no preimage in K; finally, one adds nodes and edges found in R but not in K
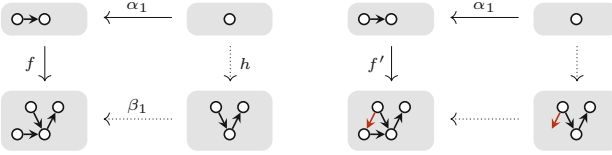

Fig. 5. The pair f, α_1 of monos has a POC, while the pair f', α_1 does not

is that the required left-hand PO does not exist for every pair $f\colon L \to G$, $\alpha_1\colon L \to K$ of monos. A suitable pair $h\colon K \to D$, $\beta_1\colon D \to G$ of monos, called *pushout complement* (POC) of f and α_1, exists iff the removal of nodes from G does not result in dangling edges. Fig. 5 illustrates this point. If a POC h, β_1 exists for f, α_1, then it is unique up to (unique) iso on D. Secondly, given a pair $h\colon K \to D$, $\alpha_2\colon K \to R$ of monos, the right-hand PO always exists, and the corresponding pair $g\colon R \to H$, $\beta_2\colon D \to H$ of monos is unique up to (unique) iso on H. Thirdly, if there is a derivation of g from f by α, then by symmetry, there is also a derivation of f from g via $\alpha^\dagger$.

Importantly, derivations *compose* and *split* (Fig. 6). Given a derivation of g_1 from f_1 by α with corule γ (the top DPO) and a derivation of g_2 from f_2 by γ with corule β (the bottom DPO), one obtains a composite derivation of $g = g_2 \circ g_1$ from $f = f_2 \circ f_1$ via α with corule β, by pasting together the two DPO diagrams. Conversely, derivations split along factorizations of matches: given the outer and top derivations of g from f and g_1 from f_1 via α, with corules β and γ, respectively, there is, for every f_2 such that $f = f_2 \circ f_1$, a unique bottom derivation with rule γ, comatch g_2 and corule β, such that $g = g_2 \circ g_1$.

Whenever there is a derivation with match f, comatch g and rule α, we say g is α-*derivable* from f and write $f \Rightarrow_\alpha g$. A pair of matches $f\colon G \to H$, $f'\colon G \to H'$ of a graph G is said to be isomorphic if there is an iso $u\colon H \to H'$ such that $f' = u \circ f$. It follows directly from the above discussion that the binary relation $\Rightarrow_\alpha$ is *(i)* partial, in that not every match of the LHS of α extends to a derivation, *(ii)* functional up to iso, that is, if $f \Rightarrow_\alpha g$ and $f \Rightarrow_\alpha g'$ then $g \simeq g'$, and *(iii)* injective up to iso, that is, if $f \Rightarrow_\alpha g$ and $f' \Rightarrow_\alpha g$ then $f \simeq f'$.

The fact that derivations are only defined up to iso is convenient as it allows us to invert them without paying attention to the concrete naming of nodes and
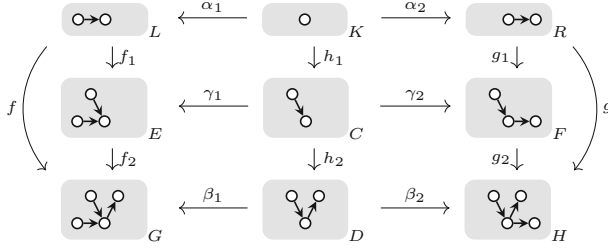
Fig. 6. A vertical composition of derivations

edges. Indeed, the inverse of $\Rightarrow_\alpha$ is just $(\Rightarrow_\alpha)^{-1} = (\Rightarrow_{\alpha^\dagger})$. On the other hand, when defining the stochastic semantics of rule-based systems, it is more convenient to restrict $\Rightarrow_\alpha$ to a properly functional relation. To this end, we fix once and for all a (countable) set $\mathscr{G}$ of *representatives* from every isomorphism class of graphs, and denote by $\alpha(f)$ and $f(\alpha)$, for any given rule $\alpha\colon L \rightharpoonup R$ and match $f \in \mathrm{dom}(\Rightarrow_\alpha)$, the unique comatch $\alpha(f)\colon R \to H$ and corule $f(\alpha)_1\colon D \to G$, $f(\alpha)_2\colon D \to H$ of the unique derivation for which both D and H are in $\mathscr{G}$. Note that the partial maps $\alpha(-)$ and $\alpha^\dagger(-)$ need not be inverses. Indeed, if α involves node deletions, one cannot choose $\alpha(-)$ and $\alpha^\dagger(-)$ such that $\alpha^\dagger(\alpha(f)) = f$ for all $f \in \mathrm{dom}(\Rightarrow_\alpha)$ because the node and edge sets of the respective codomains differ in general.

Given a rule $\alpha\colon L \rightharpoonup R$, define the rate matrix Q_α over $\mathscr{G}$ as

$$q^\alpha_{GH} = |\{f \in [L; G] \mid f \in \mathrm{dom}(\Rightarrow_\alpha) \text{ and } \alpha(f) \in [R; H]\}| \qquad \text{for } G \neq H,$$
$$q^\alpha_{GG} = \sum_{H \neq G} -q^\alpha_{GH} \qquad\qquad \text{otherwise.}$$

Given a finite set of rules $\mathscr{R}$ and a *rate map* $k\colon \mathscr{R} \to \mathbb{R}^+$, let

$$Q(\mathscr{R}, k) = \sum_{\alpha \in \mathscr{R}} k(\alpha) Q_\alpha.$$

This defines a CTMC over $\mathscr{G}$. As $\mathscr{R}$ is finite, the rate matrix $Q(\mathscr{R}, k)$ is row-finite.

3.2   Marked Graphs

So far, our notion of pattern matching is rather limited. While monos identify (isomorphic) images of a "pattern" (i.e. their domain) in other graphs, they provide no way of imposing additional conditions on the image of the pattern. We have seen examples in §1 and §2 where such conditions were used to count the number of nodes with a particular in-degree. But conditions are also useful for rewriting: by equipping the LHS and RHS of rules with conditions, one obtains more expressive rewrite formalisms. A particular case which has been studied in the DPO setting is that of *negative application conditions* (NAC), where the
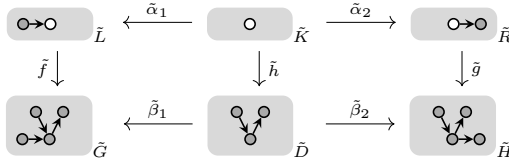
Fig. 7. A marked derivation

set of derivations under a rule α is restricted to instances where the match and comatch respect conditions associated with the LHS and RHS of α [14, 26]. In this section, we extend graphs with simple degree conditions which we call *marks*. Marks can be seen as a very simple type of NAC. Yet, they are expressive enough to cover the example patterns[2] form §1 and §2, and to internalize the dangling-edge conditions seen in the previous section into the LHS of rules.

A *marked graph* $\tilde{G}$ is a graph G together with a *marking* predicate $M_G \subseteq V_G$ over nodes. We say a node $x \in V_G$ is *marked* if $x \in M_G$ and *unmarked* otherwise. We write $\overline{M}_G$ for the complement $V_G \setminus M_G$ of the marking M_G. We say a marked graph $\tilde{G}$ is *complete* if all its nodes are marked, that is, $M_G = V_G$.

A *marked morphism* $\tilde{f} \colon \tilde{G} \to \tilde{H}$ is a graph morphism $f \colon G \to H$ that *preserves* marks, that is, $f_V(M_G) \subseteq M_H$. Marked morphisms compose as their underlying graph morphisms. The definitions of subgraphs, inclusions and monos generalize straightforwardly to marked graphs and morphisms. It is easy to see that isos *reflect* marks, that is, a marked morphism $\tilde{f} \colon \tilde{G} \to \tilde{H}$ is a marked iso if its underlying graph morphism f is an iso and $f_V(M_G) = M_H$. We define the marking of the direct image $\tilde{f}(\tilde{G})$ of a marked morphism $\tilde{f} \colon \tilde{G} \to \tilde{H}$ by $M_{\tilde{f}(\tilde{G})} = f_V(M_G)$, that is, marks on nodes in $f(H)$ that are provided only by $\tilde{H}$ are not considered part of the marking of $\tilde{f}(\tilde{G})$. Fig. 7 shows a marked version of Fig. 4 where marked nodes are colored in dark gray, unmarked nodes in white. The graphs in the bottom row of Fig. 7 are complete while those in the top row are not. All morphisms in the figure are marked monos.

We interpret marks as conditions on node degrees: a *marked match* of $\tilde{G}$ in $\tilde{H}$ is a marked mono $\tilde{f} \colon \tilde{G} \to \tilde{H}$ that preserves and reflects the degrees of marked nodes, that is, for all $x \in M_G$, $\mathrm{indeg}_H(f_V(x)) = \mathrm{indeg}_G(x)$ and $\mathrm{outdeg}_H(f_V(x)) = \mathrm{outdeg}_G(x)$. Given a marked graph $\tilde{G}$, and an unmarked match $f \colon G \to H$, we say f *extends* to a marked match $\tilde{f} \colon \tilde{G} \to \tilde{H}$ if there is at least one marking on H for which $\tilde{f}$ is a marked match. Such an extension will exist iff nodes marked in $\tilde{G}$ are mapped by f to nodes of same degree in H. The composition of marked matches is again a marked match. We write $[\tilde{G}; \tilde{H}]$

---

[2] Strictly speaking, the example patterns in §1–2 require a slightly more expressive type of NACs than the one described in this section. In particular, they require separate conditions on in and out-degrees of nodes, where as we only consider conditions on the overall node degree here. However, the theory extends straight-forwardly to the case where in and out-degrees are represented by separate markings. We have implemented the more general case in [1].

for the set of marked matches of $\tilde{G}$ in $\tilde{H}$. All the morphisms in Fig. 7 except $\tilde{\beta}_1$ and $\tilde{\beta}_2$ are marked matches.

By allowing rules to be marked, we obtain a simple form of NACs. Marks in the LHS act as *preconditions* on a rule, marks in the RHS as *postconditions*. We further restrict the type of morphisms that may appear in rules to *rigid monos*: we say a marked mono $\tilde{f}\colon \tilde{G} \to \tilde{H}$ is rigid if every node in $\tilde{H}$ is marked unless it is the image of an unmarked node in $\tilde{G}$, that is, $\overline{M}_H = f_V(\overline{M}_G)$. A *marked rule* $\tilde{\alpha}\colon \tilde{L} \rightharpoonup \tilde{R}$ is a pair $\tilde{\alpha}_1\colon \tilde{K} \to \tilde{L}$, $\tilde{\alpha}_2\colon \tilde{K} \to \tilde{R}$ of rigid marked monos. We say $\tilde{\alpha}$ is *minimally marked* if $M_K = \varnothing$. We write $\tilde{\alpha}^\dagger$ for the inverse $(\tilde{\alpha}_2, \tilde{\alpha}_1)$ of a marked rule $\tilde{\alpha}$. The rules $\tilde{\alpha}$ and $\tilde{\beta}$ in Fig. 7 are marked but only $\tilde{\alpha}$ is minimally marked. (Minimally marked rules are used later to embed normal DPO rewriting into our marked graph rewriting.)

The rigidity condition on rules ensures that rules are well-behaved in the following way: given a marked rule $\tilde{\alpha}\colon \tilde{L} \rightharpoonup \tilde{R}$ with underlying (unmarked) rule α and an unmarked match $f\colon L \to G$ of its LHS,

1. there is a g such that $f \Rightarrow_\alpha g$ if f extends to a marked match, and
2. g extends to a marked match if f does.

The first point internalizes the no-dangling-edge condition into rules and roughly corresponds to a notion of type safety: if both the rule α and its "argument" f are "well-marked", then α can be applied safely to f. The second point is reminiscent of a predicate transformer semantics: the marking on $\tilde{R}$ is the "strongest postcondition" given the marking on the LHS $\tilde{L}$. Conversely, the marking on $\tilde{L}$ is the "weakest precondition" given $\tilde{R}$. Note that, by symmetry, the same holds for $\tilde{\alpha}^\dagger$ so that the pre and postcondition uniquely determine each other.

A PO of a pair $\tilde{f}\colon \tilde{G} \to \tilde{H}$, $\tilde{f}'\colon \tilde{G} \to \tilde{H}'$ of marked morphisms, is a commuting square of marked morphisms (see (5) below), where the underlying graph morphisms form a PO (6), and $M_F = g_V(M_H) \cup g'_V(M_{H'})$.

$$
\begin{array}{ccc}
\tilde{G} \xrightarrow{\tilde{f}'} \tilde{H}' \\
{\scriptstyle \tilde{f}}\downarrow \qquad \downarrow{\scriptstyle \tilde{g}'} \quad (5) \\
\tilde{H} \xrightarrow{\tilde{g}} \tilde{F}
\end{array}
\qquad\qquad
\begin{array}{ccc}
G \xrightarrow{f'} H' \\
{\scriptstyle f}\downarrow \qquad \downarrow{\scriptstyle g'} \quad (6) \\
H \xrightarrow{g} F
\end{array}
$$

A *marked derivation* is a DPO of marked monos such as in Fig. 7, with $\tilde{f}$, $\tilde{g}$, $\tilde{h}$ marked matches and $\tilde{\alpha}$, $\tilde{\beta}$ marked rules. We call $\tilde{g}$ the *marked comatch* and $\tilde{\beta}$ the *marked corule* of the derivation.

POs of marked monos have interesting properties when one mono is rigid.

Lemma 1. *A pair of marked monos $\tilde{f}\colon \tilde{G} \to \tilde{H}$, $\tilde{g}\colon \tilde{H} \to \tilde{F}$ where $\tilde{f}$ is rigid, has a POC if $\tilde{g}$ is a marked match. (If in addition, $\tilde{G}$ is markless, then $\tilde{g}$ must be a marked match for a POC to exist.)*

Lemma 2. *Let* (5) *be a PO of marked monos, with $\tilde{f}'$ rigid. Then*

1. *$\tilde{g}$ is rigid,*
2. *$\tilde{f}$ is a marked match iff $\tilde{g}'$ is,*
3. *the pair $\tilde{f}, \tilde{g}$ is the unique (up to unique iso) POC of $\tilde{f}', \tilde{g}'$ and*

$$M_H = g_V^{-1}(M_F \setminus g_V'(M_{H'})) \cup f_V(M_G).$$

The two lemmas above are sufficient to establish the properties of marked rules outlined before; besides they also allow one to prove that marked derivations compose and split along marked matches (which is key to the proof in the next section).

Write $\tilde{f} \Rightarrow_{\tilde{\alpha}} \tilde{g}$ if $\tilde{g}$ is $\tilde{\alpha}$-derivable from $\tilde{f}$. Just as its unmarked counterpart, the binary relation $\Rightarrow_{\tilde{\alpha}}$ is functional and injective up to iso. However, *contrary* to $\Rightarrow_{\alpha}$, the relation $\Rightarrow_{\tilde{\alpha}}$ is also total (by Lemma 1). We fix again a (countable) set $\mathscr{G}$ of representatives from every isomorphism class of marked graphs. Given any marked rule $\tilde{\alpha} \colon \tilde{L} \rightharpoonup \tilde{R}$ and match $\tilde{f} \colon \tilde{L} \to \tilde{G}$, the definitions of the maps $\tilde{\alpha}(-)$ and $\tilde{f}(-)$, as well as the $\mathscr{G} \times \mathscr{G}$ rate matrix $Q_{\tilde{\alpha}}$ associated with $\tilde{\alpha}$, are completely analogous to those for the unmarked case.

At this point, the reader might be wondering whether the restrictions imposed on marked rules cause any loss of expressivity with respect to unmarked DPO rewriting. They do not. Indeed, one can build an embedding of unmarked DPO rewriting in the marked variant. We start by noticing that for every (unmarked) graph G there is a unique complete marked graph $\tilde{G}_K$, and for every (unmarked) rule $\alpha \colon L \rightharpoonup R$ there is a unique minimally marked rule $\tilde{\alpha}_{\min} \colon \tilde{L}_{\min} \rightharpoonup \tilde{R}_{\min}$. To every match $f \colon L \to G$ of an LHS L in some G corresponds a unique marked mono $\tilde{f} \colon \tilde{L}_{\min} \to \tilde{G}_K$. By Lemma 1, there is a derivation $f \Rightarrow_{\alpha} g$ for some $g \colon R \to H$ iff $\tilde{f}$ is a marked match and by Lemma 2 this uniquely determines an extension $\tilde{g} \colon \tilde{R}_{\min} \to \tilde{H}_K$ of g to a marked match. Note that due to the rigidity of $\tilde{\alpha}_{\min}$, the codomain $\tilde{H}_K$ of $\tilde{g}$ must be complete.

Although we have only presented the case of directed multigraphs, marked stochastic DPO rewriting also straight-forwardly extends to other graph-like structures, such as typed graphs or hypergraphs. Another example is that of PNs, when seen as rewriting discrete typed graphs (graphs with no edges). Since nodes in a discrete graph have degree zero, any match trivially extends to a marked match.

3.3   Minimal Gluings and the Algebra of Marked Graph Observables

Given a marked graph $\tilde{G}$ in $\mathscr{G}$, define the *marked graph observable* $[\tilde{G}] \colon \mathscr{G} \to \mathbb{N}$ to be the integer-valued function $[\tilde{G}](\tilde{H}) = |[\tilde{G}; \tilde{H}]|$ counting the number of occurrences (i.e. marked matches) of $\tilde{G}$ in any given graph $\tilde{H}$ in $\mathscr{G}$. Marked graph observables turn out to be the natural choice of observable functions over which to construct moment semantics for marked DPO rewriting. In this section, we present their algebraic structure, establishing the connection to the framework developed in §2. To do so, we first need to introduce a key ingredient: *minimal gluings* (MG).
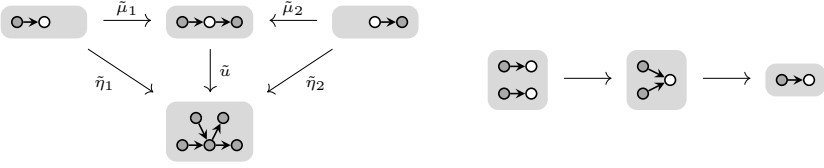
Fig. 8. Left: a gluing $\tilde{\eta}$ of the marked graphs $\tilde{L}$ and $\tilde{R}$ from Fig. 7 and the corresponding MG $\tilde{\mu}$. Right: the codomains of the MGs in $\tilde{L} * \tilde{L}$ ordered by (non mono) marked graph morphisms among them

Given a pair of subgraphs $G_1 \subseteq H$, $G_2 \subseteq H$ of a graph H, the *union* of G_1 and G_2 in H is the unique subgraph $G_1 \cup G_2$ of H, such that $V_{(G_1 \cup G_2)} = V_{G_1} \cup V_{G_2}$ and $E_{(G_1 \cup G_2)} = E_{G_1} \cup E_{G_2}$. The union $\tilde{G}_1 \cup \tilde{G}_2 \subseteq \tilde{H}$ of a pair of marked subgraphs $\tilde{G}_1 \subseteq \tilde{H}$, $\tilde{G}_2 \subseteq \tilde{H}$ has $G_1 \cup G_2$ as its underlying graph and $M_{G_1} \cup M_{G_2}$ as its marking. A *gluing* $\tilde{\eta}$ of a pair $\tilde{G}_1$, $\tilde{G}_2$ of marked graphs is a pair of marked matches $\tilde{\eta}_1 \colon \tilde{G}_1 \to \tilde{U}$, $\tilde{\eta}_2 \colon \tilde{G}_2 \to \tilde{U}$ with common codomain $\mathrm{cod}(\tilde{\eta}) = \tilde{U}$. We say $\tilde{\eta}$ is *minimal* if $\mathrm{cod}(\tilde{\eta}) = \tilde{\eta}_1(\tilde{G}_1) \cup \tilde{\eta}_2(\tilde{G}_2)$. Two gluings $\tilde{\eta}$ and $\tilde{\mu}$ are *isomorphic* if there is a marked iso $\tilde{u} \colon \mathrm{cod}(\tilde{\eta}) \to \mathrm{cod}(\tilde{\mu})$, such that $\tilde{\mu}_1 = \tilde{u} \circ \tilde{\eta}_1$ and $\tilde{\mu}_2 = \tilde{u} \circ \tilde{\eta}_2$. Write $\tilde{G}_1 *_{\simeq} \tilde{G}_2$ for the set of isomorphism classes of minimal gluings of $\tilde{G}_1$ and $\tilde{G}_2$, and $\tilde{G}_1 * \tilde{G}_2$ for a choice of representatives from each class in $\tilde{G}_1 *_{\simeq} \tilde{G}_2$ such that $\mathrm{cod}(\tilde{\mu}) \in \tilde{\mathcal{G}}$ for all $\tilde{\mu}$ in $\tilde{G}_1 * \tilde{G}_2$. It is easy to verify the following:

Lemma 3. *Let $\tilde{G}_1$, $\tilde{G}_2$ be marked graphs, then* (i) *$\tilde{G}_1 * \tilde{G}_2$ has $O(2^{N+M})$ elements, with $N = |V_{G_1}| + |V_{G_2}|$, $M = |E_{G_1}| + |E_{G_2}|$, and* (ii) *for every gluing $\tilde{\eta}$ of $\tilde{G}_1$ and $\tilde{G}_2$, there is a unique MG $\tilde{\mu}$ in $G_1 * G_2$ and marked match $\tilde{u} \colon \mathrm{cod}(\tilde{\eta}) \to \mathrm{cod}(\tilde{\mu})$ such that $\tilde{\eta}_1 = \tilde{u} \circ \tilde{\mu}_1$ and $\tilde{\eta}_2 = \tilde{u} \circ \tilde{\mu}_2$.*

Fig. 8 shows a gluing $\tilde{\eta}$ and its corresponding MG $\tilde{\mu}$ on the left, and the codomains of a set of minimal self-gluings on the right.

Thanks to MGs, marked graph observables form an algebra.

Theorem 1. *Let $\mathscr{B}$ be the set of marked graph observables. The linear space $\mathscr{A} = \mathrm{lin}(\mathscr{B})$ spanned by $\mathscr{B}$ is a sub-algebra of $\mathbb{R}^{\tilde{\mathcal{G}}}$, that is, $\mathrm{poly}(\mathscr{A}) = \mathscr{A}$.*

Proof. As $\mathscr{A}$ is a linear subspace of $\mathbb{R}^{\tilde{\mathcal{G}}}$, it suffices to show that $\mathscr{B}$ is closed under product. First, note that the product of any two marked graph observables $[\tilde{G}_1]$ and $[\tilde{G}_2]$ in $\mathscr{B}$ counts exactly the number of gluings $\tilde{G}_1 \to \tilde{H} \leftarrow \tilde{G}_2$ of marked matches in some common $\tilde{H} \in \tilde{\mathcal{G}}$. By Lemma 3, we can express such products as a (finite) linear combination of observables $[\mathrm{cod}(\tilde{\mu})]$ corresponding to the (codomains of) the MGs under $\tilde{G}_1$ and $\tilde{G}_2$:

$$[\tilde{G}_1](\tilde{H})[\tilde{G}_2](\tilde{H}) = |[\tilde{G}_1; \tilde{H}]| \cdot |[\tilde{G}_2; \tilde{H}]| = |\{\tilde{G}_1 \xrightarrow{\tilde{\eta}_1} \tilde{H} \xleftarrow{\tilde{\eta}_2} \tilde{G}_2 \mid \tilde{\eta} \text{ a gluing}\}|$$
$$= \sum_{\tilde{\mu} \in \tilde{G}_1 * \tilde{G}_2} [\mathrm{cod}(\tilde{\mu})](\tilde{H}).$$

Since $\mathrm{cod}(\tilde{\mu}) \in \tilde{\mathcal{G}}$ for all $\tilde{u}$ in $\tilde{G}_1 * \tilde{G}_2$, the result is again in $\mathscr{A}$. $\qquad\square$

Note that the set $\tilde{\mathscr{G}}$ contains the empty graph $\varnothing$, which makes $\mathscr{A}$ a *unitary* algebra with unit $[\varnothing]$.

4   Jump-Closure of Marked Graph Observables

We now have all the ingredients in place to derive moment semantics for (DPO-based) marked graph rewriting. The set $\tilde{\mathscr{G}}$ forms a countable state space over which we generate CTMCs from finite sets of marked rules and associated rate maps. The space $\mathscr{A}$ spanned by marked graph observables provides us with a candidate sub-algebra of $\mathbb{R}^{\tilde{\mathscr{G}}}$. It remains to show that $\mathscr{A}$ is jump-closed with respect to the CTMCs generated by marked rules.

Theorem 2. *Let $\mathscr{R}$ be a finite set of marked rules with associated rate map $k\colon \mathscr{R} \to \mathbb{R}^+$, and $\mathscr{B}$ the set of marked graph observables. The linear subspace $\mathscr{A}$ of $\mathbb{R}^{\tilde{\mathscr{G}}}$ spanned by $\mathscr{B}$ is closed under the action of the infinitesimal generator $Q(\mathscr{R}, k)$. In particular, for each marked rule $\tilde{\alpha}\colon \tilde{L} \rightharpoonup \tilde{R}$ in $\mathscr{R}$ and marked graph $\tilde{G}$ in $\tilde{\mathscr{G}}$, we have*

$$Q_{\tilde{\alpha}}[\tilde{G}] = \sum_{\tilde{\mu} \in \tilde{R}*\tilde{G}} [\operatorname{cod}(\tilde{\alpha}^{\dagger}(\tilde{\mu}_1))] - \sum_{\tilde{\mu} \in \tilde{L}*\tilde{G}} [\operatorname{cod}(\tilde{\mu})].$$

Proof. Let $\tilde{F}$ be some marked graph in $\tilde{\mathscr{G}}$ and $\tilde{\alpha}\colon \tilde{L} \rightharpoonup \tilde{R}$ a rule in $\mathscr{R}$. By (2) and the definition of $Q_{\tilde{\alpha}}$, we have

$$\begin{aligned}
(Q_{\tilde{\alpha}}[\tilde{F}])(\tilde{G}) &= \sum_{\tilde{H} \in \tilde{\mathscr{G}}} q_{\tilde{G}\tilde{H}}^{\tilde{\alpha}}([\tilde{F}](\tilde{H}) - [\tilde{F}](\tilde{G})) \\
&= \sum_{\tilde{f} \in [\tilde{L}; \tilde{G}]} |[\tilde{F}; \operatorname{cod}(\tilde{\alpha}(\tilde{f}))]| - \sum_{\tilde{f} \in [\tilde{L}; \tilde{G}]} |[\tilde{F}; \tilde{G}]|.
\end{aligned}$$

Recall that $\tilde{\alpha}(\tilde{f})$ denotes the representative comatch derived from $\tilde{f}$ by $\tilde{\alpha}$, and hence $\operatorname{cod}(\tilde{\alpha}(\tilde{f}))$ is just the marked graph derived from $\tilde{G}$ via $\tilde{\alpha}$ and $\tilde{f}$. The action of $Q_{\tilde{\alpha}}$ at $\tilde{G}$ thus naturally decomposes into two terms $Q_{\tilde{\alpha}}^+$ and $Q_{\tilde{\alpha}}^-$ describing, respectively, the production and consumption of instances of $\tilde{F}$. By Lemma 3, the consumption term $Q_{\tilde{\alpha}}^-$ is equal to

$$Q_{\tilde{\alpha}}^-[\tilde{F}](\tilde{G}) = |[\tilde{L}; \tilde{G}]| \cdot |[\tilde{F}; \tilde{G}]| = \sum_{\tilde{\mu} \in \tilde{L}*\tilde{F}} |[\operatorname{cod}(\tilde{\mu}); \tilde{G}]|$$

which is a linear combination of a finite number of elements in $\mathscr{B}$.

Applying the same decomposition Lemma 3 to the production term $Q_{\tilde{\alpha}}^+$, we obtain a more complicated expression:

$$Q_{\tilde{\alpha}}^+[\tilde{F}](\tilde{G}) = \sum_{\tilde{\mu} \in \tilde{R}*\tilde{F}} \sum_{\tilde{f} \in [\tilde{L}; \tilde{G}]} |\{\tilde{u} \in [\operatorname{cod}(\tilde{\mu}); \operatorname{cod}(\tilde{\alpha}(\tilde{f}))] \mid \tilde{u} \circ \tilde{\mu}_1 = \tilde{\alpha}(\tilde{f})\}|.$$

To simplify this expression, we use properties of marked derivations seen in §3.2. First, recall that the relation $\Rightarrow_{\tilde{\alpha}}$ between marked matches under $\tilde{L}$ and $\tilde{R}$ has $\Rightarrow_{\tilde{\alpha}^\dagger}$ as its inverse. The first match $\tilde{\mu}_1$ of any MG $\tilde{\mu}$ in $\tilde{R} * \tilde{F}$ thus has a preimage $\tilde{\alpha}^\dagger(\tilde{\mu}_1)$ under $\Rightarrow_{\tilde{\alpha}}$, as well as an associated corule $\tilde{\mu}_1(\tilde{\alpha}^\dagger)$. Write $\tilde{\mu}_1^\dagger$ for $\tilde{\alpha}^\dagger(\tilde{\mu}_1)$, $\tilde{\alpha}_\mu^\dagger$ for $\tilde{\mu}_1(\tilde{\alpha}^\dagger)$, and $\tilde{U}$, $\tilde{U}^\dagger$ for the codomains of $\tilde{\mu}_1$, $\tilde{\mu}_1^\dagger$. Recall that $\tilde{\alpha}_\mu^\dagger : \tilde{U} \rightharpoonup \tilde{U}^\dagger$ is again a marked rule, and hence there is an associated map $\tilde{\alpha}_\mu^\dagger(-)$ between matches under $\tilde{U}$ and $\tilde{U}^\dagger$. As $\tilde{\alpha}(\tilde{f}) \Rightarrow_{\tilde{\alpha}^\dagger} \tilde{f}$, the fact that derivations split along factorizations of matches means that $\tilde{\alpha}_\mu^\dagger(-)$ restricts to a bijection

$$\{\tilde{u} \in [\tilde{U}; \mathrm{cod}(\tilde{\alpha}(\tilde{f}))] \mid \tilde{u} \circ \tilde{\mu}_1 = \tilde{\alpha}(\tilde{f})\} \quad \simeq \quad \{\tilde{v} \in [\tilde{U}^\dagger; \tilde{G}] \mid \tilde{v} \circ \tilde{\mu}_1^\dagger = \tilde{f}\}$$

which allows us to simplify our previous expression for $Q_{\tilde{\alpha}}^+[\tilde{F}]$ to

$$Q_{\tilde{\alpha}}^+[\tilde{F}](\tilde{G}) = \sum_{\tilde{\mu} \in \tilde{R} * \tilde{F}} \sum_{\tilde{f} \in [\tilde{L}; \tilde{G}]} |\{\tilde{u} \in [\tilde{U}; \mathrm{cod}(\tilde{\alpha}(\tilde{f}))] \mid \tilde{u} \circ \tilde{\mu}_1 = \tilde{\alpha}(\tilde{f})\}|$$

$$= \sum_{\tilde{\mu} \in \tilde{R} * \tilde{F}} \sum_{\tilde{f} \in [\tilde{L}; \tilde{G}]} |\{\tilde{v} \in [\tilde{U}^\dagger; \tilde{G}] \mid \tilde{v} \circ \mu_1^\dagger = \tilde{f}\}| = \sum_{\tilde{\mu} \in \tilde{R} * \tilde{F}} |[\tilde{U}^\dagger; \tilde{G}]|.$$

This is again a linear combination of a finite number of elements in $\mathscr{B}$, concluding the proof of Theorem 2.   □

Before we move on, a few remarks about the above theorem and its proof are in order. Firstly, the theorem is a statement about marked graph observables rather than individual derivations. Although the observables in question depend on the LHS and RHS of rules, the explicit dependency of (2) on the rate matrix $Q_{\tilde{\alpha}}$ has vanished along with the corresponding dependencies on derivations in the definition of $Q_{\tilde{\alpha}}$. This is made possible by two key insights from §3.2 and §3.3, namely *(i)* that rules internalize application conditions and hence every marked match of an LHS extends to a derivation, and *(ii)* that application conditions in rules can be combined algebraically, by means of minimal gluings, with those in marked graph observables.

Secondly, the proof makes use of reversibility of derivations in several places. Nevertheless, Theorem 2 also holds for CTMCs generated by a significant class of irreversible graph transformation systems. In particular, our approach extends to single pushout (SPO) and sesqui-pushout (SqPO) rewriting, both of which deal with irreversible derivations [6,15,31]. In both cases the rigidity constraints on (marked) rules need to be relaxed as they would otherwise force rules to be reversible. For Theorem 2 to hold, it is sufficient to restrict rules to spans of monos and impose a strongest postcondition on their RHS (note that the latter does not restrict the expressivity of rules). Although $\Rightarrow_{\tilde{\alpha}^\dagger}$ is no longer the inverse of $\Rightarrow_{\tilde{\alpha}}$ in this setting, it can still be used to split derivations "backwards" along factorizations of comatches, leading to a bijection argument akin to that in the above proof.

Combining Theorems 1 and 2, we obtain the moment semantics for any finite rule set $\mathscr{R}$ and associated rate map k. In particular, the expected value of a

marked graph observable $\tilde{G}$ will satisfy the following differential equation:

$$\frac{d}{dt}\,\mathbf{E}_p([\tilde{G}]) = \sum_{\tilde{\alpha}\in\mathscr{R}} k(\tilde{\alpha}) \sum_{\tilde{\mu}\in\tilde{R}(\tilde{\alpha})*\tilde{G}} \mathbf{E}_p([\mathrm{cod}(\tilde{\alpha}^{\dagger}(\tilde{\mu}_1))]) - \sum_{\tilde{\alpha}\in\mathscr{R}} k(\tilde{\alpha}) \sum_{\tilde{\mu}\in\tilde{L}(\tilde{\alpha})*\tilde{G}} \mathbf{E}_p([\mathrm{cod}(\tilde{\mu})])$$

(7)

where $\tilde{L}(\tilde{\alpha})$ and $\tilde{R}(\tilde{\alpha})$ denote, respectively, the LHS and RHS of the marked rule $\tilde{\alpha}$. Since, $\mathrm{poly}(\mathscr{A}) = \mathscr{A}$, ODEs for the higher moments can be generated by the exact same procedure.

The number of terms in (7) depends on the size of the relevant sets of left and right-hand MGs, which is worst-case exponential in the size of the graphs involved (Lemma 3). In practice, one often finds many pairs of *irrelevant* MGs, the terms of which cancel out exactly. This reduces the effective size of the equations but not the overall complexity of generating them.

Second, as said in §1.3, the repeated application of (7) will lead to an infinite expansion in general. In practice, the system of ODEs needs to be truncated. For concrete models, static analysis might help finding invariants in the underlying rewrite system and find a finite closure even for models where the set of reachable states is demonstrably infinite [10]. We have seen a simple example in §1.

5   Conclusion

Consider again the example of preferential attachment presented in the first two sections. In this case, we can automatically derive systems of ODEs that are finite; however, we have to cope with the combinatorial blow-up. This bring us to the most exciting direction for future work: mean field approximations of moment semantics. In the literature, one often finds graphical approximation techniques based on conditional independence assumptions to control the size of patterns used in observables, such as so-called pair approximation [13,21]. It is known that these methods can be more accurate than naive truncation of ODEs. In a natural next step, we would like to understand if and how these can be brought inside our formal approach.

References

1. The Preferential Attachment ODE Generator (2015). https://github.com/sstucki/pa-ode-gen/
2. Bapodra, M., Heckel, R.: From graph transformations to differential equations. ECEASST **30** (2010)
3. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)
4. Barabási, A.L., Albert, R., Jeong, H.: Mean-field theory for scale-free random networks. Physica A: Statistical Mechanics and its Applications **272**(1), 173–187 (1999)
5. Chaput, P., Danos, V., Panangaden, P., Plotkin, G.D.: Approximating Markov processes by averaging. Journal of the ACM **61**(1), 5 (2014)

6. Corradini, A., Heindel, T., Hermann, F., König, B.: Sesqui-pushout rewriting. In: Corradini, A., Ehrig, H., Montanari, U., Ribeiro, L., Rozenberg, G. (eds.) ICGT 2006. LNCS, vol. 4178, pp. 30–45. Springer, Heidelberg (2006)
7. Corradini, A., Montanari, U., Rossi, F., Ehrig, H., Heckel, R., Löwe, M.: Algebraic approaches to graph transformation. Part I: basic concepts and double pushout approach. In: Handbook of Graph Grammars and Computing by Graph Transformation, pp. 163–245 (1997)
8. Danos, V., Heindel, T., Honorato-Zimmer, R., Stucki, S.: Approximations for stochastic graph rewriting. In: Merz, S., Pang, J. (eds.) ICFEM 2014. LNCS, vol. 8829, pp. 1–10. Springer, Heidelberg (2014)
9. Danos, V., Honorato-Zimmer, R., Jaramillo-Riveri, S., Stucki, S.: Deriving rate equations for site graph rewriting systems. In: SASB (2013)
10. Danos, V., Honorato-Zimmer, R., Jaramillo-Riveri, S., Stucki, S.: Coarse-graining the dynamics of ideal branched polymers. In: Electronic Notes in Theoretical Computer Science, Workshop on Static Analysis and Systems Biology, SASB 2012, Deauville, pp. 47–64, April 2015
11. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of networks with aging of sites. Phys. Rev. **E 62**, 1842–1845 (2000)
12. Dorogovtsev, S.N., Mendes, J.F.F., Samukhin, A.N.: Structure of growing networks with preferential linking. Phys. Rev. Lett. **85**, 4633–4636 (2000)
13. Durrett, R., Gleeson, J.P., Lloyd, A.L., Mucha, P.J., Shi, F., Sivakoff, D., Socolar, J.E., Varghese, C.: Graph fission in an evolving voter model. Proceedings of the National Academy of Sciences **109**(10), 3682–3687 (2012)
14. Ehrig, H., Ehrig, K., Habel, A., Pennemann, K.H.: Theory of constraints and application conditions: From graphs to high-level structures. Fundamenta Informaticae **74**(1), 135–166 (2006)
15. Ehrig, H., Heckel, R., Korff, M., Löwe, M., Ribeiro, L., Wagner, A., Corradini, A.: Algebraic approaches to graph transformation. Part II: Single pushout approach and comparison with double pushout approach. In: Rozenberg, G. (ed.) Handbook of Graph Grammars and Computing by Graph Transformation, pp. 247–312. World Scientific, River Edge (1997)
16. Ehrig, H., Pfender, M., Schneider, H.J.: Graph-grammars: an algebraic approach. In: 14th Annual IEEE Symposium on Switching and Automata Theory, pp. 167–180 (1973)
17. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley (1986)
18. Evans, M.R., Ferrari, P.A., Mallick, K.: Matrix representation of the stationary measure for the multispecies TASEP. Journal of Statistical Physics **135**(2), 217–239 (2009)
19. Fages, F., Soliman, S.: Formal cell biology in Biocham. In: Bernardo, M., Degano, P., Zavattaro, G. (eds.) SFM 2008. LNCS, vol. 5016, pp. 54–80. Springer, Heidelberg (2008)
20. Feret, J., Danos, V., Harmer, R., Krivine, J., Fontana, W.: Internal coarse-graining of molecular systems. PNAS **106**(16), 6453–6458 (2009)
21. Gleeson, J.P.: High-accuracy approximation of binary-state dynamics on networks. Physical Review Letters **107**(6), 068701 (2011)
22. Harmer, R., Danos, V., Feret, J., Krivine, J., Fontana, W.: Intrinsic information carriers in combinatorial dynamical systems. Chaos **20**(3) (2010)
23. Hayman, J., Heindel, T.: Pattern graphs and rule-based models: the semantics of Kappa. In: Pfenning, F. (ed.) FOSSACS 2013 (ETAPS 2013). LNCS, vol. 7794, pp. 1–16. Springer, Heidelberg (2013)

24. Heckel, R.: DPO transformation with open maps. In: Ehrig, H., Engels, G., Kreowski, H.-J., Rozenberg, G. (eds.) ICGT 2012. LNCS, vol. 7562, pp. 203–217. Springer, Heidelberg (2012)
25. Heckel, R., Lajios, G., Menge, S.: Stochastic graph transformation systems. Fundam. Inform. **74**(1), 63–84 (2006)
26. Heckel, R., Wagner, A.: Ensuring consistency of conditional graph grammars - a constructive approach. Electronic Notes in Theoretical Computer Science **2**(0), 118–126 (1995)
27. van Kampen, N.: Stochastic processes in physics and chemistry, 3rd edition, North-Holland (2007)
28. Lack, S., Sobociński, P.: Adhesive categories. In: Walukiewicz, I. (ed.) FOSSACS 2004. LNCS, vol. 2987, pp. 273–288. Springer, Heidelberg (2004)
29. Lack, S., Sobocinski, P.: Adhesive and quasiadhesive categories. Theoretical Informatics and Applications **39**(2), 522–546 (2005)
30. Lopez, C.F., Muhlich, J.L., Bachman, J.A., Sorger, P.K.: Programming biological models in Python using PySB. Molecular Systems Biology **9**(1) (2013)
31. Löwe, M.: Algebraic Approach to Single-Pushout Graph Transformation. Theoretical Computer Science **109**(1&2), 181–224 (1993)
32. Lynch, J.F.: A logical characterization of individual-based models. In: 23rd Annual IEEE Symposium on Logic in Computer Science, LICS 2008, pp. 379–390. IEEE (2008)
33. Norris, J.R.: Markov chains. Cambridge series in statistical and probabilistic mathematics. Cambridge University Press (1998)
34. Shkarin, S.A.: Some results on solvability of ordinary linear differential equations in locally convex spaces. Mathematics of the USSR-Sbornik **71**(1), 29 (1992)
35. Stukalin, E.B., Phillips III, H., Kolomeisky, A.B.: Coupling of two motor proteins: a new motor can move faster. Physical Review Letters **94**(23), 238101 (2005)
36. Thomas, P., Matuschek, H., Grima, R.: Intrinsic noise analyzer: a software package for the exploration of stochastic biochemical kinetics using the system size expansion. PloS ONE **7**(6), e38518 (2012)