A compositional account of motifs, mechanisms, and dynamics in biochemical regulatory networks

Rebekah Aduddell[1], James P. Fairbanks[2], Amit Kumar[3], Pablo S. Ocal[4], Evan Patterson[5], and Brandon T. Shapiro[6]

[1]University of Texas at Arlington, Mathematics Department, Arlington, TX 76019, USA

[2]University of Florida, Computer & Information Science & Engineering, Gainesville, FL 32611, USA

[3]Louisiana State University, Mathematics Department, Baton Rouge, LA 70803, USA

[4]University of California, Los Angeles (UCLA), Mathematics Department, Los Angeles, CA 90095, USA

[5]Topos Institute, Berkeley, CA 94704, USA

[6]University of Virginia, Mathematics Department, Charlottesville, VA 22904, USA

Regulatory networks depict promoting or inhibiting interactions between molecules in a biochemical system. We introduce a category-theoretic formalism for regulatory networks, using signed graphs to model the networks and signed functors to describe occurrences of one network in another, especially occurrences of network motifs. With this foundation, we establish functorial mappings between regulatory networks and other mathematical models in biochemistry. We construct a functor from reaction networks, modeled as Petri nets with signed links, to regulatory networks, enabling us to precisely define when a reaction network could be a physical mechanism underlying a regulatory network. Turning to quantitative models, we associate a regulatory network with a Lotka-Volterra system of differential equations, defining a functor from the category of signed graphs to a category of parameterized dynamical systems. We extend this result from closed to open systems, demonstrating that Lotka-Volterra dynamics respects not only inclusions and collapsings of regulatory networks, but also the process of building up complex regulatory networks by gluing together simpler pieces. Formally, we use the theory of structured cospans to produce a lax double functor from the double category of open signed graphs to that of open parameterized dynamical systems. Throughout the paper, we ground the categorical formalism in examples inspired by systems biology.

1 Introduction

The genes, proteins, and RNA molecules that comprise living cells interact in complex, varied ways to sustain the cell throughout its lifecycle and respond to changes in its environment. Intensive experimental study of these interactions is distilled in an idealized form as *regulatory networks*, a kind of directed graph in which vertices represent molecules and edges represent interactions between molecules (Figure 1). The edges are labeled with a positive or negative sign according to whether the interaction is activating or inhibiting. Regulatory networks are the subject of a large body of experimental and theoretical work, notably reviewed by Alon [Alo07; Alo19] and

Rebekah Aduddell: rjaduddell@gmail.com

James P. Fairbanks: fairbanksj@ufl.edu

Amit Kumar: akuma25@lsu.edu

Pablo S. Ocal: socal@math.ucla.edu

Evan Patterson: evan@epatters.org

Brandon T. Shapiro: brandonshapiro@virginia.edu

Tyson et al. [TN10; TLK19] among others. Particular attention has been paid to *network motifs* [Alo07; TN10], the simple but functionally meaningful patterns that recur frequently in regulatory networks, and to various quantitative *dynamics* [TLK19] that can be assigned to the networks.

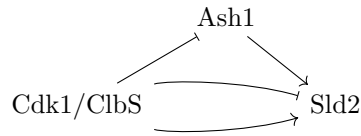Ash1

Cdk1/ClbS   Sld2

Figure 1: A small biochemical regulatory network: regulation of Sld2 by Cdk1 or ClbS with Ash1 as a predicted transcription factor. Adapted from Csikász-Nagy et al. [Csi+09, Figure 3C].

Although regulatory networks are simple enough to define mathematically—we shall define them to be directed graphs, possibly with multiple edges and loops, whose edges are assigned a positive or negative sign—important scientific concepts involving them, such as occurrences of motifs in networks and biochemical mechanisms generating networks, are often treated imprecisely. Likewise for relationships between regulatory networks and other mathematical models in biochemistry, particularly dynamical models based on ordinary or stochastic differential equations. Hence a first aim of this paper is to put certain concepts and relations concerning regulatory networks on a firm mathematical footing. To do so, we will use methods from category theory.

Category theory, in both the small and the large, is a natural tool for this study. In saying that a motif *occurs* in a network, one should allow for the possibility that the occurrence is indirect, involving a sequence of appropriately signed interactions. For example, positive autoregulation can occur directly but also indirectly through a double-negative feedback loop. Since a small category is nothing other than a graph in which consecutive edges can be composed, subject to certain laws, regulatory networks should be viewed not only as signed graphs (Section 2.1) but also as signed categories freely generated by those (Section 2.2). Sign-preserving functors, unlike sign-preserving graph homomorphisms, can express indirect occurrences and are in this sense a better notion of morphism for regulatory networks. Here we are doing category theory *in the small*, using categories as algebraic structures comparable to familiar ones like graphs, groups, and monoids.

Having laid these foundations for regulatory networks, we turn to category theory *in the large*, a mathematical theory of structure well suited to describe the passages between regulatory networks and other mathematical models of biochemical systems. Formally speaking, these passages are functors into or out of the category of regulatory networks. Making a functor is significantly stronger than making an objects-only mapping, as is typically done in the literature, since if morphisms of signed graphs formalize relationships between different regulatory networks, then functorality requires that these relationships be transported to or from other models of interest. By contrast, an objects-only mapping is, abstractly speaking, entirely unconstrained and so is capable of acting highly irregularly across different models of a given class. Functorality thus serves as a kind of safeguard for model transformation: it does not, on its own, ensure that a transformation makes good scientific sense but it does impose nontrivial logical constraints and coherences.

A first illustration of this principle is the connection between regulatory networks and biochemical reaction networks (Section 2.3). When modeling the complex biochemical systems that constitute a living cell, it is often practically necessary to abstract away certain details of the underlying chemical processes. Regulatory networks generally do not capture all the species or reactions involved in a given system; nor can they capture multispecies reactions faithfully because they describe only pairwise interactions. Given that regulatory networks are, to some degree, phenomenological models, it is natural to ask whether a given network could arise as a summary of a specific chemical process. The latter are described by *biochemical reaction networks*, graph-like structures allowing reactions or transitions with multiple inputs and outputs. Inspired by graphical syntax from systems biology [Voi00; Voi13], we formalize reaction networks as "Petri nets with links," and we construct a functor from the category of Petri nets with signed links to the category of signed graphs. This functor enables us to propose a formal definition for when a reaction network could be a *mechanism* for a regulatory network, a concept that is rarely if ever treated in a precise way.

This concludes the content of Section 2. In Section 3, we turn from qualitative to quantitative analysis, seeking a functorial assignment of continuous dynamics to regulatory networks. Although

rarely made explicitly functorial, systematic ways to formulate a model belonging to a mathematically homogeneous class of models are ubiquitous in science. Voit calls these "canonical representations" or "canonical models,"[1] and identifies Lotka-Volterra models and BST models/S-systems as two prominent examples in biology [Voi13, §3]. Reflecting their phenomenological status, regulatory networks do not admit a single, obvious dynamical interpretation, and so a wide variety of dynamical models have been considered, spanning the discrete and continuous, deterministic and stochastic [TLK19]. We consider Lotka-Volterra systems of ordinary differential equations. While not necessarily the most biologically plausible, Lotka-Volterra systems are among the simplest possible continuous models and so are a natural place to begin a functorial study.

A *Lotka-Volterra system* of equations has the form

$$\dot{x}_i = \rho_i\, x_i + \sum_{j=1}^n \beta_{i,j}\, x_i\, x_j, \qquad i = 1, \dots, n.$$

or equivalently, has logarithmic derivatives that are affine functions of the state variables:

$$\frac{d}{dt}[\log x_i(t)] = \rho_i + \sum_{j=1}^n \beta_{i,j}\, x_j(t), \qquad i = 1, \dots, n.$$

The coefficients ρ_i specify baseline rates of growth or decay, according to their sign, and the coefficients $\beta_{i,j}$ rates of activation or inhibition, according to their sign. We construct a functor that sends a signed graph (regulatory network) to a Lotka-Volterra model that constrains the signs of the rate coefficients (Section 3.2). As a prerequisite, we define a category of parameterized dynamical systems (Section 3.1), a construction of intrinsic interest that is by no means restricted to Lotka-Volterra dynamics. By working with signed graphs, rather than merely graphs, we ensure that scientific knowledge about whether interactions are promoting or inhibiting is reflected in both the syntax and the quantitative semantics.

In order to comprehend complex biological systems, we must decompose them into small, readily understandable pieces and then compose them back together to reproduce the behavior of the original system. This is the mantra of systems biology, which stresses that compositionality is no less important than reductionism in biology. With this motivation, a secondary aim of this paper is to extend the above constructions from closed systems to open ones, which can be composed together by gluing them along their interfaces. Mathematically, we pass from categories to *double categories*[2] [Gra19], two-dimensional categorical structures in which the usual morphisms of systems compose along one direction (by convention, the "vertical" one) and open systems compose along the other direction (the "horizontal" one). The double categories of open systems are further equipped with *monoidal products*, enabling systems to be composed not just in sequence but also in parallel. Among other results, we show that the Lotka-Volterra dynamics functor extends to a monoidal lax double functor from the monoidal double category of open signed graphs to that of open parameterized dynamical systems (Section 3.3).

The mathematics developed here is motivated by biochemistry but need not be restricted to it. Famously, Lotka-Volterra systems originated in ecology to model predator-prey dynamics [Lot25]. Regulatory networks and Lotka-Volterra systems can be used as generic models of entities that "regulate" each other in some manner, be it at the scale of individual cells or animal ecosystems. Regulatory networks are highly reminiscent of the *causal loop diagrams* in system dynamics [Ste00, Chapter 5], where the latter explicitly label feedback loops and their polarities.

The language of category theory is indispensable to this work but the level of knowledge assumed of the reader is not constant. We assume throughout that the reader is familiar with the basic notions of category theory, such as categories, functors, and natural transformations. Our main reference for facts about category theory is Riehl's text [Rie16], although there are many others. In the definitions and theorem statements, we have tried to minimize the technical level and explicate

---

[1]"Canonical" models in systems biology should not be confused with the unrelated, in fact incompatible, notion of "canonical" constructions in category theory.

[2]Early work on categorical systems theory, including on structured cospans [FS07], was based on *bicategories*. For technical reasons explained in [Pat23], it is increasingly common to use double categories instead, as in recent work on structured cospans [BC20]. This entails no loss since every double category has an underlying bicategory.

the ideas in concrete terms. In the proofs, we have aimed for efficiency and freely use concepts and results from the literature that do not appear in the main text. The reader can omit the proofs without disrupting the continuity of the paper.

2 Qualitative analysis: motifs and mechanisms

2.1 Regulatory networks as signed graphs

To begin, we clarify the notion of graph to be used throughout in this paper. The following definition is standard among category theorists. In other fields, it might be called a "directed multigraph," but we will call it simply a "graph."

Definition 2.1 (Graphs). The **schema for graphs** is the category $\mathsf{Sch}(\mathsf{Graph})$ freely generated by two parallel morphisms:

$$V \underset{\text{tgt}}{\overset{\text{src}}{\rightleftarrows}} E \;.$$

A **graph** is a functor $X : \mathsf{Sch}(\mathsf{Graph}) \to \mathsf{Set}$, also known as a *copresheaf* on $\mathsf{Sch}(\mathsf{Graph})$.[3] A **graph homomorphism** from a graph X to another graph Y is a natural transformation $\phi : X \to Y$. Graphs and graph homomorphisms form the category Graph.

To restate the definition in explicit terms, a graph X consists of

- a set $X(V)$ of **vertices**;
- a set $X(E)$ of **edges**; and
- functions $X(\text{src}), X(\text{tgt}) : X(E) \to X(V)$, assigning to each edge its **source** and **target**.

A graph homomorphism $\phi : X \to Y$ consists of a function $\phi_V : X(V) \to Y(V)$, the **vertex map**, and another function $\phi_E : X(E) \to Y(E)$, the **edge map**. These maps must preserve sources and targets, meaning that the following squares commute:

$$
\begin{array}{ccc}
X(E) & \xrightarrow{X(\text{src})} & X(V) \\
\phi_E \downarrow & & \downarrow \phi_V \\
Y(E) & \xrightarrow[Y(\text{src})]{} & Y(V)
\end{array}
\qquad
\begin{array}{ccc}
X(E) & \xrightarrow{X(\text{tgt})} & X(V) \\
\phi_E \downarrow & & \downarrow \phi_V \\
Y(E) & \xrightarrow[Y(\text{tgt})]{} & Y(V)
\end{array}
\;.
$$

We now turn to the main notion of this section, signed graph. Write Sgn for the set of (nonzero) signs, whose two elements may be denoted $\{1, -1\}$ or $\{+, -\}$. The set of signs is an abelian group, isomorphic to the cyclic group $\mathbb{Z}_2$, under the usual multiplication.

Definition 2.2 (Signed graphs). The **category of signed graphs** is the slice category

$$\mathsf{SgnGraph} := \mathsf{Graph}/\mathsf{Sgn},$$

where, by abuse of notation, Sgn is regarded as a graph with one vertex and two loops.

---

[3]Applied category theorists often take set-valued functors to be covariant (i.e., as copresheaves) rather than contravariant (i.e., as the more traditional presheaves), for reasons of convenience visible in works such as [Spi21].

Unpacking the definition, a **signed graph** is seen to be a graph X equipped with a function $X(\mathrm{sgn}) : X(E) \to \mathsf{Sgn}$ that assigns a sign to each edge. Given signed graphs X and Y, a **morphism of signed graphs** from X to Y is a graph homomorphism ϕ that preserves signs, meaning that the following triangle commutes:

$$
\begin{array}{ccc}
X(E) & \xrightarrow{\phi_E} & Y(E) \\
& \searrow^{X(\mathrm{sgn})} \quad \swarrow_{Y(\mathrm{sgn})} & \\
& \mathsf{Sgn} &
\end{array} \qquad .
$$

Signed graphs are a mathematical description of the regulatory networks studied in systems biology [Alo07; TN10]. For the purposes of this paper, we will simply define a **regulatory network** to be a signed graph. The vertices of the graph represent the components of the network, which could be proteins, genes, or RNA molecules. Signed edges represent interactions between components, where the source has the effect of either *activating/promoting* the target (positive sign) or *inhibiting/repressing* it (negative sign). As is customary, we denote activation interactions by arrows with pointed heads ($\longrightarrow$) and inhibition interactions by arrows with flat heads ($\longmapsto$). For instance, the two drawings

$$
\overset{+}{x \rightleftarrows y} \quad \overset{}{\underset{-}{}} \quad \longleftrightarrow \quad x \rightleftarrows y
$$
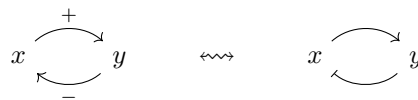
represent the same network, a negative feedback loop in which x activates y, which in turn inhibits x [TN10, Scheme 1, Motif B].

In the literature [TN10], regulatory networks are often represented mathematically as sign-valued matrices. This approach is a special case of ours in that an n-by-n matrix valued in $\{+1, -1, 0\}$ can be interpreted as a simple signed graph on n vertices, with signed edges defined by the nonzero matrix elements. Unlike the matricial formalism, our formalism allows multiple edges between the same pair of edges, which can model multiple interactions based on different mechanisms. Allowing multiple edges and self-loops also ensures that graphs and signed graphs form well behaved categories, as the following proposition shows.

Proposition 2.3. *The category of signed graphs is complete (has all limits) and cocomplete (has all colimits).*

Proof. Because Graph is a copresheaf category, it is complete and cocomplete [Rie16, Proposition 3.3.9]. The slice category $\mathsf{SgnGraph} = \mathsf{Graph}/\mathsf{Sgn}$ is hence also complete and cocomplete [Rie16, Proposition 3.5.5]; alternatively, this follows because slices of copresheaf categories are again (equivalent to) copresheaf categories [Str00, Remark p. 303]. □

A morphism of signed graphs can do two things. Most obviously, it can pick out a signed graph as a subobject of another one, via a sign-preserving subgraph embedding. A signed graph morphism can also collapse multiple vertices onto a single vertex, and multiple edges onto a single edge with the same sign, in the restrictive sense permitted by a graph homomorphism. To illustrate, consider the following morphism inspired by Alon's review [Alo07, Figure 5].

$$
\begin{array}{ccc}
\begin{array}{c}
\curvearrowright \\
\mathrm{argR} \\
\swarrow \downarrow \searrow \\
\mathrm{argCBH} \quad \mathrm{argD} \quad \mathrm{argE} \quad \mathrm{argF} \quad \mathrm{argI}
\end{array}
& \longrightarrow &
\begin{array}{c}
\curvearrowright \\
\mathrm{argR} \\
\downarrow \\
\mathrm{arg*}
\end{array}
\end{array} \qquad (1)
$$

The network in the domain is a "single-input module" in the arginine biosynthesis system, in which the regulator argR represses five different enzymes (argCHB, argD, etc.) involved in producing arginine. The morphism above forgets the distinction between these enzymes, collapsing them into a catch-all entity labeled "arg*". These two functions—embedding and collapsing—are *all* that a signed graph morphism can do. More precisely, any morphism of signed graphs factors

essentially uniquely as an epimorphism (morphism with surjective vertex and edge maps) followed by a monomorphism (morphism with injective vertex and edge maps), using the epi-mono factorization available in any copresheaf category, or more generally in any topos [MM94, §IV.6]. In the next section we will consider a more flexible notion of morphism between signed graphs.

Colimits of signed graphs can be used to construct a category, or rather a double category, of *open* signed graphs. Composition of open signed graphs formalizes the process of building large regulatory networks from smaller pieces, including network motifs.

Proposition 2.4 (Open signed graphs). *There is a symmetric monoidal double category of open signed graphs, $\mathbb{O}\text{pen}(\mathsf{SgnGraph})$, having*

- *as objects, sets $A, B, C, \ldots$;*
- *as vertical morphisms, functions $f : A \to B$;*
- *as horizontal morphisms, **open signed graphs**, which consist of a signed graph X together with a cospan of sets $A_0 \xrightarrow{\ell_0} X(V) \xleftarrow{\ell_1} A_1$;*
- *as cells, **morphisms of open signed graphs** $(X, \ell_0, \ell_1) \to (Y, m_0, m_1)$, which consist of a map of signed graphs $\phi : X \to Y$ along with functions $f_i : A_i \to B_i$, $i = 0, 1$, making the following diagram commute:*

$$
\begin{array}{ccccc}
A_0 & \xrightarrow{\ell_0} & X(V) & \xleftarrow{\ell_1} & A_1 \\
{\scriptstyle f_0}\downarrow & & {\scriptstyle \phi_V}\downarrow & & \downarrow{\scriptstyle f_1} \\
B_0 & \xrightarrow{m_0} & Y(V) & \xleftarrow{m_1} & B_1
\end{array} \;\cdot
$$

Vertical composition is by composition in Set and in $\mathsf{SgnGraph}$. Horizontal composition and monoidal products are by pushouts and coproducts in $\mathsf{SgnGraph}$, respectively, viewing the sets in the feet of the cospans as discrete signed graphs.

Proof. To construct this symmetric monoidal double category, we use the method of structured cospans [FS07] in its double-categorical form [BC20]. The categories of sets and of signed graphs are related by an adjoint pair of functors

$$
\mathsf{Set} \underset{\mathrm{ev}_V}{\overset{\mathrm{Disc}}{\rightleftarrows}} \mathsf{SgnGraph} \;.
$$

Here $\mathrm{ev}_V : \mathsf{SgnGraph} \to \mathsf{Set}$ is the **evaluation at V** functor, sending a signed graph X to its set of vertices $X(V)$ and a morphism of signed graphs ϕ to its vertex map ϕ_V, and $\mathrm{Disc} : \mathsf{Set} \to \mathsf{SgnGraph}$ is the **discrete signed graph** functor, sending a set A to the signed graph with vertex set A and no edges. We obtain a symmetric monoidal double category of open signed graphs as the L-structured cospans for the functor $L \coloneqq \mathrm{Disc} : \mathsf{Set} \to \mathsf{SgnGraph}$ [BC20, Theorems 2.3 and 3.9].

To show that this symmetric monoidal double category is the same one in the proposition statement, suppose that $L \dashv R : \mathsf{A} \to \mathsf{X}$ is an adjoint pair of functors, where in our application $L = \mathrm{Disc}$ and $R = \mathrm{ev}_V$. By the defining bijection of an adjunction, L-structured cospans, i.e., objects A_0 and A_1 in A together with a cospan $L(A_0) \to X \leftarrow L(A_1)$ in X, correspond exactly to "R-decorated cospans," i.e., an object X in X together with a cospan $A_0 \to R(X) \leftarrow A_1$ in A. Furthermore, by the naturality of this bijection [Rie16, Lemma 4.1.3], morphisms of L-structured and R-decorated cospans

$$
\begin{array}{ccccc}
L(A_0) & \longrightarrow & X & \longleftarrow & L(A_1) \\
{\scriptstyle L(f_0)}\downarrow & & {\scriptstyle \phi}\downarrow & & \downarrow{\scriptstyle L(f_1)} \\
L(B_0) & \longrightarrow & Y & \longleftarrow & L(B_1)
\end{array}
\qquad \longleftrightarrow \qquad
\begin{array}{ccccc}
A_0 & \longrightarrow & R(X) & \longleftarrow & A_1 \\
{\scriptstyle f_0}\downarrow & & {\scriptstyle R(\phi)}\downarrow & & \downarrow{\scriptstyle f_1} \\
B_0 & \longrightarrow & R(Y) & \longleftarrow & B_1
\end{array}
$$

related by the adjunction are equivalent in that one diagram commutes if and only if the other does. We will tacitly reuse this reasoning in future constructions, such as Proposition 2.10 below. □

Let us illustrate compositionality as a means of building larger regulatory networks from smaller ones. The following example is adapted from Keurentjes et al. [Keu+11, Figure 1], later reproduced in [Voi18, Figure 1.7].

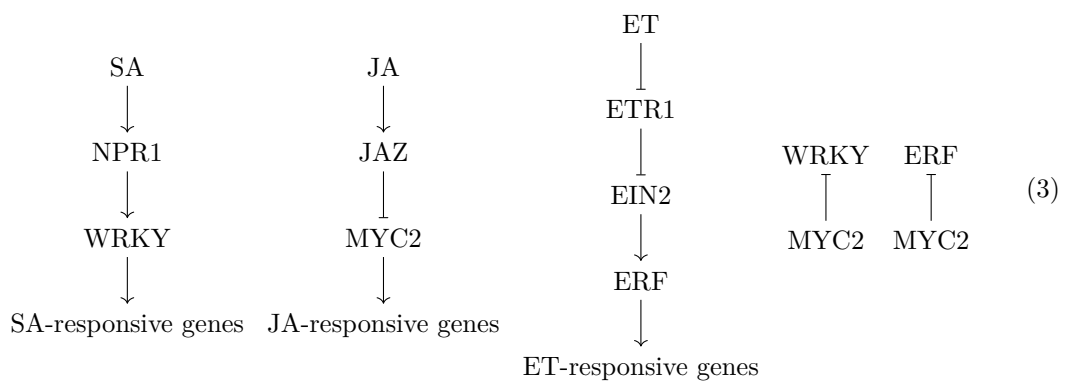*Example* 2.5 (Stress response system in plants)*.* When a plant perceives stress factors, it has three main biotic responses, called salicylic acid (SA) signaling, jasmonic acid (JA) signaling, and ethylene (ET) signaling. These processes promote the transcription factors WRKY, MYC2, and ERF, respectively, which in turn activate the genes responsible for responding to SA, JA, and ET. As it appears in [Keu+11, Figure 1] and [Voi18, Figure 1.7], the regulatory network governing these biotic responses

$$
\begin{array}{c}
\text{SA} \qquad \text{JA} \qquad \text{ET} \\
\downarrow \qquad \downarrow \qquad \downarrow \\
\text{ETR1} \\
\text{NPR1} \quad \text{JAZ} \quad \text{EIN2} \\
\downarrow \qquad \downarrow \qquad \downarrow \\
\text{WRKY} \dashleftarrow \text{MYC2} \dashrightarrow \text{ERF} \\
\downarrow \qquad \downarrow \qquad \downarrow \\
\text{SA-responsive genes} \quad \text{JA-responsive genes} \quad \text{ET-responsive genes}
\end{array} \tag{2}
$$
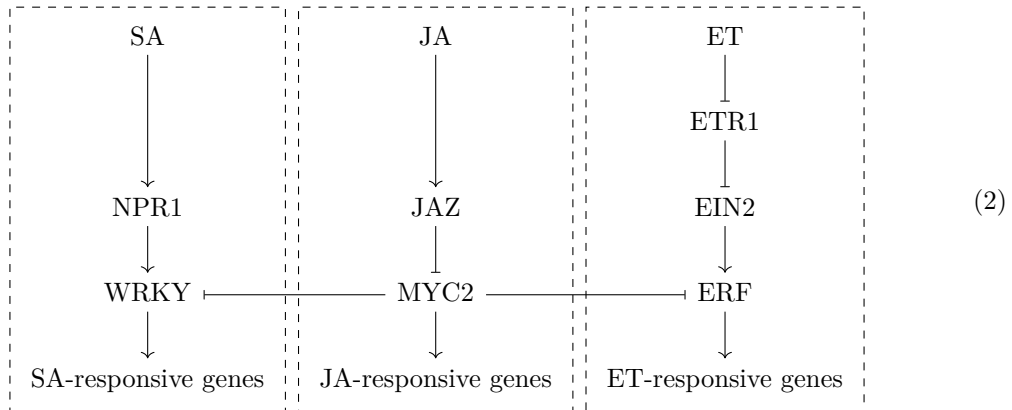
has five discernible subsystems: three signaling processes involving SA, JA, and ET, indicated by the dashed boxes, and two interactions *between* these processes, namely the inhibitions of WRKY and ERF by MYC2 that constitute SA-JA and JA-ET interactions. These subsystems and their functions were identified empirically, and it is known that they interact in the prescribed manner.

Following the decomposition identified by the biologists, we build up the overall system from smaller subsystems. Namely, we identify five subsystems and compose them from left to right. We could equally well have identified other subsystems or done the compositions in a different order, yielding an equivalent system. Horizontal associativity in the double category $\mathbb{O}$pen($\mathsf{SgnGraph}$), constructed in Proposition 2.4, ensures that any order of composition yields the same result, up to isomorphism.

We choose to separate out the five subsystems into small, individual regulatory networks:

$$
\begin{array}{c}
\text{ET} \\
\downarrow \\
\text{SA} \qquad \text{JA} \qquad \text{ETR1} \\
\downarrow \qquad \downarrow \qquad \downarrow \\
\text{NPR1} \quad \text{JAZ} \quad \text{EIN2} \quad \text{WRKY} \quad \text{ERF} \\
\downarrow \qquad \downarrow \qquad \downarrow \qquad \dashv \qquad \dashv \\
\text{WRKY} \quad \text{MYC2} \quad \text{ERF} \quad \text{MYC2} \quad \text{MYC2} \\
\downarrow \qquad \downarrow \qquad \downarrow \\
\text{SA-responsive genes} \quad \text{JA-responsive genes} \quad \text{ET-responsive genes}
\end{array} \tag{3}
$$

To turn these into *open* regulatory networks, as defined in Proposition 2.10 above, we will regard

the SA signaling subsystem as having no inputs and one output as follows.

$$
\begin{array}{c}
\text{SA} \\
\downarrow \\
\text{NPR1} \\
\downarrow \\
\text{WRKY} \longleftarrow \longleftarrow \bullet \\
\downarrow \\
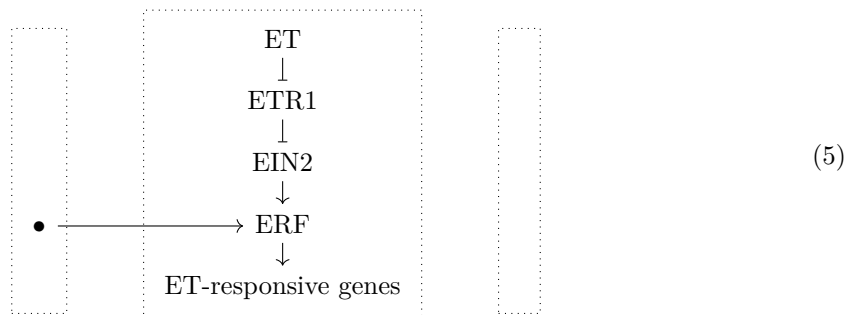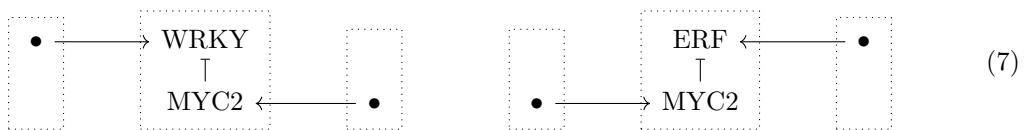\text{SA-responsive genes}
\end{array}
\tag{4}
$$
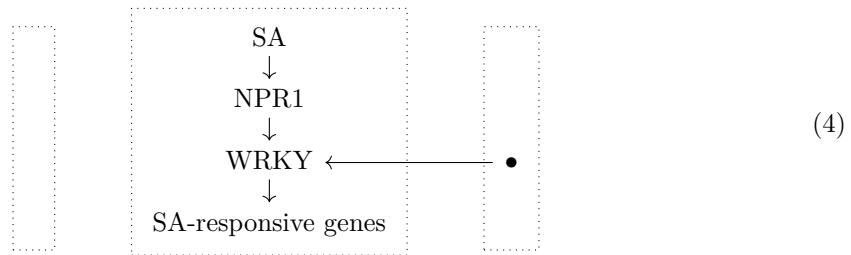
Similarly, ET signaling will admit one input and no outputs.

$$
\begin{array}{c}
\text{ET} \\
\perp \\
\text{ETR1} \\
\perp \\
\text{EIN2} \\
\downarrow \\
\bullet \longrightarrow \text{ERF} \\
\downarrow \\
\text{ET-responsive genes}
\end{array}
\tag{5}
$$

JA signaling will admit one input and one output.

$$
\begin{array}{c}
\text{JA} \\
\downarrow \\
\text{JAZ} \\
\perp \\
\bullet \longrightarrow \text{MYC2} \longleftarrow \bullet \\
\downarrow \\
\text{JA-responsive genes}
\end{array}
\tag{6}
$$
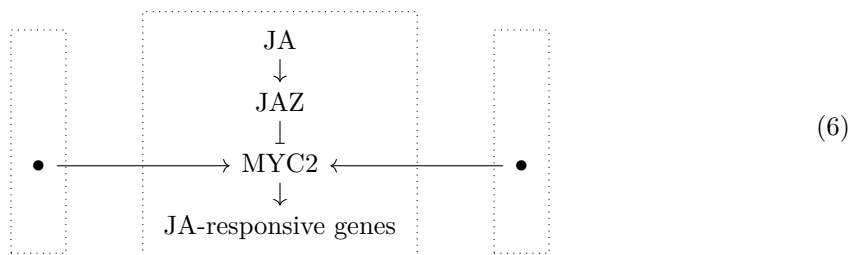
Finally, the inhibition of WRKY by MYC2 will have one input to WRKY and one output from MYC2, whereas the inhibition of ERF by MYC2 will have one output from ERF and one input from MYC2.

$$
\begin{array}{cc}
\bullet \longrightarrow \text{WRKY} & \text{ERF} \longleftarrow \bullet \\
\perp & \perp \\
\text{MYC2} \longleftarrow \bullet & \bullet \longrightarrow \text{MYC2}
\end{array}
\tag{7}
$$

The only thing left to do is compose (from left to right) these open regulatory nets. Starting with Equation (6), we compose it with the left part of Equation (7) by first putting them side by side and identifying the intermediate sets

$$
\begin{array}{c}
\text{SA} \\
\downarrow \\
\text{NPR1} \\
\downarrow \\
\text{WRKY} \longleftarrow \bullet \longrightarrow \text{WRKY} \\
\downarrow \qquad\qquad\qquad \perp \\
\text{SA-responsive genes} \qquad \text{MYC2} \longleftarrow \bullet
\end{array}
\tag{8}
$$
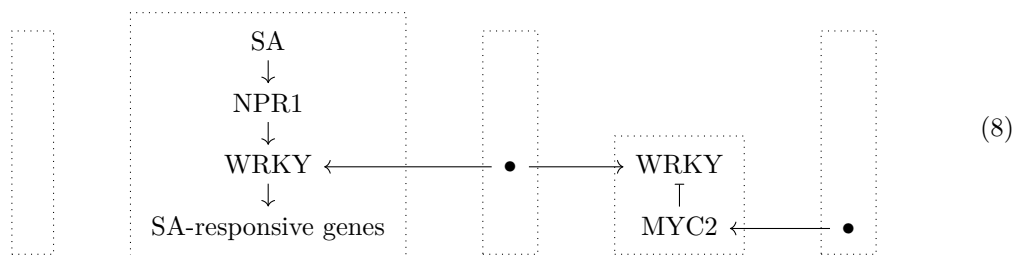
which we then remove, together with its outgoing arrows, while identifying the vertexes they

connect.

$$
\begin{array}{c}
\text{SA} \\
\downarrow \\
\text{NPR1} \\
\searrow \\
\text{WRKY} \\
\swarrow \qquad \nwarrow \\
\text{SA-responsive genes} \qquad \text{MYC2} \leftarrow \bullet
\end{array}
\tag{9}
$$
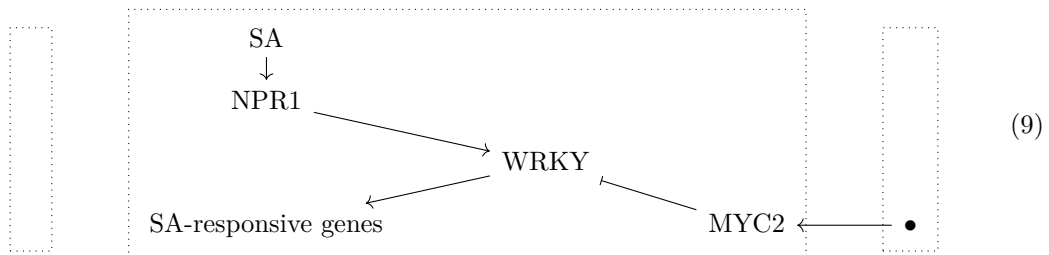
Rearranging the picture slightly for a better visualization, we obtain the open regulatory network that results from the composition:

$$
\begin{array}{c}
\text{SA} \\
\downarrow \\
\text{NPR1} \\
\downarrow \\
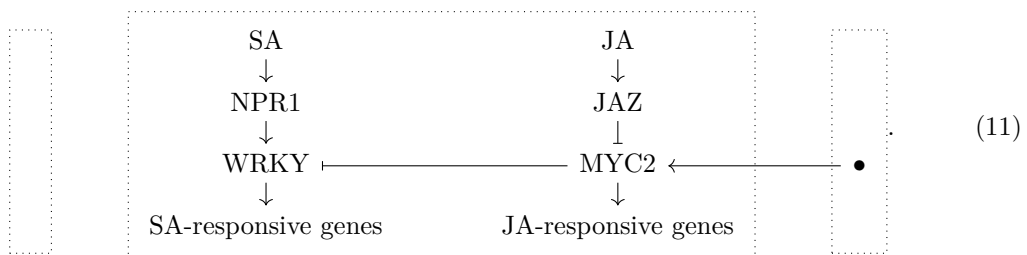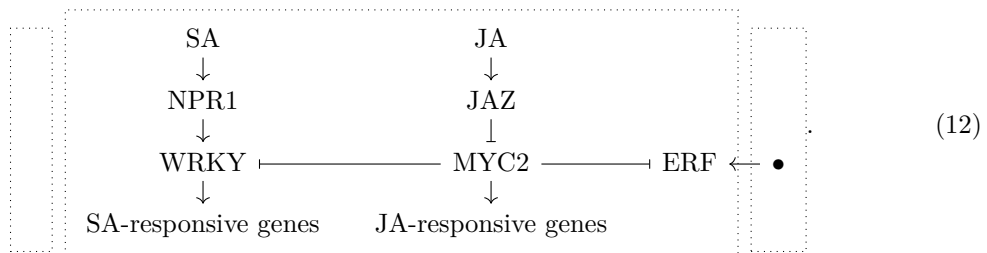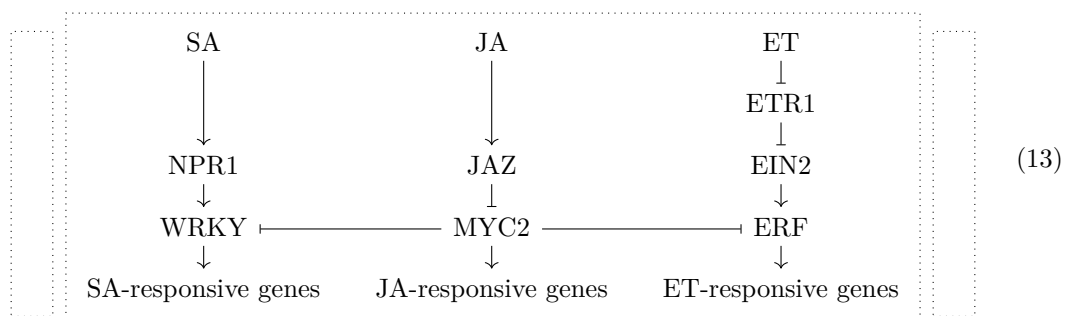\text{WRKY} \longmapsto \text{MYC2} \leftarrow \bullet \\
\downarrow \\
\text{SA-responsive genes}
\end{array} \; .
\tag{10}
$$
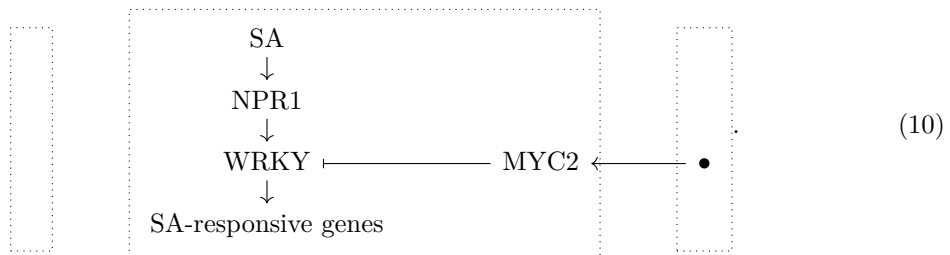
Next, composing Equations (6) and (10) yields

$$
\begin{array}{cc}
\text{SA} & \text{JA} \\
\downarrow & \downarrow \\
\text{NPR1} & \text{JAZ} \\
\downarrow & \dashv \\
\text{WRKY} \longmapsto & \text{MYC2} \leftarrow \bullet \\
\downarrow & \downarrow \\
\text{SA-responsive genes} & \text{JA-responsive genes}
\end{array} \; .
\tag{11}
$$

Then composing Equation (11) with the right part of Equation (7) yields

$$
\begin{array}{ccc}
\text{SA} & \text{JA} & \\
\downarrow & \downarrow & \\
\text{NPR1} & \text{JAZ} & \\
\downarrow & \dashv & \\
\text{WRKY} \longmapsto & \text{MYC2} \longrightarrow \dashv \text{ERF} \leftarrow \bullet \\
\downarrow & \downarrow & \\
\text{SA-responsive genes} & \text{JA-responsive genes} &
\end{array} \; .
\tag{12}
$$

Finally, composing Equations (5) and (13) yields

$$
\begin{array}{ccc}
\text{SA} & \text{JA} & \text{ET} \\
& & \downarrow \\
& & \text{ETR1} \\
& & \dashv \\
\text{NPR1} & \text{JAZ} & \text{EIN2} \\
\downarrow & \dashv & \downarrow \\
\text{WRKY} \longmapsto & \text{MYC2} \longrightarrow \dashv & \text{ERF} \\
\downarrow & \downarrow & \downarrow \\
\text{SA-responsive genes} & \text{JA-responsive genes} & \text{ET-responsive genes}
\end{array}
\tag{13}
$$

which corresponds to the open regulatory network governing the biotic responses (cf. Equation (2)). As mentioned at the beginning of the example, although we made a choice of subsystems and a choice in the order of compositions, any other choice will give an isomorphic result by the horizontal associativity of the double category $\mathbb{O}\mathrm{pen}(\mathsf{SgnGraph})$.

2.2 Refining regulatory networks using signed categories and functors

While morphisms of signed graph have their uses, they do not capture the important idea of *refining* regulatory networks, in which an interaction in one network is realized as a composite of several interactions in another. To express refinement, we must generalize our notion of morphism from graph homomorphisms to functors. This, in turn, requires the concept of a *signed category*.

Definition 2.6 (Signed categories)**.** The **category of signed categories** is the slice category

$$\mathsf{SgnCat} \coloneqq \mathsf{Cat}/\mathsf{Sgn},$$

where Cat is the category of small categories and the group of signs, Sgn, is regarded as a category with one object and two morphisms.

Unpacking the definition, a **signed category** is a category C in which every morphism f is assigned a sign $\mathrm{sgn}(f) \in \{1, -1\}$ in a functorial way, meaning that

$$\mathrm{sgn}(x_0 \xrightarrow{f_1} x_1 \xrightarrow{f_2} \cdots \xrightarrow{f_n} x_n) = \prod_{i=1}^{n} \mathrm{sgn}(f_i)$$

for every $n \geq 0$ and every sequence of composable morphisms $f_1, \ldots, f_n$. In particular ($n = 0$), the identity morphisms have positive sign. A **morphism of signed categories**, or **signed functor**, is a functor $F : \mathsf{C} \to \mathsf{D}$ between signed categories that preserves the signs, meaning that

$$\mathrm{sgn}_{\mathsf{D}}(F(f)) = \mathrm{sgn}_{\mathsf{C}}(f)$$

for every morphism f in C.

Since our aim is to have a more flexible notion of morphism between signed graphs, we will mostly restrict ourselves to those signed categories that are freely generated by a signed graph. The **free signed category** or **signed path category** functor

$$\mathrm{Path} : \mathsf{SgnGraph} \to \mathsf{SgnCat}$$

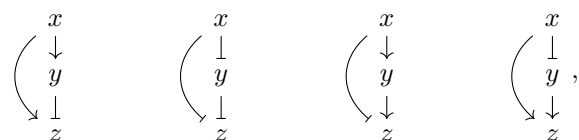sends a signed graph X to the signed category $\mathrm{Path}(X)$ having

- as objects, the vertices of X;
- as morphisms from x to y, the paths in X from x to y, whose sign is defined to be the product of the signs of the edges comprising the path.

Composition of paths is by concatenation, which clearly preserves the sign. The identity morphism at x is the empty path at x, which has positive sign. The functor Path on signed graphs is completely analogous to the usual free category functor on graphs, and as such is a left adjoint to the forgetful functor from signed categories to signed graphs. Signed categories are likewise algebras for the corresponding monad on the category $\mathsf{SgnGraph}$.

By convention, if X and Y are signed graphs, we say that a **signed functor** from X to Y is a signed functor $F : \mathrm{Path}(X) \to \mathrm{Path}(Y)$ between the corresponding signed path categories. Since the morphisms of $\mathrm{Path}(X)$ are freely generated by the edges in X, a signed functor from X to Y is uniquely determined by a morphism of signed graphs from X to the underlying signed graph of $\mathrm{Path}(Y)$. This means that each edge in X is sent to an appropriately signed *path* of edges in Y, which can be regarded as a refinement of the relationship that the edge represents.

Definition 2.7 (Category of refinements)**.** The category $\mathsf{SgnGraph}_{\mathrm{Path}}$ has as objects signed graphs and as morphisms signed functors between them; in other words, it is the Kleisli category for the free signed category monad on $\mathsf{SgnGraph}$.

We now have a precise language with which to classify network motifs and their occurrences. As a first example, Alon identifies four types of *incoherent feedforward loop (FFL)* involving three components,

$$
\left(
\begin{array}{c}
x \\ \downarrow \\ y \\ \dashv \\ z
\end{array}
\right)
\quad
\left(
\begin{array}{c}
x \\ \dashv \\ y \\ \dashv \\ z
\end{array}
\right)
\quad
\left(
\begin{array}{c}
x \\ \downarrow \\ y \\ \downarrow \\ z
\end{array}
\right)
\quad
\left(
\begin{array}{c}
x \\ \dashv \\ y \\ \downarrow \\ z
\end{array}
\right),
$$

Motif	Generic instance
Positive autoregulation	$L_+ := \left\{\, \bullet\, \circlearrowright\, \right\}$
Negative autoregulation	$L_- := \left\{\, \bullet\, \circlearrowleft\, \right\}$
Coherent feedforward loop	$I_{++} := \left\{\, \bullet \rightrightarrows \bullet\, \right\}$
Incoherent feedforward loop	$I_\pm := \left\{\, \bullet \rightrightarrows \bullet\, \right\}$
Positive feedback loop	$L_{++} := \left\{\, \bullet \rightleftarrows \bullet\, \right\}$
Negative feedback loop	$L_\pm := \left\{\, \bullet \rightleftarrows \bullet\, \right\}$
Double-negative feedback loop	$L_{--} := \left\{\, \bullet \rightleftarrows \bullet\, \right\}$

Table 1: Common motifs in biochemical regulation networks [Alo07; TN10]

those of *type 1, 2, 3, and 4*, respectively [Alo07, Figure 2a]. Besides having three components, what these motifs have in common is that there exists a signed functor into each of them from the signed graph $I_\pm := \left\{\, \bullet \rightrightarrows \bullet\, \right\}$ having two parallel arrows of opposite sign. The network $I_\pm$ is thus the "generic" incoherent feedforward loop, in the sense that signed functors out of it refine the pattern in specific ways. A similar situation holds for other common network motifs (Table 1), which motivates the following definition.

Definition 2.8 (Motif instance). Given a signed graph A, regarded as a motif, an **instance** or **occurrence** of the motif A in a network X is a monic signed functor $A \rightarrowtail X$.

Note that a signed functor is a monomorphism exactly when the functor is an embedding of categories, i.e., an injective-on-objects, faithful functor. Requiring the functor in the definition to be monic excludes "degenerate instances" of motifs where vertices or edges are identified.

Now, should the incoherent FFL be regarded as a network motif, or is it the more specific types, such as the incoherent FFL of type 1, that are motifs? From our point of view, they are all equally motifs but they have different degrees of specificity, and the functorial language clarifies how motifs are iteratively refined. Specifically, an instance of an incoherent FFL of type 1 in a network X also gives an instance of an incoherent FFL in X (of unspecified type), simply by composing the monomorphisms involved:

$$I_\pm \cong \left\{ x \rightrightarrows z \right\} \quad \rightarrowtail \quad \left\{ x \xrightarrow{} y \xrightarrow{} z \right\} \quad \rightarrowtail \quad X.$$

Similarly, in the notation of Table 1, any instance of double-negative feedback (L_{--}) also gives an instance of positive autoregulation (L_+) [CP09], via the monomorphism $L_+ \rightarrowtail L_{--}$ that sends the positive loop to the double-negative 2-cycle.

For any choice of motif A, the mapping that sends a regulatory network X to the set of occurrences of A in X is a representable functor

$$\mathrm{Hom}(A, -) : (\mathsf{SgnGraph}_{\mathrm{Path}})_m \to \mathsf{Set},$$

where $(\mathsf{SgnGraph}_{\mathrm{Path}})_m$ denotes the wide subcategory of monomorphisms in $\mathsf{SgnGraph}_{\mathrm{Path}}$. The functorality implies that, for any motif A, a monomorphism between regulatory networks induces a map between instances of A in those networks.

We now extend the construction of open signed graphs to open signed categories.

Proposition 2.9. *The category of signed categories is complete and cocomplete.*

Proof. Because the category Cat is complete and cocomplete [Rie16, Proposition 3.5.6], its slice $\mathsf{SgnCat} = \mathsf{Cat}/\mathsf{Sgn}$ is also [Rie16, Proposition 3.5.5]. $\qquad\square$

Proposition 2.10 (Open signed categories). *There is a symmetric monoidal double category of open signed categories, $\mathbb{O}\mathsf{pen}(\mathsf{SgnCat})$, having*

- *as objects, sets $A, B, C, \ldots$;*
- *as vertical morphisms, functions $f : A \to B$;*
- *as horizontal morphisms, **open signed categories**, which consist of a signed category C together with a cospan of sets $A_0 \xrightarrow{\ell_0} \mathrm{Ob}(\mathsf{C}) \xleftarrow{\ell_1} A_1$;*
- *as cells, **morphisms of open signed categories** $(\mathsf{C}, \ell_0, \ell_1) \to (\mathsf{D}, m_0, m_1)$, which consist of a signed functor $F : \mathsf{C} \to \mathsf{D}$ along with functions $f_i : A_i \to B_i$, $i = 0, 1$, making the diagram commute:*

$$
\begin{array}{ccc}
A_0 & \xrightarrow{\ell_0} \mathrm{Ob}(\mathsf{C}) \xleftarrow{\ell_1} & A_1 \\
f_0 \downarrow & \mathrm{Ob}(F) \downarrow & \downarrow f_1 \\
B_0 & \xrightarrow[m_0]{} \mathrm{Ob}(\mathsf{D}) \xleftarrow[m_1]{} & B_1
\end{array} \cdot
$$

Vertical composition is by composition in Set and in SgnCat. Horizontal composition and monoidal products are by pushouts and coproducts in SgnCat, respectively, viewing the sets in the feet of cospans as discrete signed categories.

Moreover, the signed path category functor extends to a symmetric monoidal double functor

$$\mathrm{Path} : \mathbb{O}\mathsf{pen}(\mathsf{SgnGraph}) \to \mathbb{O}\mathsf{pen}(\mathsf{SgnCat}).$$

Proof. We take $\mathbb{O}\mathsf{pen}(\mathsf{SgnCat})$ to be the symmetric monoidal double category of L'-structured cospans for the functor $L' := \mathrm{Disc} : \mathsf{Set} \to \mathsf{SgnCat}$ involved the composite adjunction

$$
\mathsf{Set} \underset{\mathrm{Ob}}{\overset{\mathrm{Disc}}{\rightleftarrows}} \perp \mathsf{SgnCat}
\qquad = \qquad
\mathsf{Set} \underset{\mathrm{ev}_V}{\overset{\mathrm{Disc}}{\rightleftarrows}} \perp \mathsf{SgnGraph} \underset{U}{\overset{\mathrm{Path}}{\rightleftarrows}} \perp \mathsf{SgnCat} \ .
$$

On the right hand side, the first adjunction was already used in the proof of Proposition 2.4, and the second adjunction is the free-forgetful adjunction between signed graphs and signed categories.

To prove the last statement, we notice that all functors involved in the commutative square

$$
\begin{array}{ccc}
\mathsf{Set} & \xrightarrow{L = \mathrm{Disc}} & \mathsf{SgnGraph} \\
\| & & \downarrow \mathrm{Path} \\
\mathsf{Set} & \xrightarrow[L' = \mathrm{Disc}]{} & \mathsf{SgnCat}
\end{array}
$$

are left adjoints, hence preserve colimits. We can therefore appeal to [BC20, Theorem 4.3] to obtain a symmetric monoidal double functor

$$\mathbb{O}\mathsf{pen}(\mathsf{SgnGraph}) \cong {}_L\mathbb{C}\mathsf{sp}(\mathsf{SgnGraph}) \to {}_{L'}\mathbb{C}\mathsf{sp}(\mathsf{SgnCat}) \cong \mathbb{O}\mathsf{pen}(\mathsf{SgnCat}). \qquad \square$$

2.3 Mechanistic models as Petri nets with links

However challenging they may be to identify through experiments and data analysis, regulatory networks still only summarize how the components of a complex biochemical system interact. Regulatory networks typically include only a subset of the system's components, and they do not model individual reactions and processes, only pairwise promoting or inhibiting interactions between components. In this sense, regulatory networks are not fully mechanistic models, even if they have a stronger causal interpretation than, say, a correlation matrix.

By contrast, mechanistic models in biochemistry model individual reactions, which requires a different formalism. Pictures like the following, adapted from Voit's review [Voi13, Figure 4], are common in systems biology.

$$
\begin{array}{c}
\longrightarrow A \longrightarrow B \longrightarrow D \longrightarrow \\
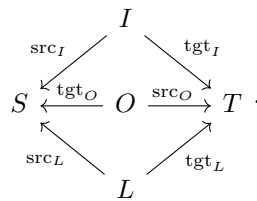\longrightarrow C
\end{array}
\tag{14}
$$
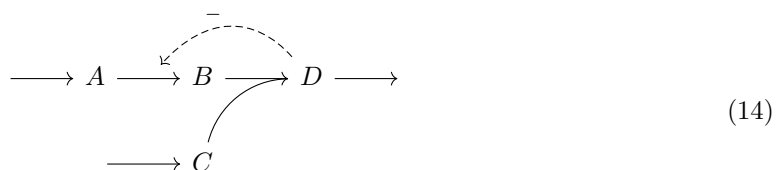
This diagram possesses two distinctive features. First, directed hyperedges represent reactions having a number of inputs or outputs different than one. There are, for example, hyperedges from B *and* C to D, from nothing to A (an inflow), and from D to nothing (an outflow). If, in lieu of hyperedges, we introduce a second type of vertex, we obtain a structure similar to a Petri net

$$(15)$$
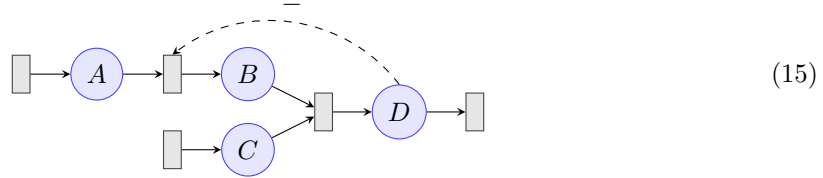
but with the second distinctive feature of having *signed links* from the first type of vertices (*species*) to the second type (*transitions*), whose signs indicate promotion or inhibition.

In this section, we explain how a Petri net with signed links can provide a *mechanism* for a regulatory network. This involves constructing a functor from Petri nets with signed links to signed graphs, approximating the former as the latter. As a prerequisite, we need a precise definition of a Petri net with links.

Definition 2.11 (Petri net with links)**.** The **schema for Petri nets with links** is the category $\mathsf{Sch}(\mathsf{LPetri})$ freely generated by these objects and morphisms:

$$
\begin{array}{ccc}
 & I & \\
\mathrm{src}_I \nearrow & & \searrow \mathrm{tgt}_I \\
S \xleftarrow{\mathrm{tgt}_O} O \xrightarrow{\mathrm{src}_O} T & & \\
\mathrm{src}_L \searrow & & \nearrow \mathrm{tgt}_L \\
 & L &
\end{array} \cdot
$$

A **Petri nets with links** is a functor $P : \mathsf{Sch}(\mathsf{LPetri}) \to \mathsf{Set}$, and a **morphism** of these is a natural transformation. A morphism $\phi : P \to Q$ **preserves arities** if the naturality squares associated with the morphisms $I \to T$ and $O \to T$ are also pullback squares:

$$
\begin{array}{ccc}
P(I) & \xrightarrow{P(\mathrm{tgt}_I)} & P(T) \\
\phi_I \downarrow & \quad \lrcorner & \downarrow \phi_T \\
Q(I) & \xrightarrow[P(\mathrm{tgt}_I)]{} & Q(T)
\end{array}
\qquad
\begin{array}{ccc}
P(O) & \xrightarrow{P(\mathrm{src}_O)} & P(T) \\
\phi_O \downarrow & \quad \lrcorner & \downarrow \phi_T \\
Q(O) & \xrightarrow[P(\mathrm{src}_O)]{} & Q(T)
\end{array} \cdot
$$

Petri nets with links and their morphisms form the category LPetri.

To explicate the definition, a Petri net with links P consists of

- a set $P(S)$ of **species**;
- a set $P(T)$ of **transitions**;
- a set $P(I)$ of **input arcs** going from species to transitions, via maps $P(\mathrm{src}_I) : P(I) \to P(S)$ and $P(\mathrm{tgt}_I) : P(I) \to P(T)$;
- a set $P(O)$ of **output arcs** going from transitions to species, via maps $P(\mathrm{src}_O) : P(O) \to P(T)$ and $P(\mathrm{tgt}_O) : P(O) \to P(S)$; and finally
- a set $P(L)$ of **links** going from species to transitions, via maps $P(\mathrm{src}_L) : P(L) \to P(S)$ and $P(\mathrm{tgt}_L) : P(L) \to P(T)$.

The property of preserving arities, called "etale" by Kock [Koc22, §3.4], means that a morphism $\phi : P \to Q$ of Petri nets with links preserves the input and output arities of all transitions. Namely, for each transition t in the net P the map $\phi_I : P(I) \to Q(I)$ restricts to a bijection between the input arcs to t and to $\phi_T(t)$, and similarly the map $\phi_O : P(O) \to Q(O)$ restricts to a bijection between the output arcs from t and from $\phi_T(t)$. This property seems appropriate for many purposes, including in biochemistry, but for mathematical convenience we will not always assume it.

Remark 2.12 (Related literature). While not appearing in the literature in precisely this form, our definition of a Petri net with links is closely related to several existing concepts. Kock has described Petri nets as copresheaves on a category with objects S, T, I, O [Koc22], calling them *whole-grain Petri nets* to distinguish them from classical Petri nets, whose semantics are subtly different [Bae+21]. Meanwhile, the concept of a link is essential to *stock and flow diagrams*, originating in the field of system dynamics [For61; Ste00] and recently given a rigorous categorical account [Bae+22]. The classical literature on Petri nets has also considered Petri nets equipped with positive/context and negative/inhibitory arcs [AF73], called *inhibitor nets* [Hac76; Bal+04]. Roughly speaking, our Petri nets with signed links are related to inhibitor nets ([Bal+04, Definition 1]) in the same way that Kock's whole-grain Petri nets are related to classical Petri nets; in both cases, working with copresheaves has technical advantages. Finally, we note that the *Petri nets with catalysts* proposed by Baez, Foley, and Moeller [BFM19] differ significantly from Petri nets with links: the former fix a subset of the species to be catalysts throughout the net, whereas the latter uses links to make catalyzation specific to individual reactions.

Remark 2.13 (Petri nets as typed graphs). Like bare Petri nets, Petri nets with links can be described as graphs with two types of vertices. To see this, take the graph

$$T_{\mathsf{LPetri}} \coloneqq \left\{ \; S \underset{L}{\overset{I}{\rightleftarrows}} \; O \longrightarrow T \; \right\}$$

with vertices S and T and edges I, O, and L. The category of Petri nets with links and natural transformations is isomorphic to the slice category $\mathsf{Graph}/T_{\mathsf{LPetri}}$. Moreover, the schema category $\mathsf{Sch}(\mathsf{LPetri})$ is isomorphic to the category of elements of the functor $T_{\mathsf{LPetri}} : \mathsf{Sch}(\mathsf{Graph}) \to \mathsf{Set}$, exemplifying a general fact about slices of copresheaf categories [Str00, Remark p. 303].

Petri nets with *signed* links are defined analogously to signed graphs (Definition 2.2).

Definition 2.14 (Petri nets with signed links). The **category of Petri nets with signed links** is the slice category

$$\mathsf{SgnPetri} \coloneqq \mathsf{LPetri}/P_{\mathsf{Sgn}},$$

where P_{Sgn} is the Petri net with links having one species, one transition, one input arc, one output arc, and two links, namely the elements of Sgn.

Incidentally, the morphism $P \to P_{\mathsf{Sgn}}$ defining a Petri net with signed links does *not* preserve arities unless every transition in P has exactly one input and one output.

We now turn to the main construction of this section, a functor that sends a Petri net with signed links to the signed graph of interactions implied by the net. On the example in Equations (14) and (15), this functor gives the signed graph

$$\begin{array}{c} A \overset{\longleftarrow}{\underset{}{\longrightarrow}} B \overset{\longrightarrow}{\underset{}{\rightrightarrows}} D \\ \circlearrowleft \qquad\quad \circlearrowleft \qquad \circlearrowleft \\ C \\ \circlearrowleft \end{array} \quad . \tag{16}$$

In general, the resulting signed graph has the Petri net's species as vertices, and has signed edges of the following four types:

(a) for every input-output pair to a transition, a positive edge from input to output;
(b) for every input to a transition, a negative self loop, representing consumption by the reaction;
(c) for every signed link, an edge of *opposite* sign going from the linked species to each input to the linked transition;
(d) for every signed link, an edge of *equal* sign going from the linked species to each output from the linked transition.
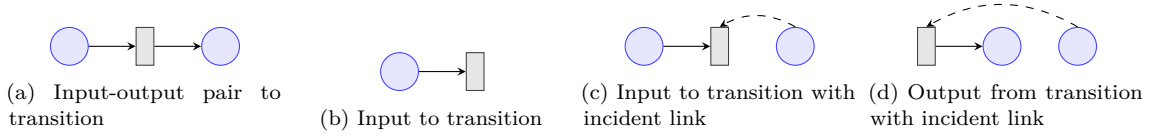
(a) Input-output pair to transition

(b) Input to transition

(c) Input to transition with incident link

(d) Output from transition with incident link

Figure 2: Four different Petri nets with links. For each of these instances P, evaluating the representable functor $\mathrm{Hom}_{\mathsf{LPetri}}(P, -) : \mathsf{LPetri} \to \mathsf{Set}$ gives the edges for one case in the case analysis that defines the functor from Petri nets with signed links to signed graphs (Proposition 2.15).

All four cases are visible in the example of Equation (16). Each of these cases is detected by a representable functor of the form

$$\mathrm{Hom}_{\mathsf{LPetri}}(P, -) : \mathsf{LPetri} \to \mathsf{Set},$$

where P is one of the four Petri nets with links shown in Figure 2. Each net represents a "pattern" whose matches produce edges in the signed graph.

The resulting functor $\mathsf{SgnPetri} \to \mathsf{SgnGraph}$, which restricts on vertices to the functor picking out the set of species of the Petri net (representable by the Petri net with one species and no transitions) and restricts on edges to the disjoint union of the four representable functors described above, has the form of a *familially representable functor* between copresheaf categories. This type of functor has been studied extensively within both pure (see, for instance, [Lei04, Appendix C]) and applied category theory (see [Spi21]). To be consistent with our existing notation and to emphasize that such functors have a useful computational interpretation, we adopt Spivak's perspective that they are *data migration functors*. In this framing, the schema category of a copresheaf category is regarded as the schema for a database, a copresheaf on that category is regarded as a database instance on the schema, and representable functors from the copresheaf category to Set are regarded as *conjunctive queries*. A data migration functor between copresheaf categories is then one which restricts on each component of the codomain to a disjoint union of conjunctive queries, or *duc-query*.

In order to apply this theory and precisely define the functor $\mathsf{SgnPetri} \to \mathsf{SgnGraph}$, we fully schematize the definitions of signed graphs and Petri nets with signed links. The **schema for signed graphs** is the category $\mathsf{Sch}(\mathsf{SgnGraph})$ freely generated by these objects and morphisms:

$$V \underset{\mathrm{tgt}}{\overset{\mathrm{src}}{\rightleftarrows}} E \xrightarrow{\mathrm{sgn}} A \circlearrowleft \mathrm{neg}.$$

A signed graph as in Definition 2.2 is equivalent to a functor $X : \mathsf{Sch}(\mathsf{SgnGraph}) \to \mathsf{Set}$ such that $X(A) = \mathsf{Sgn}$, the set of signs, and $X(\mathrm{neg}) : \mathsf{Sgn} \to \mathsf{Sgn}$ is negation (i.e., multiplication by -1). Note that negation is not needed to define the data of a signed graph but is relevant to the data migration. A morphism of signed graphs $X \to Y$ is a natural transformation $\phi : X \to Y$ whose component at A is the identity function, $\phi_A = 1_{\mathsf{Sgn}}$. We thus obtain a category isomorphic to $\mathsf{SgnGraph}$.

Similarly, the **schema for Petri nets with signed links** is the category $\mathsf{Sch}(\mathsf{SgnPetri})$ freely generated by:

A Petri net with signed links, as in Definition 2.14, is equivalent to a functor $P : \mathsf{Sch}(\mathsf{SgnPetri}) \to \mathsf{Set}$ such that $P(A) = \mathsf{Sgn}$, the map $P(\mathrm{neg}) : \mathsf{Sgn} \to \mathsf{Sgn}$ is negation, and $P(\mathrm{one}) : P(T) \to \mathsf{Sgn}$ is the constant map at $+1$. Again, these maps are needed for data migration, not for the data itself. A morphism of Petri nets with signed links is a natural transformation $\phi : P \to Q$ such $\phi_A = 1_{\mathsf{Sgn}}$, yielding a category isomorphic to $\mathsf{SgnPetri}$.

With these preliminaries, we can construct the "interactions functor," called "Int" for short, from Petri nets with signed links to signed graphs.

Proposition 2.15 (Interactions functor). *A functor*

$$\text{Int} : \mathsf{SgnPetri} \to \mathsf{SgnGraph}$$

is specified by the following functor from $\mathsf{Sch}(\mathsf{SgnGraph})$ *to the category of duc-queries on* $\mathsf{Sch}(\mathsf{SgnPetri})$*:*

$$\mathsf{Sch}(\mathsf{SgnGraph}) \to \mathrm{II}\left(\left(\mathsf{Set}^{\mathsf{Sch}(\mathsf{SgnPetri})}\right)^{\mathrm{op}}\right)$$

$$V \mapsto S$$
$$E \mapsto I \times_T O + I + I \times_T L + O \times_T L$$
$$A \mapsto A \tag{17}$$
$$\mathrm{src} \mapsto [\mathrm{src}_I \circ \pi_I, \ \mathrm{src}_I, \ \mathrm{src}_L \circ \pi_L, \ \mathrm{src}_L \circ \pi_L]$$
$$\mathrm{tgt} \mapsto [\mathrm{tgt}_O \circ \pi_O, \ \mathrm{src}_I, \ \mathrm{src}_I \circ \pi_I, \ \mathrm{tgt}_O \circ \pi_O]$$
$$\mathrm{sgn} \mapsto [\mathrm{one} \circ \pi_T, \ \mathrm{neg} \circ \mathrm{one} \circ \mathrm{tgt}_I, \ \mathrm{neg} \circ \mathrm{sgn} \circ \pi_L, \ \mathrm{sgn} \circ \pi_L]$$
$$\mathrm{neg} \mapsto \mathrm{neg} .$$

Proof. We will define the functor $\text{Int} : \mathsf{SgnPetri} \to \mathsf{SgnGraph}$ as the restriction of a functor $\mathsf{Set}^{\mathsf{D}} \to \mathsf{Set}^{\mathsf{C}}$ between the categories of copresheaves on $\mathsf{D} := \mathsf{Sch}(\mathsf{SgnPetri})$ and $\mathsf{C} := \mathsf{Sch}(\mathsf{SgnGraph})$. In fact, the functor $\mathsf{Set}^{\mathsf{D}} \to \mathsf{Set}^{\mathsf{C}}$ is of the special kind known as a *parametric right adjoint* [Str00, Definition p. 311] or a *familially representable functor* [Lei04, Appendix C].

According to the theory of data migration [Spi21, Corollary 2.3.6], giving a parametric right adjoint $\mathsf{Set}^{\mathsf{D}} \to \mathsf{Set}^{\mathsf{C}}$ is equivalent to giving a functor from C to $\mathrm{II}((\mathsf{Set}^{\mathsf{D}})^{\mathrm{op}})$, the free coproduct completion of the free limit completion of D. Our functor $\mathsf{C} \to \mathrm{II}((\mathsf{Set}^{\mathsf{D}})^{\mathrm{op}})$ is defined by Equation (17). The assignment of $E \in \mathsf{C}$ can also be described as the sum of the four representables associated with the Petri nets with links in Figure 2.

Finally, the assignments $A \mapsto A$ and $\mathrm{neg} \mapsto \mathrm{neg}$ ensure that if P is a copresheaf on D with $P(A) = \mathsf{Sgn}$ and $P(\mathrm{neg})$ is negation, then applying this parametric right adjoint functor to P yields a copresheaf X on C where again $X(A) = \mathsf{Sgn}$ and $X(\mathrm{neg})$ is negation. Thus, this functor between copresheaf categories restricts to a functor $\mathsf{SgnPetri} \to \mathsf{SgnGraph}$ as claimed. $\square$

Using the interactions functor, we can give a formal account of what it means to have a mechanistic model for a regulatory network.

Definition 2.16 (Mechanism). A **mechanistic model** for a regulatory network X is a Petri net with signed links P together with an occurrence of X in $\text{Int}(P)$, i.e., a monic signed functor $X \rightarrowtail \text{Int}(P)$.

For example, the Petri net with signed links in Equation (15) is a mechanistic model for a regulatory network in which A and D participate in a positive feedback loop:

$$A \, \overbrace{\underbrace{}} \, D \ .$$
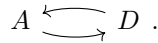
The next example is inspired by activation and inhibition of enzymes [Ing13, Chapter 3]. Besides being of greater biological interest, it shows why Petri nets with links are better suited than bare Petri nets to describe mechanisms.

Example 2.17 (Enzyme inhibition). Suppose that an enzyme E catalyzes the conversion of a substrate S into a product P, and that an inhibiting agent I transforms the chemical configuration of the reactive enzyme E into the inert compound F. Then the agent I inhibits the product P:

$$I \longmapsto P \ .$$

As a first attempt to model these interactions, consider the following Petri net without links.
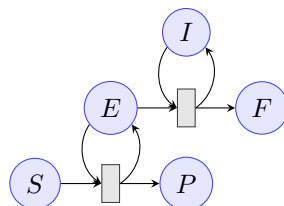
One might expect this Petri net to provide a mechanism for the inhibition of P by I, but at least according to our definition, it does not. Applying the functor $\mathsf{Int} : \mathsf{SgnPetri} \to \mathsf{SgnGraph}$ yields the regulatory network

in which there is no occurrence of the desired inhibition. If instead we allow links, the reaction catalyzed by E should be written as having input S, output P, and a link from E to the transition. A similar change is required for the reaction involving E, F, and I. A better model of the reactions is thus the following Petri net with links.

Applying the functor Int now yields the network

which indeed has an occurrence of I inhibiting P.

3 Quantitative analysis: parameters and dynamics

3.1 Parameterized dynamical systems

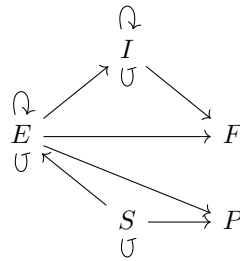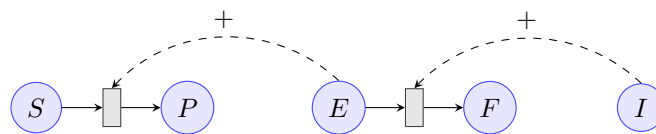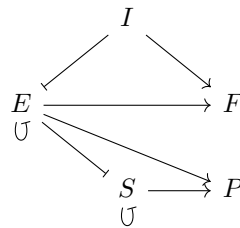Pioneering the idea of functorial semantics for scientific models, Baez and Pollard extended the mass-action model of reaction networks to a functor from the category of Petri nets with rates into a category of dynamical systems [BP17]. In this picture, the reaction rate coefficients are known constants associated with the reaction network. In practice, however, rate coefficients are often unknown and must be extracted from existing literature or estimated from experimental data. We therefore change our perspective slightly and consider dynamical systems not in isolation but as *parameterized families*. This shift also turns out to have formal advantages: the category of parameterized dynamical systems is better behaved than the category of dynamical systems, which has too few morphisms.

To begin, we recall the dynamics functor, nearly identical to Baez-Pollard's [BP17, Lemma 15]:

Lemma 3.1 (Dynamics). *There is a functor* $\mathsf{Dynam} : \mathsf{FinSet} \to \mathsf{Vect}_{\mathbb{R}}$ *that sends*

- *a finite set S to the vector space of algebraic vector fields $v : \mathbb{R}^S \to \mathbb{R}^S$, where **algebraic** means that the components of the vector field are polynomials in the state variables;*

- *a function $f : S \to S'$ between finite sets to the linear transformation*

$$(v : \mathbb{R}^S \to \mathbb{R}^S) \quad \mapsto \quad (f_* \circ v \circ f^* : \mathbb{R}^{S'} \to \mathbb{R}^{S'}),$$

*where the linear map $f^* : \mathbb{R}^{S'} \to \mathbb{R}^S$ is the **pullback** along f*

$$f^*(x')(i) \coloneqq x'(f(i)), \qquad x' \in \mathbb{R}^{S'}, \ i \in S,$$

*and the linear map $f_* : \mathbb{R}^S \to \mathbb{R}^{S'}$ is the **pushforward** along f*

$$f_*(x)(i') \coloneqq \sum_{i \in f^{-1}(i')} x(i), \qquad x \in \mathbb{R}^S, \ i' \in S'.$$

Proof. The functor $\mathsf{Dynam} : \mathsf{FinSet} \to \mathsf{Vect}_\mathbb{R}$ can be constructed as the composite functor

$$\mathsf{FinSet} \xrightarrow{\langle D, F \rangle} \mathsf{Vect}_\mathbb{R}^{\mathrm{op}} \times \mathsf{Vect}_\mathbb{R} \xrightarrow{\mathsf{Poly}_\mathbb{R}(-,-)} \mathsf{Vect}_\mathbb{R},$$

where $F : \mathsf{FinSet} \to \mathsf{Vect}_\mathbb{R}$, $X \mapsto \mathbb{R}^X$ is the free vector space functor (restricted to finite sets); $D : \mathsf{FinSet}^{\mathrm{op}} \to \mathsf{Vect}_\mathbb{R}$ is the dual vector space functor (restricted to F), whose underlying set-valued functor is

$$\mathsf{Vect}_\mathbb{R}(F(-), \mathbb{R}) \cong \mathsf{Set}(-, \mathbb{R}) : \mathsf{FinSet}^{\mathrm{op}} \to \mathsf{Set};$$

and $\mathsf{Poly}_\mathbb{R}(-, -)$ is the $\mathsf{Vect}_\mathbb{R}$-valued hom-functor that sends a pair of vector spaces to the vector space of polynomial functions between them.[4] $\qquad \square$

The dynamics functor is the same one studied by Baez and Pollard except that we take the vector space, rather than merely the set, of vector fields. That is because we are interested in *linearly* parameterized dynamical systems. In calling the functor "dynamics," we implicitly identify a vector field with the differentiable dynamical system that it generates. This common practice is not entirely innocent since even when a system of differential equations depends linearly on parameters, its solutions rarely do. We also note that the restriction to *algebraic* vector fields, as opposed to smooth or even just continuous ones, is inessential but suffices for us and agrees with Baez-Pollard.

The dynamics functor is the main building block in constructing a category of parameterized dynamical systems.

Definition 3.2 (Linear parameterizations). The **category of linearly parameterized dynamical systems** is the comma category

$$\mathsf{Para}(\mathsf{Dynam}) \coloneqq F/\mathsf{Dynam},$$

where $F : \mathsf{FinSet} \to \mathsf{Vect}_\mathbb{R}$, $X \mapsto \mathbb{R}^X$ is the free vector space functor restricted to finite sets.

So, by definition, a linearly parameterized dynamical system consists of a finite set P, the **parameter variables**, and a finite set S, the **state variables**, together with a linear map

$$v : \mathbb{R}^P \to \mathrm{Dynam}(S)$$

sending each choice of parameters $\theta \in \mathbb{R}^P$ to an algebraic vector field $v(\theta)$. In more conventional notation, we can write $v(x; \theta) \coloneqq v(\theta)(x)$ for $x \in \mathbb{R}^S$ and $\theta \in \mathbb{R}^P$. A morphism $(P, S, v) \to (P', S', v')$ of linearly parameterized dynamical systems is a pair of functions $q : P \to P'$ and $f : S \to S'$ making the square

$$
\begin{array}{ccc}
\mathbb{R}^P & \xrightarrow{\ v\ } & \mathrm{Dynam}(S) \\
{\scriptstyle q_*}\downarrow & & \downarrow{\scriptstyle f_* \circ (-) \circ f^*} \\
\mathbb{R}^{P'} & \xrightarrow[\ v'\ ]{} & \mathrm{Dynam}(S')
\end{array}
\tag{18}
$$

commute, or equivalently making the equation

$$f_*(v(f^*(x'); \theta)) = v'(x'; q_*(\theta))$$

hold for all $x' \in \mathbb{R}^{S'}$ and $\theta \in \mathbb{R}^P$.

---

[4]For a coordinate-free description of polynomial maps between vector spaces, see [Car71, §1.6].

While certainly not all dynamical models depend linearly on their parameters, a great many of them do, including several important canonical models in biology and chemistry. The law of mass action defines dynamical systems that depend linearly on the rate coefficients. The generalized Lotka-Volterra equations, studied in the next section, are linear in the rate and affinity parameters. Of course, the mass-action and Lotka-Volterra equations are nonlinear ODEs; linearity of a vector field in state or in parameters are separate matters. Nevertheless, even for nonlinear models such as Lotka-Volterra, linearity in parameters is a useful assumption that aides in the identifiability analysis of the model [SRS14, §5].

To express important physical constraints and to define a semantics for signed graphs, we will restrict the dynamical system and its parameters to be nonnegative. This is straightforward enough but requires a bit of additional formalism.

Write $\mathbb{R}_+ \coloneqq \{x \in \mathbb{R} : x \geq 0\}$ for the semiring of nonnegative real numbers. A module over $\mathbb{R}_+$ is called a **conical space**, and the category of conical spaces and conic-linear ($\mathbb{R}_+$-linear) maps is denoted $\mathsf{Con} \coloneqq \mathsf{Mod}_{\mathbb{R}_+}$. A conical space is a structure in which one can take linear combinations with nonnegative real coefficients, just as a real vector space ($\mathbb{R}$-module) is a structure in which one can take linear combinations with arbitrary real coefficients. Any convex cone in a real vector space is a conical space. Our main example is the **nonnegative orthant** of $\mathbb{R}^S$ for some set S: the function space $\mathbb{R}_+^S \coloneqq \{x : S \to \mathbb{R}_+\}$, with conical combinations taken pointwise. A real vector space can itself be regarded as a conical space; more precisely, the inclusion of semirings $\mathbb{R}_+ \hookrightarrow \mathbb{R}$ induces a forgetful functor $\mathsf{Vect}_{\mathbb{R}} \to \mathsf{Con}$ by restriction of scalars.

Recall that a dynamical system is **nonnegative** if whenever the initial condition is in the nonnegative orthant, its trajectory always remains in the nonnegative orthant. A dynamical system of form $\dot{x} = v(x)$ is nonnegative if and only if $v_i(x) \geq 0$ whenever $x \geq 0$ componentwise and $x_i = 0$ [HCH10, Proposition 2.1], in which case the vector field v is called **essentially nonnegative** [HCH10, Definition 2.1]. Using this criterion, it is easy to see that a reaction network with mass-action kinetics is nonnegative assuming the rate constants are nonnegative, as is a Lotka-Volterra system for any choice of parameters. Hence both systems satisfy the obvious physical constraint that no species should have negative concentration or population.

Lemma 3.3 (Nonnegative dynamics). *There is a functor* $\mathrm{Dynam}_+ : \mathsf{FinSet} \to \mathsf{Con}$ *that sends a finite set S to the conical space of essentially nonnegative, algebraic vector fields $v : \mathbb{R}^S \to \mathbb{R}^S$ and sends a function $f : S \to S'$ to the transformation $v \mapsto f_* \circ v \circ f^*$.*

Proof. The proof is similar to that of Lemma 3.1. It is clear that the essentially nonnegative functions are stable under pointwise conical combinations, hence form a conical space. (They are, of course, *not* stable under arbitrary linear combinations.) We just need to check that if $v : \mathbb{R}^S \to \mathbb{R}^S$ is essentially nonnegative, then so is the transformed vector field $f_* \circ v \circ f^* : \mathbb{R}^{S'} \to \mathbb{R}^{S'}$. Fix $x' \in \mathbb{R}_+^{S'}$ and $i' \in S'$, and suppose that $x'(i') = 0$. For every $i \in f^{-1}(i')$, we have $f^*(x')(i) = x'(f(i)) = x'(i') = 0$ and so $v(f^*(x'))(i) \geq 0$, whence the inequality of essential nonnegativity follows:

$$(f_* \circ v \circ f^*)(x')(i') = \sum_{i \in f^{-1}(i')} v(f^*(x'))(i) \geq 0. \qquad \square$$

We now define the conical analogue of linearly parameterized dynamical systems.

Definition 3.4 (Conical parameterizations). The **category of conically parameterized nonnegative dynamical systems** is the comma category

$$\mathsf{Para}(\mathrm{Dynam}_+) \coloneqq F_+ / \mathrm{Dynam}_+ .$$

where $F_+ : \mathsf{FinSet} \to \mathsf{Con}$, $X \mapsto \mathbb{R}_+^X$ is the free conical space functor restricted to finite sets.

So, a conically parameterized nonnegative dynamical system consists of finite sets P and S together with a conic-linear map

$$v : \mathbb{R}_+^P \to \mathrm{Dynam}_+(S).$$

Proposition 3.5 (Colimits of parameterized dynamical systems). *The categories of linearly and conically parameterized dynamical systems are finitely cocomplete. Moreover, these finite colimits are computed by colimits in* FinSet *of the parameter variables and of the state variables.*

Proof. The category FinSet has finite colimits and the functors $F : \mathsf{FinSet} \to \mathsf{Vect}_{\mathbb{R}}$ and $F_+ :$ $\mathsf{FinSet} \to \mathsf{Con}$ preserve finite colimits, since they are composites of the inclusion $\mathsf{FinSet} \hookrightarrow \mathsf{Set}$ with the left adjoints

$$\mathsf{Set} \underset{U}{\overset{F}{\underset{\perp}{\rightleftarrows}}} \mathsf{Vect}_{\mathbb{R}} \qquad \text{and} \qquad \mathsf{Set} \underset{U_+}{\overset{F_+}{\underset{\perp}{\rightleftarrows}}} \mathsf{Con}$$

to the underlying set functors on vector spaces and conical spaces. By Lemma 3.6 below, the comma categories $\mathsf{Para}(\mathrm{Dynam}) = F / \mathrm{Dynam}$ and $\mathsf{Para}(\mathrm{Dynam}_+) = F_+ / \mathrm{Dynam}_+$ have finite colimits, which are preserved by the projection functors onto FinSet. $\qquad\square$

To illustrate, we describe the initial object and binary coproducts in $\mathsf{Para}(\mathrm{Dynam})$. The initial linearly parameterized dynamical system has no parameter variables, no state variables, and the unique (trivial) vector field on the zero vector space. The coproduct of two linearly parameterized dynamical systems (P_1, S_1, v_1) and (P_2, S_2, v_2) has parameter variables $P_1 + P_2$, state variables $S_1 + S_2$, and parameterized vector field

$$\mathbb{R}^{P_1+P_2} \cong \mathbb{R}^{P_1} \oplus \mathbb{R}^{P_2} \xrightarrow{v_1 \oplus v_2} \mathrm{Dynam}(S_1) \oplus \mathrm{Dynam}(S_2) \xrightarrow{[\mathrm{Dynam}(\iota_1), \mathrm{Dynam}(\iota_2))]} \mathrm{Dynam}(S_1 + S_2),$$

where $\iota_1 : S_1 \to S_1 + S_2$ and $\iota_2 : S_2 \to S_1 + S_2$ are the canonical inclusions. In conventional notation, the coproduct system has parameterized vector field

$$v \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} ; \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \right) = \begin{bmatrix} v_1(x_1; \theta_1) \\ v_2(x_2; \theta_2) \end{bmatrix}.$$

The proof of Proposition 3.5, as well as of Theorems 3.7 and 3.8 below, depends on the following technical lemma about comma categories, which the reader can omit without loss of continuity.

Lemma 3.6 (Colimits in comma categories). *Let $\mathsf{C}_0 \xrightarrow{F_0} \mathsf{C} \xleftarrow{F_1} \mathsf{C}_1$ be a cospan of categories such that C_0 and C_1 have colimits of shape J and F_0 preserves J-shaped colimits. Then the comma category F_0/F_1 also has J-shaped colimits, and the projection functors $\pi_i : F_0/F_1 \to \mathsf{C}_i$, $i = 0, 1$, preserve those colimits.*

Furthermore, a functor $G : \mathsf{X} \to F_0/F_1$ into the comma category preserves J-shaped colimits whenever the associated functors $G_i := \pi_i \circ G : \mathsf{X} \to \mathsf{C}_i$, $i = 0, 1$, do so.

Proof. Colimits in the comma category F_0/F_1 are constructed in [RB88, §5.2]. To make the rest of the proof self-contained, we recall the construction here.

By the universal property of the comma category, a diagram $D : \mathsf{J} \to F_0/F_1$ is equivalent to diagrams $D_i := \pi_i \circ D : \mathsf{J} \to \mathsf{C}_i$, $i = 0, 1$, along with a natural transformation $\vec{D} : F_0 \circ D_0 \Rightarrow F_1 \circ D_1$. Let (c_i, λ^i) be a colimit cocone for the diagram D_i in C_i, having legs $D_i(j) \xrightarrow{\lambda^i_j} c_i$ for each $j \in \mathsf{J}$. The family of morphisms

$$F_0(D_0(j)) \xrightarrow{\vec{D}_j} F_1(D_1(j)) \xrightarrow{F_1(\lambda^1_j)} F_1(c_1), \qquad j \in \mathsf{J},$$

is then a cocone under $F_0 \circ D_0$. Since F_0 preserves J-shaped limits, $(F_0(c_0), F_0 * \lambda^0)$ is a colimit cocone for $F_0 \circ D_0$, so by its universal property, there exists a unique morphism $f : F_0(c_0) \to F_1(c_1)$ making the squares commute:

$$\begin{array}{ccc} F_0(D_0(j)) & \xrightarrow{\vec{D}_j} & F_1(D_1(j)) \\ {\scriptstyle F_0(\lambda^0_j)} \downarrow & & \downarrow {\scriptstyle F_1(\lambda^1_j)} \\ F_0(c_0) & \xrightarrow{f} & F_1(c_1) \end{array} \qquad j \in \mathsf{J}.$$

Setting $\lambda := (\lambda^0_j, \lambda^1_j)_{j \in \mathsf{J}}$, the cocone $((c_0, c_1, f), \lambda)$ can be shown to be a colimit of the diagram D.

To prove the last statement about colimit preservation, let $D : \mathsf{J} \to \mathsf{X}$ be a diagram with colimit cocone (x, λ), having legs $Dj \xrightarrow{\lambda_j} y$ for $j \in \mathsf{J}$. We must show that its image cocone $(G(x), G * \lambda)$ is

a colimit of the diagram $G \circ D$ in F_0/F_1. By the universal property of the comma category, the functor $G : \mathsf{X} \to F_0/F_1$ is equivalent to the functors $G_i : \mathsf{X} \to \mathsf{C}_i$, $i = 0, 1$, along with a natural transformation $\vec{G} : F_0 \circ G_0 \Rightarrow F_1 \circ G_1$. The image cocone $(G(x), G * \lambda)$ then amounts to cocones $(G_0(x), G_0 * \lambda)$ and $(G_1(x), G_1 * \lambda)$, which by hypothesis are colimits of the diagrams $G_0 \circ D$ and $G_1 \circ D$ in C_0 and C_1, together with a family of commutative squares in C:

$$
\begin{array}{ccc}
F_0(G_0(Dj)) & \xrightarrow{\vec{G}_{Dj}} & F_1(G_1(Dj)) \\
{\scriptstyle F_0(G_0(\lambda_j))}\Big\downarrow & & \Big\downarrow{\scriptstyle F_1(G_1(\lambda_j))} \\
F_0(G_0(x)) & \xrightarrow[\vec{G}_x]{} & F_1(G_1(x))
\end{array} \qquad j \in \mathsf{J}.
$$

But a morphism $\vec{G}_x$ making all these squares commute is already uniquely determined by the universal property of the colimit cocone $(F_0(G_0(x)), F_0 * G_0 * \lambda)$ for the diagram $F_0 \circ G_0 \circ D$, using the hypothesis that F_0 preserves J-shaped colimits. Indeed, this is precisely how one constructs the colimit of the diagram $G \circ D$ in F_0/F_1, as shown above. It follows that $(G(x), G * \lambda)$ is a colimit cocone for $G \circ D$. $\qquad\square$

3.2 The Lotka-Volterra dynamical model

A **Lotka-Volterra system** with n species has, using matrix notation, the vector field

$$
v(x; \rho, \beta) \coloneqq x \odot (\rho + \beta x) = \mathrm{diag}(x)(\rho + \beta x)
$$

with state vector $x \in \mathbb{R}^n$ and arbitrary real-valued parameters $\rho \in \mathbb{R}^n$ and $\beta \in \mathbb{R}^{n \times n}$ [SMH18, §2.2]. In coordinates, the vector field is

$$
v_i(x; \rho, \beta) = x_i \left(\rho_i + \sum_{j=1}^n \beta_{i,j} x_j \right) = \rho_i x_i + \sum_{j=1}^n \beta_{i,j} x_i x_j, \qquad i = 1, \dots, n.
$$

The parameter ρ_i sets the baseline rate of growth (when positive) or decay (when negative) for species i, whereas $\beta_{i,j}$ defines a promoting (when positive) or inhibiting (when negative) effect of species j on species i. In typical applications the signs of the parameters are fixed and known in advance of any data. For example, in the famous **predator-prey Lotka-Volterra system**

$$
\dot{x} = ax - bxy
$$
$$
\dot{y} = dxy - cy,
$$

with prey x and predators y, the parameters $\rho = \begin{bmatrix} a \\ -c \end{bmatrix}$ and $\beta = \begin{bmatrix} 0 & -b \\ d & 0 \end{bmatrix}$ are specified by nonnegative real numbers $a, b, c, d \geq 0$.

In this section, we define quantitative semantics for graphs and signed graphs using the Lotka-Volterra dynamical model. To illustrate the main ideas, we first construct a functor from finite graphs (Definition 2.1) to linearly parameterized dynamical systems (Definition 3.2), giving a semantics for unlabeled graphs. It is more useful to have a semantics for regulatory networks, which we have defined to be signed graphs. We therefore construct a second functor from finite signed graphs (Definition 2.2) to conically parameterized nonnegative dynamical systems (Definition 3.4).

Recall that a graph is **finite** if its vertex and edge sets are both finite. Let $\mathsf{FinGraph}$ denote the full subcategory of Graph spanned by finite graphs.

Theorem 3.7 (Lotka-Volterra model for finite graphs)**.** *There is a functor*

$$
\mathrm{LV} : \mathsf{FinGraph} \to \mathsf{Para}(\mathrm{Dynam})
$$

that sends

- *a finite graph X to the linearly parameterized dynamical system with parameter variables $P \coloneqq X(V) + X(E)$, state variables $S \coloneqq X(V)$, and algebraic vector field*[5]

$$v(x; \rho, \beta)(i) \coloneqq \rho(i)\, x(i) + \sum_{(e : i' \to i) \in X} \beta(e)\, x(i')\, x(i), \qquad x \in \mathbb{R}^{X(V)},\ i \in X(V),$$

  parameterized by vectors $\rho \in \mathbb{R}^{X(V)}$ and $\beta \in \mathbb{R}^{X(E)}$;
- *a graph homomorphism $\phi : X \to Y$ to a morphism of systems with parameter variable map $\phi_V + \phi_E : X(V) + X(E) \to Y(V) + Y(E)$ and state variable map $\phi_V : X(V) \to Y(V)$.*

Moreover, the functor LV *preserves finite colimits.*

Proof. By the universal property of the comma category $\mathsf{Para}(\mathsf{Dynam}) = F/\mathsf{Dynam}$, to give a functor $\mathrm{LV} : \mathsf{FinGraph} \to \mathsf{Para}(\mathsf{Dynam})$ is to give a pair of functors $\mathrm{LV}_0, \mathrm{LV}_1 : \mathsf{FinGraph} \to \mathsf{FinSet}$ along with a natural transformation

$$\vec{\mathrm{LV}} : (F \circ \mathrm{LV}_0) \Rightarrow (\mathsf{Dynam} \circ \mathrm{LV}_1) : \mathsf{FinGraph} \to \mathsf{Vect}_{\mathbb{R}}.$$

We set $\mathrm{LV}_0(X) \coloneqq X(V) + X(E)$ and $\mathrm{LV}_1(X) \coloneqq X(V)$. Using the universal property of the coproduct in $\mathsf{Vect}_{\mathbb{R}}$, the components

$$\vec{\mathrm{LV}}_X : \mathbb{R}^{X(V)} \oplus \mathbb{R}^{X(E)} \cong \mathbb{R}^{X(V)+X(E)} \to \mathsf{Dynam}(X(V)).$$

of the transformation $\vec{\mathrm{LV}}$ themselves decompose into two parts, call them

$$v_X^0 \coloneqq \vec{\mathrm{LV}}_X^0 : \mathbb{R}^{X(V)} \to \mathsf{Dynam}(X(V)) \quad \text{and} \quad v_X^1 \coloneqq \vec{\mathrm{LV}}_X^1 : \mathbb{R}^{X(E)} \to \mathsf{Dynam}(X(V)).$$

We define these to be

$$v_X^0(x; \rho)(i) \coloneqq \rho(i)\, x(i) \qquad \text{and} \qquad v_X^1(x; \beta)(i) \coloneqq \sum_{e \in X(\mathrm{tgt})^{-1}(i)} \beta(e)\, x(X(\mathrm{src})(e))\, x(i).$$

Putting the pieces back together reproduces the first statement of the theorem. We just need to check that the transformation $\vec{\mathrm{LV}}$ is, in fact, natural.

Given a graph homomorphism $\phi : X \to Y$, the naturality square for the transformation $\vec{\mathrm{LV}}$ is

$$
\begin{array}{ccc}
\mathbb{R}^{X(V)+X(E)} & \xrightarrow{\ \vec{\mathrm{LV}}_X\ } & \mathsf{Dynam}(X(V)) \\
{\scriptstyle (\phi_V + \phi_E)_*}\downarrow & & \downarrow{\scriptstyle (\phi_V)_* \circ (-) \circ (\phi_V)^*} \\
\mathbb{R}^{Y(V)+Y(E)} & \xrightarrow[\ \vec{\mathrm{LV}}_Y\ ]{} & \mathsf{Dynam}(Y(V))
\end{array}
\tag{19}
$$

which decomposes into two squares,

$$
\begin{array}{ccc}
\mathbb{R}^{X(V)} & \xrightarrow{\ v_X^0\ } & \mathsf{Dynam}(X(V)) \\
{\scriptstyle (\phi_V)_*}\downarrow & & \downarrow{\scriptstyle (\phi_V)_* \circ (-) \circ (\phi_V)^*} \\
\mathbb{R}^{Y(V)} & \xrightarrow[\ v_Y^0\ ]{} & \mathsf{Dynam}(Y(V))
\end{array}
\qquad \text{and} \qquad
\begin{array}{ccc}
\mathbb{R}^{X(E)} & \xrightarrow{\ v_X^1\ } & \mathsf{Dynam}(X(V)) \\
{\scriptstyle (\phi_E)_*}\downarrow & & \downarrow{\scriptstyle (\phi_V)_* \circ (-) \circ (\phi_V)^*} \\
\mathbb{R}^{Y(E)} & \xrightarrow[\ v_Y^1\ ]{} & \mathsf{Dynam}(Y(V))
\end{array}.
$$

Let us check that both squares commute. For the first, we have

$$(\phi_V)_*(v_X^0(\phi_V^*(y); \rho))(j) = \sum_{i \in \phi_V^{-1}(j)} v_X^0(y \circ \phi_V; \rho)(i) = \sum_{i \in \phi_V^{-1}(j)} \rho(i)\, y(\phi_V(i))$$

$$= \left(\sum_{i \in \phi_V^{-1}(j)} \rho(i) \right) y(j) = v_Y^0(y; (\phi_V)_*(\rho))(j)$$

---

[5]For a fixed graph X and vertex $i \in X(V)$, the notation $(e : i' \to i) \in X$ means any edge $e \in X(\mathrm{tgt})^{-1}(i)$ incoming to i, whose source $i' = X(\mathrm{src})(e)$ varies with e.

for all $y \in \mathbb{R}^{Y(V)}$, $\rho \in \mathbb{R}^{X(V)}$, and $j \in Y(V)$. For the second, we have

$$
\begin{aligned}
(\phi_V)_*(v_X^1(\phi_V^*(y); \beta))(j) &= \sum_{i \in \phi_V^{-1}(j)} v_X^1(y \circ \phi_V; \beta)(i) \\
&= \sum_{i \in \phi_V^{-1}(j)} \sum_{e \in X(\text{tgt})^{-1}(i)} \beta(e)\, y(\phi_V(X(\text{src})(e)))\, y(j) \\
&= \sum_{f \in Y(\text{tgt})^{-1}(j)} \sum_{e \in \phi_E^{-1}(f)} \beta(e)\, y(Y(\text{src})(\phi_E(e)))\, y(j) \\
&= \sum_{f \in Y(\text{tgt})^{-1}(j)} \left(\sum_{e \in \phi_E^{-1}(f)} \beta(e) \right) y(Y(\text{src})(f)) y(j) \\
&= v_Y^1(y; (\phi_E)_*(\beta))(j),
\end{aligned}
$$

for all $y \in \mathbb{R}^{Y(V)}$, $\beta \in \mathbb{R}^{X(E)}$, and $j \in Y(V)$. When exchanging the order of the summations we have used the facts that the graph homomorphism $\phi : X \to Y$ preserves sources and targets, the latter in its contravariant form

$$
\begin{array}{ccc}
\begin{array}{ccc}
X(E) & \xrightarrow{X(\text{tgt})} & X(V) \\
\phi_E \downarrow & & \downarrow \phi_V \\
Y(E) & \xrightarrow[Y(\text{tgt})]{} & Y(V)
\end{array}
& \rightsquigarrow &
\begin{array}{ccc}
\mathcal{P}(Y(V)) & \xrightarrow{\phi_V^{-1}} & \mathcal{P}(X(V)) \\
Y(\text{tgt})^{-1} \downarrow & & \downarrow X(\text{tgt})^{-1} \\
\mathcal{P}(Y(E)) & \xrightarrow[\phi_E^{-1}]{} & \mathcal{P}(X(E))
\end{array} ,
\end{array}
$$

where $\mathcal{P}(S)$ denotes the power set of a set S.

Finally, we must verify that the functor LV preserves finite colimits. By Lemma 3.6, that happens provided both functors $\text{LV}_0, \text{LV}_1 : \mathsf{FinGraph} \to \mathsf{FinSet}$ preserve finite colimits. The functor $\text{LV}_1 = \text{ev}_V$ is an evaluation functor on a copresheaf category, hence preserves colimits [Rie16, Proposition 3.3.9]. Since coproducts commute with colimits, the pointwise coproduct of two evaluation functors

$$
\text{LV}_0 = \left(\mathsf{FinGraph} \xrightarrow{\langle \text{ev}_V, \text{ev}_E \rangle} \mathsf{FinSet} \times \mathsf{FinSet} \xrightarrow{+} \mathsf{FinSet} \right)
$$

also preserves colimits. This completes the proof. $\qquad \square$

A quantitative semantics for signed graphs can be defined similarly, subject to a caveat about the vertex parameters. Our notion of signed graph, designed to capture regulatory networks as studied in the biochemistry literature, attaches signs only to edges. We are thus led to assume that, in the Lotka-Volterra dynamical model, all species have baseline rates of *decay* rather than growth. This assumption is valid for some, though certainly not all, regulatory networks [TN10, p. 222]. It is not suitable for predator-prey models in ecology.

Baseline rates of growth or decay could, if needed, be parameterized more flexibly. Most simply, one could attach signs to vertices as well as edges and use them in the quantitative semantics. Alternatively, at the expense of a more cumbersome formalism, one could define dynamical systems with mixed linear-conical parameterizations, allowing the vertex parameters to be arbitrary real numbers while the edge parameters are constrained to be nonnegative.[6] For simplicity and uniformity of presentation, we do not describe these extensions further.

Let $\mathsf{FinSgnGraph}$ denote the full subcategory of $\mathsf{SgnGraph}$ spanned by finite signed graphs.

Theorem 3.8 (Lotka-Volterra model for finite signed graphs). *There is a functor*

$$
\text{LV} : \mathsf{FinSgnGraph} \to \mathsf{Para}(\text{Dynam}_+)
$$

---

[6]Similar mixed parameterizations are a practical necessity in parametric *statistical* models, studied in detail in one author's PhD thesis [Pat20].

that sends a finite signed graph X to the conically parameterized nonnegative dynamical system with parameter variables $P := X(V) + X(E)$, state variables $S := X(V)$, and essentially nonnegative, algebraic vector field

$$v(x; \rho, \beta)(i) := -\rho(i)\, x(i) + \sum_{(e:i' \to i) \in X} X(\mathrm{sgn})(e)\, \beta(e)\, x(i')\, x(i), \qquad x \in \mathbb{R}^{X(V)},\ i \in X(V),$$

parameterized by $\rho \in \mathbb{R}_+^{X(V)}$ and $\beta \in \mathbb{R}_+^{X(E)}$. Moreover, the functor LV preserves finite colimits.

Proof. Similarly to the previous proof, the functor $\mathrm{LV} : \mathsf{FinSgnGraph} \to \mathsf{Para}(\mathsf{Dynam}_+)$ is defined by functors $\mathrm{LV}_0, \mathrm{LV}_1 : \mathsf{FinSgnGraph} \to \mathsf{FinSet}$ along with a natural transformation

$$\vec{\mathrm{LV}} : F_+ \circ \mathrm{LV}_0 \Rightarrow \mathrm{Dynam}_+ \circ \mathrm{LV}_1 : \mathsf{FinSgnGraph} \to \mathsf{Con},$$

now having components $\vec{\mathrm{LV}}_X$ given by the copairing of

$$v_X^0 := \vec{\mathrm{LV}}_X^0 : \mathbb{R}_+^{X(V)} \to \mathrm{Dynam}_+(X(V)) \quad \text{and} \quad v_X^1 := \vec{\mathrm{LV}}_X^1 : \mathbb{R}_+^{X(E)} \to \mathrm{Dynam}_+(X(V)),$$

where we define

$$v_X^0(x; \rho)(i) := -\rho(i)\, x(i) \quad \text{and} \quad v_X^1(x; \beta)(i) := \sum_{e \in X(\mathrm{tgt})^{-1}(i)} X(\mathrm{sgn})(e)\, \beta(e)\, x(X(\mathrm{src})(e))\, x(i).$$

The proof of naturality is essentially the same as before, using the crucial additional fact that morphisms of signed graphs preserve signs. The proof that the functor LV preserves finite colimits is unchanged. □

To exemplify the theorem, let us see how the Lotka-Volterra dynamics functor acts on a monomorphism and on an epimorphism of signed graphs.

In order to compare the dynamics of two species A and B involved in a negative feedback loop versus A and B in isolation, we take the inclusion of signed graphs

$$A \qquad B \qquad \overset{\iota}{\hookrightarrow} \qquad A \overset{\frown}{\underset{\smile}{\quad}} B$$

Labeling the edges in the feedback loop as AB and BA, the morphism $\mathrm{LV}(\iota)$ sends the conically parameterized dynamical system

$$\begin{cases} v_A(x; \rho) = -\rho_A\, x_A \\ v_B(x; \rho) = -\rho_B\, x_B \end{cases}, \qquad \rho \in \mathbb{R}_+^{\{A,B\}},$$

to the parameterized dynamical system

$$\begin{cases} v_A(x; \rho, \beta) = -\rho_A\, x_A - \beta_{BA}\, x_B\, x_A \\ v_B(x; \rho, \beta) = -\rho_B\, x_B + \beta_{AB}\, x_A\, x_B \end{cases}, \qquad \rho \in \mathbb{R}_+^{\{A,B\}},\ \beta \in \mathbb{R}_+^{\{AB,BA\}},$$

by setting the latter's interaction coefficients to zero: $\beta_{AB} = \beta_{BA} = 0$. This formalizes the commonsense fact that the first system is a degenerate case of the second.

For a more interesting example, we return to the projection map between regulatory networks given by Equation (1) of Section 2.1, inspired by the arginine biosynthesis system. Call this projection map p, and abbreviate the regulator molecule as R and the enzymes as $S := \{C, D, E, F, I\}$. The morphism $\mathrm{LV}(p)$ sends the parameterized dynamical system

$$\begin{cases} v_R(x; \rho, \beta) = -\rho_R\, x_R - \beta_R\, x_R^2 \\ v_C(x; \rho, \beta) = -\rho_C\, x_C - \beta_C\, x_R\, x_C \\ v_D(x; \rho, \beta) = -\rho_D\, x_D - \beta_D\, x_R\, x_D \end{cases} \qquad \begin{aligned} & v_E(x; \rho, \beta) = -\rho_E\, x_E - \beta_E\, x_R\, x_E \\ & v_F(x; \rho, \beta) = -\rho_F\, x_F - \beta_F\, x_R\, x_F \\ & v_I(x; \rho, \beta) = -\rho_I\, x_I - \beta_I\, x_R\, x_I \end{aligned}$$

with state variables $\{R\} + S$ and parameters $\rho, \beta \in \mathbb{R}_+^{\{R\}+S}$ to the parameterized dynamical system

$$\begin{cases} v_R(x; \rho, \beta) = -\rho_R\, x_R - \beta_R\, x_R^2 \\ v_*(x; \rho, \beta) = -\rho_*\, x_* - \beta_*\, x_R x_* \end{cases}$$

with state variables $\{R, *\}$ and parameters $\rho, \beta \in \mathbb{R}_+^{\{R,*\}}$, in two different but equivalent ways. The first way sets the latter system's coefficients equal to sums of the former's coefficients, namely

$$\rho_* = \sum_{i \in S} \rho_i \qquad \text{and} \qquad \beta_* = \sum_{i \in S} \beta_i.$$

The second way substitutes x_* for each x_i, $i \in S$, in the first system and then takes the vector field v_* to be the sum of the v_i's, $i \in S$, with these substitutions. The equivalence of these operations is precisely the condition for $\mathrm{LV}(p)$ to be a morphism of parameterized dynamical systems, cf. Equations (18) and (19).

3.3 Composing Lotka-Volterra models

To complete this part of the story, we extend the Lotka-Volterra dynamics functors between graphs and parameterized dynamical systems, constructed in Theorems 3.7 and 3.8, to *double* functors between open graphs and open parameterized dynamical systems. We begin by making parameterized dynamical systems into open systems. By "open systems," we mean dynamical systems that have specified interfaces along which they can share material with other systems.

Proposition 3.9 (Open parameterized dynamical systems)**.** *There is a symmetric monoidal double category of* open *linearly parameterized dynamical systems,* $\mathbb{O}\mathrm{pen}(\mathsf{Para}(\mathrm{Dynam}))$, *having*

- *as objects, finite sets $A, A', \ldots$;*
- *as vertical morphisms, functions $f : A \to A'$;*
- *as horizontal morphisms, **open linearly parameterized dynamical systems**, which consist of a linearly parameterized dynamical system $(P, S, v : \mathbb{R}^P \to \mathrm{Dynam}(S))$ along with a cospan $A_0 \xrightarrow{\ell_0} S \xleftarrow{\ell_1} A_1$ whose apex is the set S of state variables;*
- *as cells, **morphisms of such open systems** $(P, S, v, \ell_0, \ell_1) \to (P', S', v', \ell_0', \ell_1')$, which consist of a morphism $(q, f) : (P, S, v) \to (P', S', v')$ between linearly parameterized dynamical systems along with functions $f_0 : A_0 \to A_0'$ and $f_1 : A_1 \to A_1'$ making the diagram commute:*

$$\begin{array}{ccccc} A_0 & \xrightarrow{\ell_0} & S & \xleftarrow{\ell_1} & A_1 \\ {\scriptstyle f_0}\downarrow & & {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f_1} \\ A_0' & \xrightarrow[\ell_0']{} & S' & \xleftarrow[\ell_1']{} & A_1' \end{array}\ .$$

Vertical composition is by composition in FinSet and in $\mathsf{Para}(\mathrm{Dynam})$. Horizontal composition and monoidal products are by pushouts and coproducts in $\mathsf{Para}(\mathrm{Dynam})$, respectively, interpreting the finite sets in the feet of the cospans as linearly parameterized dynamical systems with no parameter variables and identically zero vector fields.

Similarly, there is a symmetric monoidal double category $\mathbb{O}\mathrm{pen}(\mathsf{Para}(\mathrm{Dynam}_+))$ of open conically *parameterized nonnegative dynamical systems.*

Proof. We do the construction for linearly parameterized dynamical systems. The construction for conically parameterized nonnegative dynamical systems is perfectly analogous, replacing $\mathbb{R}$ with $\mathbb{R}_+$ and vector spaces with conical spaces.

The projection functor $\pi_S : \mathsf{Para}(\mathrm{Dynam}) \to \mathsf{FinSet}$, $(P, S, v) \mapsto S$ that sends a linearly parameterized dynamical systems to its set of state variables has a left adjoint $Z : \mathsf{FinSet} \to \mathsf{Para}(\mathrm{Dynam})$ that sends a finite set S to the system $(\emptyset, S, 0)$ with empty set of parameter variables. By linearity, its parameterized vector field $0 \cong \mathbb{R}^\emptyset \to \mathrm{Dynam}(S)$ is necessarily the zero vector field. This indeed gives an adjunction $Z \dashv \pi_S$, because to any function $f : S \to S'$ and linearly

parameterized dynamical system (P', S', v') there corresponds a unique morphism $(0_{P'}, f) : Z(S) \to (P', S', v')$, where the compatibility square

$$
\begin{array}{ccc}
0 & \longrightarrow & \mathrm{Dynam}(S) \\
\downarrow & & \downarrow{\scriptstyle f_* \circ (-) \circ f^*} \\
\mathbb{R}^{P'} & \underset{v'}{\longrightarrow} & \mathrm{Dynam}(S')
\end{array}
$$

commutes trivially, since the zero vector space is initial in $\mathsf{Vect}_\mathbb{R}$.

Therefore, since $\mathsf{Para}(\mathrm{Dynam})$ has finite colimits (Proposition 3.5), we can construct a symmetric monoidal double category of Z-structured cospans [BC20, Theorem 3.9], which is isomorphic to $\mathbb{O}\mathrm{pen}(\mathsf{Para}(\mathrm{Dynam}))$ by arguments given before. $\qquad \square$

We can now construct double functors between open graphs and open parameterized dynamical systems, but the vertex parameters under Lotka-Volterra dynamics cause a twist in the story compared to Baez and Pollard's compositionality result for mass-action kinetics of open Petri nets [BP17, Theorem 18]. When composing open dynamical systems in the image of the Lotka-Volterra functor, one takes a coproduct of the parameter variables, i.e., a direct sum of the parameter spaces, belonging to identified vertices. However, if one composes the open graphs first, then the identified vertices receive a single copy of the parameters from the Lotka-Volterra functor. Thus this functor does not preserve composition of open systems, not even up to isomorphism. Nevertheless, there is a (noninvertible) comparison between the two: given a pair of vertex parameters in the direct sum, we can reduce them to a single parameter simply by summing them. In mathematical terms, we get a *lax* double functor: a double functor that strictly preserves vertical composition, as usual, but preserves horizontal composition only up to specified comparison cells.[7] While this laxness could be seen as a failure of compositionality, it is at most a very mild and well controlled failure. It is better regarded as a bookkeeping device for the vertex parameters.

Theorem 3.10 (Open Lotka-Volterra models). *There is a symmetric monoidal* lax *double functor*

$$
\mathrm{LV} : \mathbb{O}\mathrm{pen}(\mathsf{FinGraph}) \to \mathbb{O}\mathrm{pen}(\mathsf{Para}(\mathrm{Dynam}))
$$

that acts

- *on objects and vertical morphisms, as the identity;*
- *on horizontal morphisms and cells, by the functor* $\mathrm{LV} : \mathsf{FinGraph} \to \mathsf{Para}(\mathrm{Dynam})$ *on graphs and graph homomorphisms and as the identity on the associated cospans and cospan morphisms:*

$$
\left(X, A_0 \xrightarrow{\ell_0} X(V) \xleftarrow{\ell_1} A_1 \right) \mapsto \left(\mathrm{LV}(X), A_0 \xrightarrow{\ell_0} X(V) \xleftarrow{\ell_1} A_1 \right).
$$

The comparison cells are defined using the morphisms of linearly parameterized dynamical systems $\alpha_S : Z(S) \to \mathrm{LV}(\mathrm{Disc}\, S)$, *where*

$$
\alpha_S := (0_S, 1_S) : (\emptyset, S, 0) \to (S, S, \vec{\mathrm{LV}}(\mathrm{Disc}\, S)), \qquad S \in \mathsf{FinSet}.
$$

- *Given composable open graphs* $(X, A \to X(V) \leftarrow B)$ *and* $(Y, B \to Y(V) \leftarrow C)$, *the comparison cell for horizontal composition is given by the morphism of systems*

$$
\mathrm{LV}(X) +_{Z(B)} \mathrm{LV}(Y) \xrightarrow{\mathrm{id} +_{\alpha_B} \mathrm{id}} \mathrm{LV}(X) +_{\mathrm{LV}(\mathrm{Disc}\, B)} \mathrm{LV}(Y) \xrightarrow{\cong} \mathrm{LV}(X +_{\mathrm{Disc}\, B} Y).
$$

- *Given a finite set* A, *the comparison cell for the horizontal unit is given by the morphism of systems* $\alpha_A : Z(A) \to \mathrm{LV}(\mathrm{Disc}\, A)$.

Similarly, there is a symmetric monoidal lax *double functor*

$$
\mathrm{LV} : \mathbb{O}\mathrm{pen}(\mathsf{FinSgnGraph}) \to \mathbb{O}\mathrm{pen}(\mathsf{Para}(\mathrm{Dynam}_+)).
$$

---

[7]The precise definition of a lax double functor can be found, for instance, in the textbook [Gra19, §3.5].

Proof. To construct the lax double functor, we use [Pat23, Theorem 2.4], which extends the construction of a double functor in [BC20, Theorem 4.3] from the pseudo to the lax case. The family of morphisms $\alpha_S : Z(S) \to \mathrm{LV}(\mathrm{Disc}\, S)$, $S \in \mathsf{FinSet}$, in the theorem statement assemble into a natural transformation

$$
\begin{array}{ccc}
\mathsf{FinSet} & \xrightarrow{\ L=\mathrm{Disc}\ } & \mathsf{FinGraph} \\
\Big\| & \overset{\alpha}{\Longrightarrow\!\!\nearrow} & \Big\downarrow{\scriptstyle \mathrm{LV}} \\
\mathsf{FinSet} & \xrightarrow[\ L'=Z\ ]{} & \mathsf{Para}(\mathsf{Dynam})
\end{array}
\qquad .
$$

The functors involved in this cell all preserve finite colimits: the top and bottom ones because they are left adjoints and the right one by Theorem 3.7. By [Pat23, Theorem 2.4], we obtain a symmetric monoidal lax double functor

$$
\mathbb{O}\mathsf{pen}(\mathsf{FinGraph}) \cong {}_L\mathbb{C}\mathsf{sp}(\mathsf{FinGraph}) \to {}_{L'}\mathbb{C}\mathsf{sp}(\mathsf{Para}(\mathsf{Dynam})) \cong \mathbb{O}\mathsf{pen}(\mathsf{Para}(\mathsf{Dynam})).
$$

To see that this double functor is the same one in the theorem statement, we once again use the adjunctions to pass between L-structured and R-decorated cospans (recalling terminology introduced in the proof of Proposition 2.4). Notice that the natural transformation α has as its mate [CGR14, §1] the identity transformation $\bar{\alpha} = 1_{\mathrm{ev}_V}$:

$$
\begin{array}{ccc}
\mathsf{FinSet} & \xleftarrow{\ R=\mathrm{ev}_V\ } & \mathsf{FinGraph} \\
\Big\| & \overset{\bar{\alpha}}{\searrow\!\!\Longrightarrow} & \Big\downarrow{\scriptstyle \mathrm{LV}} \\
\mathsf{FinSet} & \xleftarrow[\ R'=\pi_S\ ]{} & \mathsf{Para}(\mathsf{Dynam})
\end{array}
\qquad .
$$

Thus the action of the double functor $F := \mathrm{LV}$ on L-structured cospans simplifies to the identity when translated to R-decorated cospans.

$$
\begin{array}{ccccc}
 & L(A_0) & \xrightarrow{\ell_0} & X & \xleftarrow{\ell_1} & L(A_1) \\
 & & & \big\downarrow & & \\
L'(A_0) \xrightarrow{\alpha_{A_0}} F(L(A_0)) & \xrightarrow{F(\ell_0)} & F(X) & \xleftarrow{F(\ell_1)} & F(L(A_1)) \xleftarrow{\alpha_{A_1}} L'(A_1)
\end{array}
$$

$$
\begin{array}{c} \rotatebox{90}{$\xrightarrow{\ \ \ }$} \\ \end{array}
$$

$$
\begin{array}{ccccc}
 & A_0 & \xrightarrow{\bar{\ell}_0} & R(X) & \xleftarrow{\bar{\ell}_1} & A_1 \\
 & & & \big\| & & \\
A_0 \xrightarrow{\bar{\ell}_0} R(X) & \xrightarrow{\bar{\alpha}_X} & R'(F(X)) & \xleftarrow{\bar{\alpha}_X} & R(X) \xleftarrow{\bar{\ell}_1} A_1
\end{array}
$$

A similar statement holds for the action of the double functor on morphisms of L-structured and R-decorated cospans. □

4 Conclusion

Summary. Regulatory networks are a minimalistic but widely used tool to describe the interactions between molecules in biochemical systems. We have made the first functorial study of regulatory networks, formalized as signed graphs, and their connections with other mathematical models in biochemistry. Among such models, we have studied reaction networks, formalized as Petri nets with signed links, and parameterized dynamical systems, focusing on Lotka-Volterra dynamics.

The major categories of this paper, and the functors between them, are summarized in the

following diagram, where "LV" is the Lotka-Volterra dynamics functor (§3.2).

$$\mathsf{SgnCat}\ (\S 2.2) \quad \perp \quad \mathsf{SgnGraph}\ (\S 2.1) \xleftarrow{\ \mathrm{Int}\ } \mathsf{SgnPetri}\ (\S 2.3)$$

with Path and U between SgnCat and $\mathsf{SgnGraph}$, and

$$\mathsf{FinSgnGraph} \xrightarrow{\ \mathrm{LV}\ } \mathsf{Para}(\mathsf{Dynam}_+)\ (\S 3.1)$$

Most of the main results extend from closed systems to open systems, which compose by gluing along their boundaries. Of the diagram above, we have extended the following parts to double categories of open systems and double functors between them. The Lotka-Volterra double functor is lax; the others are pseudo.

$$\mathbb{O}\mathrm{pen}(\mathsf{SgnCat}) \xleftarrow{\ \mathrm{Path}\ } \mathbb{O}\mathrm{pen}(\mathsf{SgnGraph})$$

$$\mathbb{O}\mathrm{pen}(\mathsf{FinSgnGraph}) \xrightarrow{\ \mathrm{LV}\ } \mathbb{O}\mathrm{pen}(\mathsf{Para}(\mathsf{Dynam}_+))\ (\S 3.3)$$

Outlook. Of many possible directions for future work, we mention a few. As noted in the introduction, Lotka-Volterra dynamics are only one of numerous dynamics that could be considered as a canonical model for regulatory networks, and they are not even among the most commonly studied in the biochemistry literature [TLK19]. It would be desirable to have dynamics functors for regulatory networks that draw on more flexible or more biologically plausible classes of dynamical systems. In another direction, the two halves of this paper—qualitative and quantitative—are not as tightly as integrated as one might hope. How does the presence of a motif in a regulatory network, such as an incoherent feedforward loop perhaps even of a specific type, manifest in the continuous dynamics of that network? Put in category-theoretic terms, the Lotka-Volterra dynamics functor is defined on signed graphs; how does it relate to the freely generated signed categories in which motifs are expressed? These intriguing questions are suggestive of "feedback loop analysis" in the field of system dynamics [Ric95], to which stronger connections should be made.

This project fits into a broader program by applied category theorists and other scientists that aims to systematize, in a completely precise way, the language and methods of describing, comparing, and composing scientific models in different domains. Within biology, the field of systems biology has advocated for a holistic view of complex biological systems that emphasizes composition as much as reduction. We believe that category theory has a role to play in this endeavor by bringing mathematical precision to compositional and structural aspects of modeling that are traditionally thought to be outside the realm of mathematics.

References

[AF73]   Tilak Agerwala and Mike Flynn. "Comments on capabilities, limitations and "correctness" of Petri nets". *Proceedings of the 1st Annual Symposium on Computer Architecture*. 1973, pp. 81–86. DOI: https://doi.org/10.1145/800123.803973.

[Alo07]   Uri Alon. "Network motifs: theory and experimental approaches". *Nature Reviews Genetics* 8.6 (2007), pp. 450–461. DOI: https://doi.org/10.1038/nrg2102.

[Alo19]   Uri Alon. *An introduction to systems biology: design principles of biological circuits*. 2nd ed. CRC Press, 2019. DOI: https://doi.org/10.1201/9780429283321.

[Bae+21] John C. Baez, Fabrizio Genovese, Jade Master, and Michael Shulman. "Categories of nets". *2021 Symposium on Logic in Computer Science (LICS)*. 2021, pp. 1–13. DOI: https://doi.org/10.1109/LICS52264.2021.9470566. arXiv: 2101.04238.

[Bae+22] John Baez, Xiaoyan Li, Sophie Libkind, Nathaniel Osgood, and Evan Patterson. "Compositional modeling with stock and flow diagrams". *Proceedings of the Fifth International Conference on Applied Category Theory (ACT 2022)*. 2022. DOI: https://doi.org/10.4204/EPTCS.380.5. arXiv: 2205.08373.

[Bal+04]   Paolo Baldan, Nadia Busi, Andrea Corradini, and G. Michele Pinna. "Domain and event structure semantics for Petri nets with read and inhibitor arcs". *Theoretical Computer Science* 323.1-3 (2004), pp. 129–189. DOI: https://doi.org/10.1016/j.tcs.2004.04.001.

[BC20]   John C. Baez and Kenny Courser. "Structured cospans". *Theory and Applications of Category Theory* 35.48 (2020), pp. 1771–1822. arXiv: 1911.04630. URL: http://www.tac.mta.ca/tac/volumes/35/48/35-48abs.html.

[BFM19]   John C. Baez, John Foley, and Joe Moeller. "Network models from Petri nets with catalysts". *Compositionality* 1.4 (2019). DOI: https://doi.org/10.32408/compositionality-1-4. arXiv: 1904.03550.

[BP17]   John C. Baez and Blake S. Pollard. "A compositional framework for reaction networks". *Reviews in Mathematical Physics* 29.9 (2017), p. 1750028. DOI: https://doi.org/10.1142/S0129055X17500283.

[Car71]   Henri Cartan. *Differential calculus*. Kershaw Publishing Company, 1971.

[CGR14]   Eugenia Cheng, Nick Gurski, and Emily Riehl. "Cyclic multicategories, multivariable adjunctions and mates". *Journal of K-Theory* 13.2 (2014), pp. 337–396. DOI: https://doi.org/10.1017/is013012007jkt250. arXiv: 1208.4520.

[CP09]   Stephen T. Crews and Joseph C. Pearson. "Transcriptional autoregulation in development". *Current Biology* 19.6 (2009), R241–R246. DOI: https://doi.org/10.1016/j.cub.2009.01.015.

[Csi+09]   Attila Csikász-Nagy, Orsolya Kapuy, Attila Tóth, Csaba Pál, Lars J. Jensen, Frank Uhlmann, John J. Tyson, and Béla Novák. "Cell cycle regulation by feed-forward loops coupling transcription and phosphorylation". *Molecular Systems Biology* 5 (2009), pp. 236–241. DOI: https://doi.org/10.1038/msb.2008.73.

[For61]   Jay W. Forrester. *Industrial dynamics*. MIT Press, 1961.

[FS07]   José Luiz Fiadeiro and Vincent Schmitt. "Structured co-spans: an algebra of interaction protocols". *International Conference on Algebra and Coalgebra in Computer Science (CALCO 2007)*. 2007, pp. 194–208. DOI: https://doi.org/10.1007/978-3-540-73859-6_14.

[Gra19]   Marco Grandis. *Higher dimensional categories: From double to multiple categories*. World Scientific, 2019. DOI: https://doi.org/10.1142/11406.

[Hac76]   Michel Hack. *Petri net languages*. Tech. rep. 159. Massachusetts Institute of Technology, Mar. 1976. URL: https://hdl.handle.net/1721.1/149453.

[HCH10]   Wassim M. Haddad, VijaySekhar Chellaboina, and Qing Hui. *Nonnegative and compartmental dynamical systems*. Princeton University Press, 2010.

[Ing13]   Brian P. Ingalls. *Mathematical modeling in systems biology: An introduction*. MIT Press, 2013.

[Keu+11]   Joost J. B. Keurentjes, Gerco C. Angenent, Marcel Dicke, Vítor A. P. Martins Dos Santos, Jaap Molenaar, Wim H. van der Putten, Peter C. de Ruiter, Paul C. Struik, and Bart P. H. J. Thomma. "Redefining plant systems biology: from cell to ecosystem". *Trends in Plant Science* 16.4 (2011), pp. 183–190. DOI: https://doi.org/10.1016/j.tplants.2010.12.002.

[Koc22]   Joachim Kock. "Whole-grain Petri nets and processes". *Journal of the ACM* 70.1 (2022), pp. 1–58. DOI: https://doi.org/10.1145/3559103. arXiv: 2005.05108.

[Lei04]   Tom Leinster. *Higher operads, higher categories*. Cambridge University Press, 2004. DOI: https://doi.org/10.1017/CBO9780511525896.

[Lot25]   Alfred J. Lotka. *Elements of physical biology*. Williams & Wilkins, 1925. URL: https://archive.org/details/elementsofphysic017171mbp.

[MM94]   Saunders Mac Lane and Ieke Moerdijk. *Sheaves in geometry and logic: A first introduction to topos theory*. Springer, 1994. DOI: https://doi.org/10.1007/978-1-4612-0927-0.

[Pat20]   Evan Patterson. "The algebra and machine representation of statistical models". PhD thesis. Stanford University, 2020. arXiv: 2006.08945.

[Pat23]   Evan Patterson. "Structured and decorated cospans from the viewpoint of double category theory". *Proceedings of the Sixth International Conference on Applied Category Theory (ACT 2023)*. 2023. DOI: https://doi.org/10.4204/EPTCS.397.13. arXiv: 2304.00447.

[RB88]   David E. Rydeheard and Rod M. Burstall. *Computational category theory*. Prentice Hall, 1988. URL: https://www.cs.man.ac.uk/~david/categories/.

[Ric95]   George P. Richardson. "Loop polarity, loop dominance, and the concept of dominant polarity". *System Dynamics Review* 11.1 (1995), pp. 67–88. DOI: https://doi.org/10.1002/sdr.4260110106.

[Rie16]   Emily Riehl. *Category theory in context*. Courier Dover Publications, 2016. URL: http://www.math.jhu.edu/~eriehl/context.pdf.

[SMH18]   Gabor Szederkényi, Attila Magyar, and Katalin M. Hangos. *Analysis and control of polynomial dynamic models with biological applications*. Academic Press, 2018. DOI: https://doi.org/10.1016/C2015-0-06035-X.

[Spi21]   David I. Spivak. "Functorial aggregation" (2021). arXiv: 2111.10968.

[SRS14]   S. Stanhope, Jonathan E. Rubin, and David Swigon. "Identifiability of linear and linear-in-parameters dynamical systems from a single trajectory". *SIAM Journal on Applied Dynamical Systems* 13.4 (2014), pp. 1792–1815. DOI: https://doi.org/10.1137/130937913.

[Ste00]   John D. Sterman. *Business dynamics: systems thinking and modeling for a complex world*. McGraw-Hill, 2000.

[Str00]   Ross Street. "The petit topos of globular sets". *Journal of Pure and Applied Algebra* 154.1-3 (2000), pp. 299–315. DOI: https://doi.org/10.1016/S0022-4049(99)00183-8.

[TLK19]   John J. Tyson, Teeraphan Laomettachit, and Pavel Kraikivski. "Modeling the dynamic behavior of biochemical regulatory networks". *Journal of Theoretical Biology* 462 (2019), pp. 514–527. DOI: https://doi.org/10.1016/j.jtbi.2018.11.034.

[TN10]   John J. Tyson and Béla Novák. "Functional motifs in biochemical reaction networks". *Annual Review of Physical Chemistry* 61 (2010), pp. 219–240. DOI: https://doi.org/10.1146/annurev.physchem.012809.103457.

[Voi00]   Eberhard O. Voit. *Computational analysis of biochemical systems: A practical guide for biochemists and molecular biologists*. Cambridge University Press, 2000.

[Voi13]   Eberhard O. Voit. "Biochemical systems theory: a review". *International Scholarly Research Notices* 2013 (2013). DOI: https://doi.org/10.1155/2013/897658.

[Voi18]   Eberhard O. Voit. *A first course in systems biology*. 2nd ed. Garland Science, 2018.