Rate-Independent Computation in Continuous Chemical Reaction Networks

Ho-Lin Chen
National Taiwan University
MD-718
Taipei 106, Taiwan
holinc@gmail.com

David Doty
California Institute of Technology
1200 E. California Blvd
MC 305-16
Pasadena, CA 91125
ddoty@caltech.edu

David Soloveichik
UCSF Center for Systems and Synthetic Biology
1700 4th Street
Byers Hall 401, Box 2540
San Francisco, CA 94158
david.soloveichik@ucsf.edu

ABSTRACT

Understanding the algorithmic behaviors that are *in principle* realizable in a chemical system is necessary for a rigorous understanding of the design principles of biological regulatory networks. Further, advances in synthetic biology herald the time when we'll be able to rationally engineer complex chemical systems, and when idealized formal models will become blueprints for engineering.

Coupled chemical interactions in a well-mixed solution are commonly formalized as chemical reaction networks (CRNs). However, despite the widespread use of CRNs in the natural sciences, the range of computational behaviors exhibited by CRNs is not well understood. Here we study the following problem: what functions $f : \mathbb{R}^k \to \mathbb{R}$ can be computed by a chemical reaction network, in which the CRN eventually produces the correct amount of the "output" molecule, no matter the rate at which reactions proceed? This captures a previously unexplored, but very natural class of computations: for example, the reaction $X_1+X_2 \to Y$ can be thought to compute the function $y = \min(x_1, x_2)$. Such a CRN is robust in the sense that it is correct whether its evolution is governed by the standard model of mass-action kinetics, alternatives such as Hill-function or Michaelis-Menten kinetics, or other arbitrary models of chemistry that respect the (fundamentally digital) stoichiometric constraints (what are the reactants and products?). We develop a formal definition of such computation using a novel notion of reachability, and prove that a function is computable in this manner if and only if it is *continuous piecewise linear*.

Categories and Subject Descriptors

F.1 [**Theory of Computation**]: Computation by Abstract Devices; C.1.3 [**Computer systems organization**]: Other Architecture Styles—*Analog computers*

ITCS'14, January 12–14, 2014, Princeton, New Jersey, USA.

ACM 978-1-4503-2698-8/14/01 ...$15.00.
http://dx.doi.org/10.1145/2554797.2554827.

Keywords

Chemical Reaction Networks, Mass-Action, Analog Computation, Piecewise-Linear

1. INTRODUCTION

Both the engineering of complex artificial molecular systems, as well as the understanding of the constraints imposed upon biology, require the understanding of what is, in principle, achievable in chemistry. The natural language for describing the interactions of molecular species in a well-mixed solution is that of chemical reaction networks (CRNs), i.e., finite sets of chemical reactions such as $A+B \to A+C$. The intuitive meaning of this expression is that a unit of chemical species A reacts with a unit of chemical species B, producing a unit of a new chemical species C and regenerating a unit of A back. Typically (in mass-action kinetics) the rate with which this occurs is proportional to the product of the amounts of the reactants A and B.

Traditionally CRNs have been used as a descriptive language to analyze naturally occurring chemical reactions, as well as various other systems with a large number of interacting components such as gene regulatory networks and animal populations. However, CRNs also constitute a natural choice of programming language for engineering artificial systems. For example, nucleic-acid networks can be rationally designed to implement arbitrary chemical reaction networks [7, 10, 20]. Thus, since in principle any CRN can be physically built, hypothetical CRNs with interesting behaviors are becoming of more than theoretical interest. One day artificial CRNs may underlie embedded controllers for biochemical, nanotechnological, or medical applications, where environments are inherently incompatible with traditional electronic controllers. However, to effectively program chemistry, we must understand the computational power at our disposal. In turn, the computer science approach to CRNs is also beginning to generate novel insights regarding natural cellular regulatory networks [8].

Informally speaking we can identify two sources of computational power in CRNs. First, the reaction *stoichiometry* transforms some specific ratios of reactants to products. For example, $X \to 2Y$ makes two units of Y for every unit of X. Second, in mass-action kinetics the reaction *rate laws* effectively perform multiplication of the reactant concentrations. In this work, we seek to disentangle the contributions of these two computational ingredients by focusing on the computational power of stoichiometry alone.

One reason to focus on *stoichiometric computation* is that algorithms that rely on stoichiometry make easier design targets. The rates of reactions are real-valued quantities that can fluctuate with reaction conditions and temperature, while the stoichiometries are immutable whole numbers set by the nature of the reaction. Methods for physically implementing CRNs naturally yield systems with digital stoichiometry that can be set exactly [7,20]. Further, relying on specific rate laws can be problematic: many systems do not follow mass-action rate laws[1] and chemists have developed an array of alternative rate laws such as Michaelis-Menten and Hill-function kinetics. Indeed, robustness of rate laws is a recurring motif in systems biology due to much evidence that biological regulatory networks tend to be robust to the form of the rate laws and the rate parameters [5]. Thus we are interested in what computations can be understood or engineered without regard for the reaction rate laws.

There are two well-studied models of chemical kinetics: *continuous* and *discrete*. In the discrete model, the amount of a species is a nonnegative *integer* representing the total count of molecules of that species in a given reaction vessel. In the continuous model, the amount of a species is a nonnegative *real number* representing its average count per unit volume or concentration. The discrete model is *stochastic* and reactions are modeled by a Markov jump process [11], while the continuous model is *deterministic*, governed by a system of ordinary differential equations with a unique solution. When the volume and counts are large, the discrete model converges to the continuous model [14]. While many cellular chemical processes operate at single-molecule precision, a significant amount of regulation is well-understood by continuous models [1]. Further, because of the difficulty of working at molecular resolution most experimental implementations of rationally designed chemical computation have been in bulk solution (e.g. [10,16,18]).

Here we study the continuous setting, and characterize the class of real-valued functions computable by CRNs when reaction rates are permitted to vary arbitrarily (and possibly adversarially) over time. Any computation in this setting must rely on stoichiometry alone. Our work is related to the study of *deterministic*[2] *computation* in stochastic CRNs: making reaction rates unreliable is the natural analog to error in the stochastic setting (where error corresponds to reactions occurring in an undesired order). (See Section 1.1.)

How can rate laws "preserve stoichiometry" while varying "arbitrarily over time"? Formally, preserving stoichiometry means that if we reach state $\mathbf{d}$ from state $\mathbf{c}$, then $\mathbf{d} = \mathbf{c}+\mathbf{M}\mathbf{u}$ for stoichiometry matrix $\mathbf{M}$, and some non-negative vector $\mathbf{u}$ of reaction fluxes. Subject to this constraint, the widest class of trajectories that still satisfies the intuitive meaning of the reaction semantics can be described informally as follows: (1) concentrations can't become negative, (2) all reactants must be non-zero for a reaction to occur (e.g. if a reaction uses a catalyst[3], it must be present).

[1]Although it is generally taken for granted that they would if properly decomposed into truly elementary reactions.

[2]Deterministic computation in the discrete/stochastic model should not be confused with the determinism of the continuous model, where any stochasticity is absent.

[3]A species acts catalytically in a reaction if it is both a reactant and product: e.g. C in reaction $A + C \to B + C$. Note that executing this reaction without C does not by itself violate condition (1).

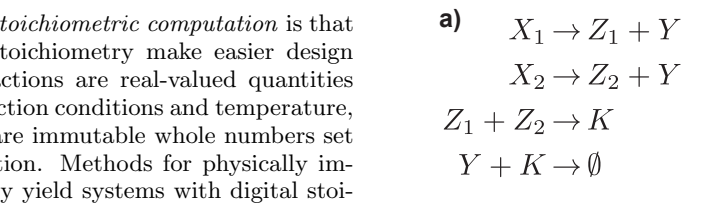

Figure 1: Examples of (a) *direct* and (b) *dual-rail* rate-independent computation of $f(x_1, x_2) = \max(x_1, x_2)$.

The example shown in Fig 1(a) illustrates the style of computation studied here. Let $f : \mathbb{R}^2_{\geq 0} \to \mathbb{R}_{\geq 0}$ be the max function $f(x_1, x_2) = \max(x_1, x_2)$ restricted to non-negative x_1 and x_2. The CRN shown computes this function in the following sense. Inputs x_1 and x_2 are given as initial concentrations of input species X_1 and X_2. Then the CRN converges to the output value of species Y, under a very wide interpretation of rate laws. Intuitively, the first two reactions must eventually produce $x_1 + x_2$ of Y, and x_1, x_2 of Z_1 and Z_2, respectively. This is enforced by the stoichiometric constraint that the amount of Z_1 and Y produced is equal to the amount of X_1 consumed (and analogously for the second reaction). Stoichiometric constraints require the third reaction to produce the amount of K that is the minimum of the amount of Z_1 and Z_2 eventually produced in the first two reactions. Thus $\min(x_1, x_2)$ of K is eventually produced. Therefore, the fourth reaction eventually consumes $\min(x_1, x_2)$ molecules of Y leaving $x_1 + x_2 - \min(x_1, x_2) = \max(x_1, x_2)$ of Y behind. We can imagine an adversary pushing flux through these four reactions in any devious stratagem, but as long as he cannot take any concentration negative, the CRN can only converge to the correct output.

In this paper we further consider the natural extension of such computation to handle negative real values. The example shown in Fig. 1(b) computes $f(x_1, x_2) = \max(x_1, x_2)$ ($f : \mathbb{R}^2 \to \mathbb{R}$). In order to handle negative input and output values, we represent the value of each input and output by a pair of species (so-called "dual-rail" representation). For example, in state $\mathbf{c}$, $x_1 = \mathbf{c}(X_1^+) - \mathbf{c}(X_1^-)$ — i.e. the difference between the concentrations of species X_1^+ and X_1^-. Note that when X_1^- and X_2^- are initially absent, the CRN becomes equivalent to the first three reactions of Fig. 1(a) under relabeling of species. We do not need the last reaction of (a) because the output is represented as the difference of Y^+ and Y^- by our convention. For the argument that the computation is correct even if X_1^- and X_2^- are initially present, we refer the reader to Section 3.1.

In addition to handling negative values, the dual-rail representation has the benefit of allowing composition. Specifically, the dual-rail representation allows CRNs to never consume their output species (e.g. rather than consuming Y^+, it can produce Y^-). This monotonicity in the production of output allows directly composing CRN computations simply by mixing CRNs and relabeling species (e.g. to make the output of one be input to the other). Since the upstream CRN never consumes its output species, the downstream CRN is free to consume them without interfering with the upstream computation.

In order to formally delineate the class of functions (direct and dual-rail) computable in a rate-independent manner, we take the following approach. We first define a reachability

relation that captures motion along trajectories satisfying the two intuitive properties above. Then we define rate-independent computation using this reachability relation. Roughly, to say that a function is correctly computed, it must be the case that from every reachable state, the correct output can be reached, and once reached the output cannot change. Our main results are that exactly functions that are positive-continuous, piecewise linear (direct) or continuous, piecewise linear (dual-rail) can be computed. (Positive-continuous means that the only discontinuities occur on a "face" of $\mathbb{R}^k_{\geq 0}$ — i.e., the function may discontinuously jump only at a point where some input goes from 0 to positive.)

In the constructive portion of our results, we supply rate-independent CRNs for computing any function in the above classes. Further, our constructions have the following property: If the same CRNs were simulated under mass-action kinetics from any reachable state, as the time $t \to \infty$ the concentration of the output species would approach its intended value. In other words, no matter how an adversary might have pushed us initially, letting the system evolve by mass-action will yield the correct answer. Although we prove that only the correct output is reachable, and mass-action trajectories converge there, it remains an open problem to delineate the entire class of rate laws that force our constructions to converge.

1.1 Relation to Previous Work

The computational abilities of discrete CRNs have been investigated more thoroughly than of continuous CRNs, and have been shown to have a surprisingly rich computational structure. Of most relevance here is the work in the discrete setting showing that the class of functions that can be computed depends strongly on whether the computation must be deterministic (guaranteed to be correct), or just likely to be correct. While Turing universal computation is possible with an arbitrarily small, non-zero probability of error over all time [19], forbidding error altogether limits the computational power severely: Error-free computation by stochastic CRNs is limited to semilinear predicates and functions [4,9]. (Intuitively, semilinear functions are expressible as a finite union of affine functions, with "simple, periodic" domains of each affine function [9].)

Our paper was motivated by trying to extend the results on error-free computation to the continuous CRN model. As mentioned above, our notion of rate-independent computation is the natural extension of deterministic computation in the discrete model. However, there are many differences in the two settings. As broached in ref. [9], many of the CRNs that work in the discrete setting appear not to work in the continuous setting. For example, the CRNs computing discontinuous functions such as "$f(x_1, x_2) = x_2$ if $x_1 > x_2$ and 0 otherwise" provably fail to work in the continuous setting. Indeed, discrete CRNs can, for example, distinguish between even and odd inputs, whereas it does not even make sense to talk about the "parity" of a real-valued input. Further, the proof techniques appear to require very different machinery.

The relationship between the discrete and continuous CRN models is a complex and much studied one in the natural sciences [17]. While computational differences in the models are less understood, the distributed computing community is beginning to investigate the subject as well [6].

Our notion of reachability captures a wide diversity of possible rate laws. A related idea in the literature, albeit restricted to mass-action, is differential inclusion [12]. Generalized rate laws (extending mass-action, Michaelis-Menten, etc) have been previously studied, although not in a computational setting. For example, certain conditions were identified on global convergence to equilibrium based on properties intuitively similar to ours [2].

2. PRELIMINARIES

Let $\mathbb{N}$ and $\mathbb{R}$ denote the set of nonnegative integers and the set of real numbers, respectively. If $x \in \mathbb{R}$, let $\mathbb{R}_{\geq x} = \{ x' \in \mathbb{R} \mid x' \geq x \}$, and similarly for $\mathbb{R}_{>x}$. Given a vector $\mathbf{x} \in \mathbb{R}^k$, let $\|\mathbf{x}\| = \|\mathbf{x}\|_1 = \sum_{i=1}^k \mathbf{x}(i)$, where $\mathbf{x}(i)$ denotes the ith coordinate of $\mathbf{x}$. We abuse notation and consider the sets $\mathbb{R}^k \times \mathbb{R}^l$ and $\mathbb{R}^{k+l}$ to be the same, because it is sometimes convenient to treat an ordered pair of vectors as being concatenated into a single longer vector. If Λ is a finite set (in this paper, of chemical species), and A is any set, we write A^Λ to denote the set of functions $f : \Lambda \to A$. Equivalently, we view an element $\mathbf{c} \in A^\Lambda$ as a vector of $|\Lambda|$ elements of A, each coordinate "labeled" by an element of Λ. Write $\mathbf{c} \upharpoonright \Delta$ to denote the vector $\mathbf{d} \in \mathbb{N}^\Delta$ such that $\mathbf{c}(S) = \mathbf{d}(S)$ for all $S \in \Delta$.

2.1 Chemical reaction networks

Given $S \in \Lambda$ and $\mathbf{c} \in \mathbb{R}^\Lambda_{\geq 0}$, we refer to $\mathbf{c}(S)$ as the *concentration of S in $\mathbf{c}$*. For any $\mathbf{c} \in \mathbb{R}^\Lambda_{\geq 0}$, let $[\mathbf{c}] = \{S \in \Lambda \mid \mathbf{c}(S) > 0\}$, the set of species *present* in $\mathbf{c}$. We write $\mathbf{c} \leq \mathbf{c}'$ to denote that $\mathbf{c}(S) \leq \mathbf{c}'(S)$ for all $S \in \Lambda$. Given $\mathbf{c}, \mathbf{c}' \in \mathbb{R}^\Lambda_{\geq 0}$, we define the vector component-wise operations of addition $\mathbf{c} + \mathbf{c}'$, subtraction $\mathbf{c} - \mathbf{c}'$, and scalar multiplication $x\mathbf{c}$ for $x \in \mathbb{R}$. If $\Delta \subset \Lambda$, we view a vector $\mathbf{c} \in \mathbb{R}^\Delta_{\geq 0}$ equivalently as a vector $\mathbf{c} \in \mathbb{R}^\Lambda_{\geq 0}$ by assuming $\mathbf{c}(S) = 0$ for all $S \in \Lambda \setminus \Delta$. For $\Delta \subset \Lambda$, we write $\mathbf{c} \upharpoonright \Delta$ to denote $\mathbf{c}$ *restricted to* Δ; in particular, $\mathbf{c} \upharpoonright \Delta = \mathbf{0} \iff (\forall S \in \Delta)\ \mathbf{c}(S) = 0$.

Given a finite set of chemical species Λ, a *reaction* over Λ is a pair $\alpha = \langle \mathbf{r}, \mathbf{p} \rangle \in \mathbb{N}^\Lambda \times \mathbb{N}^\Lambda$, specifying the stoichiometry of the reactants and products, respectively.[4] In this paper, we assume that $\mathbf{r} \neq \mathbf{0}$, i.e., we have no reactions of the form $\varnothing \to \ldots$.[5] For instance, given $\Lambda = \{A, B, C\}$, the reaction $A + 2B \to A + 3C$ is the pair $\langle (1, 2, 0), (1, 0, 3) \rangle$. Note that we represent reversible reactions such as $A \rightleftharpoons B$ as two irreversible reactions $A \to B$ and $B \to A$. A *(finite) chemical reaction network (CRN)* is a pair $\mathcal{C} = (\Lambda, R)$, where Λ is a finite set of chemical *species*, and R is a finite set of reactions over Λ. A *state* of a CRN $\mathcal{C} = (\Lambda, R)$ is a vector $\mathbf{c} \in \mathbb{R}^\Lambda_{\geq 0}$. Given a state $\mathbf{c}$ and reaction $\alpha = \langle \mathbf{r}, \mathbf{p} \rangle$, we say that α is *applicable* in $\mathbf{c}$ if $[\mathbf{r}] \subseteq [\mathbf{c}]$ (i.e., $\mathbf{c}$ contains positive concentration of all of the reactants).

2.2 Reachability

[4] It is customary to define, for each reaction, a *rate constant* $k \in \mathbb{R}_{>0}$ specifying a constant multiplier on the mass-action rate (i.e., the product of the reactant concentrations), but as we are studying CRNs whose output is independent of the reaction rates, we leave the rate constants out of the definition.

[5] We allow high order reactions; i.e., those that consume more than two reactants. It is not difficult to show that such a reaction $S_1 + S_2 + \ldots + S_n \to P_1 + \ldots + P_m$ can be replaced by bimolecular reactions $S_1 + S_2 \rightleftharpoons S_{12}, S_{12} + S_3 \rightleftharpoons S_{123}, S_{123} + S_4 \rightleftharpoons S_{1234}, \ldots, S_n + S_{12\ldots n-1} \to P_1 + \ldots + P_m$ without affecting the correctness of the CRN.

In the previous section we defined the syntax of CRNs. We wish to define a semantic interpretation of them suitable for defining how they compute functions. In particular, we want to consider CRNs guaranteed to correctly compute a function of their inputs, regardless of the rate at which reactions proceed, calling this *stable computation*, defined in Section 2.4. This definition rests on another, which we must take care in defining, namely what it means for one state to be *reachable* from another, which is the focus of this section. Intuitively, $\mathbf{d}$ is reachable from $\mathbf{c}$ if applying some amount of reactions to $\mathbf{c}$ results in $\mathbf{d}$, such that no reaction is ever applied when any of its reactants are concentration 0. Formalizing this concept is a bit tricky and constitutes one of the contributions of this paper. Intuitively, we'll think of reachability via straight line segments. This may appear overly limiting; after all mass-action and other rate laws trace out smooth curves. However we show a number of properties of our definition that support its reasonableness.

Throughout this section, fix a CRN $\mathcal{C} = (\Lambda, R)$. All states $\mathbf{c}$, etc., are assumed to be states of $\mathcal{C}$. Let $m = |R|$ be the number of reactions in CRN $\mathcal{C}$, and let $n = |\Lambda|$ be the number of species in $\mathcal{C}$. We define the the $n \times m$ *reaction stoichiometry matrix* $\mathbf{M}$ such that $\mathbf{M}(i, j)$ is the net amount of the i'th species that is produced by the j'th reaction (negative if the species is consumed).[6] For example, if we have the reactions $X \to Y$ and $X + A \to 2X + 3Y$, and if the three rows correspond to X, A, and Y, in that order, then

$$\mathbf{M} = \begin{pmatrix} -1 & 1 \\ 0 & -1 \\ 1 & 3 \end{pmatrix}$$

Definition 2.1. *State* $\mathbf{d}$ *is* straight-line reachable (aka 1-segment reachable) *from* $\mathbf{c}$*, written* $\mathbf{c} \to^1 \mathbf{d}$*, if* $(\exists \mathbf{u} \in \mathbb{R}^m_{\geq 0})$ $\mathbf{c} + \mathbf{M}\mathbf{u} = \mathbf{d}$ *and* $\mathbf{u}(j) > 0$ *only if reaction* j *is applicable at* $\mathbf{c}$*. In this case write* $\mathbf{c} \to^1_{\mathbf{u}} \mathbf{d}$.

Intuitively, by a single segment we mean running the reactions applicable at $\mathbf{c}$ at a constant (possibly 0) rate to get from $\mathbf{c}$ to $\mathbf{d}$. In the definition, $\mathbf{u}(j)$ represents the flux of reaction j.

Definition 2.2. *State* $\mathbf{d}$ *is* l-segment reachable*, written* $\mathbf{c} \rightsquigarrow^l \mathbf{d}$ *from* $\mathbf{c}$ *if* $(\exists \mathbf{b}_1, \ldots, \mathbf{b}_{l+1})$ $\mathbf{c} = \mathbf{b}_1 \to^1 \mathbf{b}_2 \to^1 \mathbf{b}_3 \to^1 \ldots \to^1 \mathbf{b}_{l+1} = \mathbf{d}$.

Definition 2.3. *State* $\mathbf{d}$ *is* segment-reachable *from* $\mathbf{c}$*, written* $\mathbf{c} \rightsquigarrow \mathbf{d}$*, if* $(\exists l \in \mathbb{N})$ $\mathbf{c} \rightsquigarrow^l \mathbf{d}$.

Suppose the reactions are $X \to C$ and $C + Y \to C + Z$, and we are in state $\{0C, 1X, 1Y, 0Z\}$. With straight-line segments, any state with a positive amount of Z must be reached in at least two segments: first to produce C, which allows the second reaction to occur, and then any combination of the first and second reactions. For example, $\{0C, 1X, 1Y, 0Z\} \to^1 \{0.1C, 0.9X, 1Y, 0Z\} \to^1 \{1C, 0X, 0Y, 1Z\}$. This is a simple example showing that more states are reachable with $\rightsquigarrow$ then $\to^1$. Often Definition 2.3 is used implicitly, when we make statements such as, "Run reaction 1 until its first reactant is 0, then run reaction 2", which implicitly defines two straight lines in concentration space.

[6]Note that $\mathbf{M}$ does not fully specify $\mathcal{C}$, since catalysts are not modeled: reactions $Z + X \to Z + Y$ and $X \to Y$ both correspond to the column vector $(-1, 1, 0)^\top$.

Given a curve, we can think about approximating it to an arbitrary accuracy using straight-line segments. But it may seem that we can never achieve the "full diversity" of states reachable with curves if we use only a finite number of line segments. However, the following lemma shows that increasing the number of straight-line segments beyond a certain point does not make any additional states reachable. Thus using a few line segments captures all the states reachable in the limit of infinitely many line segments.

Lemma 2.4. *If* $\mathbf{c} \rightsquigarrow \mathbf{d}$*, then* $\mathbf{c} \rightsquigarrow^{m+1} \mathbf{d}$*, where* $m = |R|$ *is the number of reactions.*

Proof. The intuitive idea of the proof is as follows. If $\mathbf{c} \rightsquigarrow \mathbf{d}$, then the path could use a huge number of segments, much greater than $m + 1$. We don't care precisely what these segments are, only that they give us an *ordering* $\prec$ *on species that need to become positive to reach* $\mathbf{d}$, such that, calling the species $S_1 \prec S_2 \prec \ldots \prec S_p$, for each i, there is a reaction to produce S_i that has only reactants in $[\mathbf{c}] \cup \{S_1, \ldots, S_{i-1}\}$. Therefore we can produce a small amount of each species $S_1, \ldots, S_n$, starting only from species in $[\mathbf{c}]$, by running these reactions in order, and running each reaction a smaller amount than the previous to ensure that nothing positive goes to 0. Once all these species are produced, a single straight line can take the CRN to state $\mathbf{d}$. The trick is that we should not produce *every* producible species, but only those needed to run the reactions corresponding to the final straight line. Otherwise, for example, we could create a species $X \notin [\mathbf{c}] \cup [\mathbf{d}]$, and if no reaction consumes X, $\mathbf{d}$ would become unreachable.

If $\mathbf{d}$ is reachable from $\mathbf{c}$ via fewer than $m+1$ straight line segments, then this can be re-expressed with precisely $m+1$ line segments by using length-0 line segments. It remains to show that at most $m+1$ straight line segments are necessary.

Suppose that $\mathbf{c} \rightsquigarrow \mathbf{d}$ via $l > m + 1$ straight line segments $\mathbf{c} = \mathbf{b}_0 \to^1 \mathbf{b}_1 \to^1 \ldots \to^1 \mathbf{b}_l = \mathbf{d}$. Let the reactions R be ordered as $\alpha_0, \alpha_1, \ldots$. For each $\mathbf{b}_i$, there exists $\mathbf{u}_i$ such that $\mathbf{b}_i + \mathbf{M}\mathbf{u}_i = \mathbf{b}_{i+1}$. Let $F_j = \sum_{i=0}^{l-1} \mathbf{u}_i(j)$ be the total flux through reaction α_j to get from $\mathbf{c}$ to $\mathbf{d}$ along the l straight line segments defined by the $\mathbf{b}_i$'s. Let $P = \{\ j \mid F_j > 0\ \}$ be the indices of reactions that occurred with strictly positive flux. If $\alpha_j = (\mathbf{r}_j, \mathbf{p}_j)$, let $\mathbf{v}_j = \mathbf{p}_j - \mathbf{r}_j$ be the vector representing the net change in concentrations if reaction α_j experiences unit flux. Then $\mathbf{d} = \mathbf{c} + \sum_{j \in P} F_j \mathbf{v}_j$. We use this fact to show that $\mathbf{c} \rightsquigarrow^{m+1} \mathbf{d}$.

Let $R(P) = \{\ \alpha_j \in R \mid j \in P\ \}$ be the reactions with indices in P. We first argue that there is an ordering $\prec$ of reactions in $R(P)$, such that for any reaction $\alpha = (\mathbf{r}, \mathbf{p})$, $[\mathbf{r}] \subseteq [\mathbf{c}] \cup \bigcup_{\alpha' = (\mathbf{r}', \mathbf{p}') \prec \alpha} [\mathbf{p}']$. In other words, if all products of reactions preceding α have positive concentration, as well as all species present in $\mathbf{c}$, this guarantees α is applicable. This in turn allows the reactions to proceed in the given order, so long as we are careful that no positive concentration species is allowed to be consumed back to 0. At the end, all products of all reactions in $R(P)$ are positive. To see that this ordering exists, it suffices to observe that the straight line segments given by the hypothesis $\mathbf{c} \rightsquigarrow \mathbf{d}$ give a *partial* order that respects the constraints, since any reactions with positive flux in the k'th straight line segment, which were not positive flux in any previous straight line segments, must consume only species that were either positive in $\mathbf{c}$, or that were produced by a reaction with positive flux in the l'th straight line segment, where $l < k$. Given

all reactions having positive flux for the first time in the k'th straight line segment, they can be ordered arbitrarily relative to each other while obeying the constraint $\prec$. For the sake of brevity, assume without loss of generality that $P = \{0, 1, \ldots, |P|-1\}$, and that $\alpha_0 \prec \alpha_1 \prec \ldots \prec \alpha_{|P|-1}$.

Let $K = \max_{S \in \Lambda} \max_{\alpha=(\mathbf{r},\mathbf{p}) \in R(P)} \mathbf{r}(S)$ be the maximum stoichiometry coefficient of any possible reactant. In the worst case, running a reaction by δ flux will decrease the concentration of any species by at most δK. Let

$$\epsilon = \min \left\{ \min_{S \in [\mathbf{c}]} \mathbf{c}(S), \min_{j \in P} F_j \right\}.$$

The significance of ϵ is that if we allow all reactions in $R(P)$ to go forward from $\mathbf{c}$ by strictly less than ϵ of flux, this will be less than the total flux experienced by each reaction, and also each species present in $\mathbf{c}$ will remain in positive concentration. Formally, define the following basis vectors (representing individual reaction vectors) $\mathbf{w}_j \in \mathbb{R}^m_{\geq 0}$ for $0 \leq j < |P|$. Let $\mathbf{w}_j(j) = \frac{\epsilon}{(2K)^{j+1}}$ and $\mathbf{w}_i(j) = 0$ if $i \neq j$. Then $\mathbf{M}\mathbf{w}_j$ is the j'th line segment to define. Let $\mathbf{a} = \mathbf{c} + \sum_{j=0}^{|P|-1} \mathbf{M}\mathbf{w}_j$ be the point reached after following these line segments. For each j, let $\mathbf{a}_j = \mathbf{c} + \sum_{i=0}^{j-1} \mathbf{M}\mathbf{w}_i$ be the state reached after following the first j straight line segments, letting $\mathbf{a}_0 = \mathbf{c}$.

We prove by induction on j that each $\mathbf{M}\mathbf{w}_j$ satisfies Definition 2.1 by showing that for all j, $S \in [\mathbf{a}_j] \implies \mathbf{a}_j(S) \geq \frac{\epsilon}{(2K)^{j+1}}$. I.e., all species are nonnegative at the start and end of the line (hence along the entire line), and the reactions experiencing positive flux along the line are applicable at the starting state.

We first establish the base case, where $\mathbf{w}_0(0) = \frac{\epsilon}{2K}$. By the definition of $\prec$, which ensures that for $\alpha_0 = (\mathbf{r}_0, \mathbf{p}_0)$, $[\mathbf{r}_0] \subseteq [\mathbf{c}]$, α_0 is applicable at $\mathbf{c}$. Note that $\epsilon \leq \mathbf{c}(S)$ for all $S \in [\mathbf{c}]$, i.e., running α_0 with flux $\frac{\epsilon}{2K}$ can cut the concentration of any species $S \in [\mathbf{c}]$ at most in half, so that in state $\mathbf{a}_1 = \mathbf{c} + \mathbf{M}\mathbf{w}_0$, $\mathbf{a}_1(S) \geq \frac{\epsilon}{2} \geq \frac{\epsilon}{2K}$, satisfying the base case for species in $[\mathbf{c}]$. Because α_0 experiences $\frac{\epsilon}{2K}$ of flux, each species produced $S \in [\mathbf{p}_0] \setminus [\mathbf{c}]$ initially absent in $\mathbf{c}$ but produced in $\mathbf{a}_1$ has increased in concentration by at least $\frac{\epsilon}{2K}$. Therefore all species in $[\mathbf{a}_1] \setminus [\mathbf{c}]$ have concentration at least $\frac{\epsilon}{2K}$, satisfying the base case for these species.

For the inductive case, inductively assume each species $S \in [\mathbf{a}_j]$ satisfies $\mathbf{a}_j(S) \geq \frac{\epsilon}{(2K)^{j+1}}$. Then by the definition of α_j and $\prec$, α_j is applicable in $\mathbf{a}_j$. Similarly to the base case, by running α_j by flux $\frac{\epsilon}{(2K)^{j+2}}$, the most any such species is consumed is $\frac{\epsilon}{(2K)^{j+2}}$, and since all such species S started with concentration $\mathbf{a}_j(S) \geq \frac{\epsilon}{(2K)^{j+1}}$, their concentrations are at most cut in half, i.e. $\mathbf{a}_{j+1}(S) \geq \frac{\epsilon}{(2K)^{j+2}}$. Therefore the inductive hypothesis for case $j+1$ is satisfied for species in $\mathbf{a}_j$. To show that the inductive hypothsis is satisfied for the rest of the species (i.e., those in $[\mathbf{a}_{j+1}] \setminus [\mathbf{a}_j]$), observe that all products $S \in [\mathbf{a}_{j+1}] \setminus [\mathbf{a}_j]$ have increased in concentration by at least $\frac{\epsilon}{(2K)^{j+2}}$.

The final line segment is a straight line from $\mathbf{a}$ to $\mathbf{d}$. By our choice of P, all reactions in $R(P)$ are applicable at $\mathbf{a}$. By our choice of $\mathbf{w}_j$'s, the total flux experienced by reaction α_j to get from $\mathbf{c}$ to $\mathbf{a}$ is strictly less than F_j, the total flux α_j experienced in the original straight line segments defining $\mathbf{c} \rightsquigarrow \mathbf{d}$. Therefore the final straight line from $\mathbf{a}$ can be defined by the vector $\mathbf{w}(j) = F_j - \mathbf{w}_j(j)$, so that $\mathbf{a} + \mathbf{M}\mathbf{w} = \mathbf{d}$. This shows that $\mathbf{c} \rightsquigarrow^{m+1} \mathbf{d}$. □

2.3 Relation to mass-action

For understanding the results of this paper, it is not necessary to be familiar with mass-action kinetics. However, as mass-action is the most commonly used model of chemical kinetics, we want to be satisfied that our notion of reachability at least considers states that can be reached by mass-action trajectories. Indeed, Lemma 2.8 shows that our notion of reachability is more general than mass-action.[7]

While a thorough discussion of mass-action kinetics is beyond the scope of this paper, we remind the reader that a CRN with rate constants assigned to each reaction defines a mass-action ODE system with a variable for each species. Each reaction contributes a term to the ODEs for all the species participating in it (except catalysts that are unchanged in the reaction). The term from reaction α appearing in the ODE for x is the product of: the rate constant, the reactant concentrations, and the net stoichiometry of x in α. For example, the CRN

$$X + X \to C$$
$$C + X \to C + Y$$

corresponds to ODEs:

$$\dot{x} = -2k_1 x^2 - k_2 cx$$
$$\dot{c} = k_1 x^2$$
$$\dot{y} = k_2 cx$$

where k_1, k_2 are the rate constants of the two reactions. We restrict our attention to CRNs whose mass-action trajectories are well-defined at all times $t \geq 0$, and where concentrations remain finite for all finite t.[8]

Definition 2.5. *A state* $\mathbf{d}$ *is* mass-action reachable *from* $\mathbf{c}$ *if there are non-zero rate constants such that the corresponding mass-action trajectory starting in* $\mathbf{c}$ *passes through* $\mathbf{d}$*, or approaches* $\mathbf{d}$ *as* $t \to \infty$.

In order to prove Lemma 2.8 we need to introduce the notion of a *siphon* from the Petri net literature. This notion will be used, as well, to prove negative results in later sections. Let $\mathcal{C} = (\Lambda, R)$ be a CRN. A *siphon* is a set of species $\Omega \subseteq \Lambda$ such that, for all reactions $\alpha = (\mathbf{r}, \mathbf{p}) \in R$, $[\mathbf{p}] \cap \Omega \neq \varnothing \implies [\mathbf{r}] \cap \Omega \neq \varnothing$, i.e., every reaction that produces an element of Ω requires a positive concentration of an element of Ω to be applicable. The following lemma, due to Angeli, De Leenheer, and Sontag [3], shows that this is equivalent to the notion that "the absence of Ω is forward-invariant" under mass-action: if all species in Ω are absent, then they can never again be produced (under mass-action).[9]

[7] It is much more general, in fact. For example, the reaction $X \to \varnothing$ can take X from a positive concentration to 0, whereas in mass-action kinetics the concentration of X would only approach 0 asymptotically as time $t \to \infty$.

[8] Although mass-action systems are free of many pathologies of more general dynamical systems, there are mass-action CRNs that reach infinite concentration in finite time: e.g. $2X \to 3X$

[9] It may appear obvious that if the rates of all reactions producing a particular species are zero, the species cannot be produced. However, this is a rather deep fact about mass-action ODEs. Consider the CRN $2X \to 3X$. The corresponding mass-action ODE is $\dot{x} = x^2$, and has the property that if you start with $x(0) = 0$, it cannot become positive.

LEMMA 2.6 ([3]). *Let $\Omega \subseteq \Lambda$ be a set of species. Then Ω is a siphon if and only if, for any state $\mathbf{c}$ such that $\Omega \cap [\mathbf{c}] = \varnothing$ and any state $\mathbf{d}$ that is mass-action reachable from $\mathbf{c}$, $\Omega \cap [\mathbf{d}] = \varnothing$.*

We show that the same holds true for segment-reachability.

LEMMA 2.7. *Let $\Omega \subseteq \Lambda$ be a set of species. Then Ω is a siphon if and only if, for any state $\mathbf{c}$ such that $\Omega \cap [\mathbf{c}] = \varnothing$ and any state $\mathbf{d}$ such that $\mathbf{c} \rightsquigarrow \mathbf{d}$, $\Omega \cap [\mathbf{d}] = \varnothing$.*

PROOF. Suppose Ω is a siphon, let $\mathbf{c}$ be a state such that $[\mathbf{c}] \cap \Omega = \varnothing$, and let $\mathbf{d}$ be such that $\mathbf{c} \rightsquigarrow \mathbf{d}$. Every straight line with end points $\mathbf{b}_i$ and $\mathbf{b}_{i+1}$ in the path from $\mathbf{c}$ to $\mathbf{d}$ is expressible as a sum of reaction vectors for reactions applicable in $\mathbf{b}_i$. Assuming inductively that $[\mathbf{b}_i] \cap \Omega = \varnothing$, the definition of siphon gives that $[\mathbf{b}_{i+1}] \cap \Omega = \varnothing$ as well. Therefore $\mathbf{d} \cap \Omega = \varnothing$.

To show the reverse direction, suppose that Ω is not a siphon. Then there is a reaction $\alpha = (\mathbf{r}, \mathbf{p})$ such that $[\mathbf{p}]$ contains an element $S \in \Omega$, but $[\mathbf{r}] \cap \Omega = \varnothing$. Then from any state $\mathbf{c}$ such that $[\mathbf{c}] = \Lambda \setminus \Omega$ (i.e., all species not in Ω are present), α is applicable. Running α produces S, hence results in a state $\mathbf{d}$ such that $\mathbf{c} \rightsquigarrow \mathbf{d}$ with $\Omega \cap [\mathbf{d}] \neq \varnothing$, since $S \in \Omega$. □

LEMMA 2.8. *If $\mathbf{d}$ is mass-action reachable from $\mathbf{c}$, then $\mathbf{c} \rightsquigarrow \mathbf{d}$.*

PROOF. This proof is similar to the proof of Lemma 2.4; the main difference is that we need to define the "reaction fluxes" each reaction experiences on the mass-action trajectory from $\mathbf{c}$ to $\mathbf{d}$, and that we need to find an ordering on those reactions with positive flux that allows us to apply them one-by-one in order to show $\mathbf{c} \rightsquigarrow^{m+1} \mathbf{d}$.

First, suppose $\mathbf{d}$ is reached in finite time t_f. The trajectory followed by mass-action kinetics to get from $\mathbf{c}$ to $\mathbf{d}$ defines a differentiable curve $\rho : [0, t_f] \to \mathbb{R}^n_{\geq 0}$. Let $\rho' = \frac{d\rho}{dt}$ be the corresponding mass-action differential equations. Let the reactions R be $\alpha_0, \alpha_1, \ldots$. Let $f_j : [0, t_f] \to \mathbb{R}_{\geq 0}$ be the "instantaneous flux" through the j'th reaction at time t. In other words, $\rho'(t) = \sum_j f_j(t)\mathbf{v}_j$, where index j ranges over all reactions, $\mathbf{v}_j$ represents the j'th reaction vector ($\mathbf{p} - \mathbf{r}$ if $\alpha_j = (\mathbf{r}, \mathbf{p})$, i.e., the j'th column of $\mathbf{M}$).

By the fundamental theorem of calculus,

$$\begin{aligned} \mathbf{d} - \mathbf{c} &= \rho(t_f) - \rho(0) = \int_0^{t_f} \rho'(t)dt \\ &= \int_0^{t_f} \sum_j f_j(t)\mathbf{v}_j dt = \sum_j \mathbf{v}_j \int_0^{t_f} f_j(t)dt, \end{aligned}$$

Let $F_j = \int_0^{t_f} f_j(t)dt$ be the total flux through the j'th reaction. Then $\mathbf{d} = \mathbf{c} + \sum_{j \in P} F_j \mathbf{v}_j$.

If $\mathbf{d}$ is not reached in finite time, then by the definition of mass-action reachability, $\mathbf{d} = \lim_{t_f \to \infty} \rho(t_f)$. Now, if we define $F_j = \int_0^{\infty} f_j(t)dt$ as above, the flux F_j could be infinite (consider $X \rightleftharpoons Y$, which approaches the point with equal concentrations of X and Y but has infinite flux through each reaction in the limit). Instead, we observe that for each species S being produced by one reaction whose flux is going to infinity, there must be a corresponding reaction consuming S whose flux is also going to infinity, for $\mathbf{d}(S)$ to be finite. Therefore we can find finite, positive values for each F_j such that $\mathbf{d} = \mathbf{c} + \sum_{j \in P} F_j \mathbf{v}_j$ as above.

Let $R_\rho = \{\ \alpha_j \in R \mid F_j > 0\ \}$ be the reactions that occurred with strictly positive flux along the curve ρ. We claim that there is an ordering $\prec$ of reactions in R_ρ, such that for any reaction $\alpha = (\mathbf{r}, \mathbf{p})$, $[\mathbf{r}] \subseteq [\mathbf{c}] \cup \bigcup_{\alpha' = (\mathbf{r}', \mathbf{p}') \prec \alpha} [\mathbf{p}']$.

Let $\mathcal{S}_0, \ldots, \mathcal{S}_l \subseteq \Lambda$ be a finite sequence of sets of species, and $\beta_1, \ldots, \beta_l \in R_\rho$ be a finite sequence of reactions which are generated as follows. Let $\mathcal{S}_0 = [\mathbf{c}]$. Given $\mathcal{S}_0, \ldots, \mathcal{S}_{i-1}$ and $\beta_1, \ldots, \beta_{i-1}$, let β_i be a new reaction in R_ρ whose reactants are in $\mathcal{S}_{i-1}$, and then form $\mathcal{S}_i$ by adding the products of β_i to $\mathcal{S}_{i-1}$. Let $\mathcal{S}$ be the final $\mathcal{S}_l$ such that we can no longer extend the sequence. Call the reactions appearing in the list of β's *processed*. If all of the reactions in R_ρ are processed, then the order in which they are processed can define the ordering $\prec$. We now claim that indeed all reactions in R_ρ will be processed.

We proceed by contradiction: we argue that the existence of an unprocessed reaction in R_ρ contradicts Lemma 2.6. Let $\overline{\mathcal{S}} = \Lambda \setminus \mathcal{S}$. For all unprocessed reactions in R_ρ, at least one reactant is in $\overline{\mathcal{S}}$. We also know that for all processed reactions in R_ρ, their products are not in $\overline{\mathcal{S}}$. Thus $\overline{\mathcal{S}}$ is a siphon for the CRN consisting of reactions in R_ρ. Let S^* be a reactant of the unprocessed reaction in R_ρ, and let F^* be its total flux. We know that $F^* > 0$, and thus there must be a point in the mass-action trajectory where S^* has positive concentration. However, $\mathbf{c}$ is zero on the siphon $\overline{\mathcal{S}}$, and $S^* \in \overline{\mathcal{S}}$. This violates Lemma 2.6.

The sequence $\beta_1, \ldots, \beta_l$ gives an ordering $\prec$ on reactions in R_ρ. Using this ordering, the remainder of the proof is identical to the proof of Lemma 2.4. □

Lemma 2.8 shows that our definition of reachability is at least as general as mass-action kinetics. In the other direction, ultimately we must appeal to intuition to justify that our definition does not reach too far. Our intuition says that the following properties are required of any reasonable notion of reachability: (1) concentrations must be nonnegative in all reachable states, (2) a reaction cannot execute in any state in which some reactant is 0,[10] (3) the relation should be reflexive, transitive, and "respect addition" (if $\mathbf{x} \to \mathbf{y}$, then for all $\mathbf{c}$, $\mathbf{x} + \mathbf{c} \to \mathbf{y} + \mathbf{c}$, since the presence of additional chemicals should not *prevent* reactions from being possible). Our definition was the most general one we could conceive that satisfied these properties.

Segment-reachability will serve as our main notion of reachability. Hence, throughout the rest of the paper, we say $\mathbf{d}$ is *reachable* from $\mathbf{c}$, written $\mathbf{c} \to \mathbf{d}$, if $\mathbf{c} \rightsquigarrow \mathbf{d}$.

2.4 Stable computation

In this section, we use our notion of reachability introduced above to formally define our notion of rate-independent computation. Intuitively, devious rate laws might take the system along trajectories anywhere in the reachable space. For the computation to be correct despite the rate laws, it must be able to reach a state with the final correct output from any reachable state. Further, in this setting an output

However, the very similar non-mass-action ODEs $\dot{x} = x^{1/2}$ has a perfectly valid solution $x(t) = t^2/4$ which starts at 0 but becomes positive, despite the fact that at $t = 0$, $\dot{x} = 0$.

[10] In particular, we want siphons to behave reasonably in that their absence is forward-invariant; i.e., that Lemmas 2.6 and 2.7 should apply to any reasonable notion of reachability.

can reasonably be called final only if no rate law can falsify it.

First, to formally define what it means for such a CRN to compute a function, we must first single out some aspects of the CRN as semantically meaningful. Formally, a *chemical reaction computer (CRC)* is a tuple $\mathcal{C} = (\Lambda, R, \Sigma, Y)$, where (Λ, R) is a CRN, $\Sigma \subset \Lambda$, written as $\Sigma = \{X_1, \ldots, X_k\}$,[11] is the *set of input species*, and $Y \in \Lambda \setminus \Sigma$ is the *output species*.

DEFINITION 2.9. *A state* $\mathbf{o} \in \mathbb{R}^{\Lambda}_{\geq 0}$ *is* output stable *if, for all* $\mathbf{o}'$ *such that* $\mathbf{o} \to \mathbf{o}'$, $\mathbf{o}(Y) = \mathbf{o}'(Y)$, *i.e., once* $\mathbf{o}$ *is reached, no reactions can change the concentration of the output species* Y.

DEFINITION 2.10. *Let* $f : \mathbb{R}^k_{\geq 0} \to \mathbb{R}_{\geq 0}$ *be a function*[12] *and let* $\mathcal{C}$ *be a CRC. We say that* $\mathcal{C}$ stably computes f *if, for all* $\mathbf{x} \in \mathbb{R}^k_{\geq 0}$, *for all* $\mathbf{c}$ *such that* $\mathbf{x} \to \mathbf{c}$, *there exists an output stable state* $\mathbf{o}$ *such that* $\mathbf{c} \to \mathbf{o}$ *and* $\mathbf{o}(Y) = f(\mathbf{x})$.

2.5 Dual-rail representations

In general, the output species of an upstream CRN may be used as the inputs to a downstream CRN if the upstream CRN only produces but never consumes the output species. We say that such a CRN computes its output *monotonically*. This is impossible for general stable CRNs. For example, any CRN computing the function $f(x_1, x_2) = x_1 - x_2$ must necessarily be able to consume its output species in order to account for some amount of species X_2 that has not yet reacted. Therefore, some of our CRNs represent their output Y in a "dual-rail" fashion as the difference of two species Y^+ and Y^-, both of which are only produced but never consumed by the CRN. Since these outputs are given as input to a downstream subroutine CRN, the downstream CRN must also be designed to accept inputs in this same dual-rail representation. Furthermore, representing values in this way allows us to take negative inputs and produce negative outputs, using only nonnegative concentrations, representing for example a negative output by a higher concentration of Y^- than Y^+.

Formally, let $f : \mathbb{R}^k \to \mathbb{R}$ be a function. A function $\hat{f} : \mathbb{R}^{2k} \to \mathbb{R}^2$ is a *dual-rail representation* of f if, for all $x^+, x^- \in \mathbb{R}^k$, if $(y^+, y^-) = \hat{f}(x^+, x^-)$, then $f(x^+ - x^-) = y^+ - y^-$. In other words, $\hat{f}$ represents f as the difference of its two outputs y^+ and y^-, and it works for any input pair (x^+, x^-) whose difference is the input value to f. We can define a CRC to stably compute such a function in the same manner as in Section 2.2, but having input species $X_1^+, X_1^-, X_2^+, X_2^-, \ldots, X_k^+, X_k^-, \ldots$, and two output species Y^+ and Y^-.

DEFINITION 2.11. *We say that a CRC* stably dual-computes $f : \mathbb{R}^k \to \mathbb{R}$ *if it stably computes a dual-rail representation* $\hat{f} : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R} \times \mathbb{R}$ *of* f.

Note that if a CRC monotonically outputs Y^+ and Y^- according to the dual-rail convention, and the output value is never negative, then it can be trivially modified to produce its output according to the direct computation convention (Def. 2.10) by adding the reaction $Y^+ + Y^- \to \varnothing$. Note that a single function has an infinite number of dual-rail representations. We require only that a CRN exists to compute one of them to say that the function is stably dual-computable by a CRN.

3. FUNCTIONS OVER REALS USING DUAL-RAIL REPRESENTATION

A function $f : \mathbb{R}^k \to \mathbb{R}$ is *rational linear* if there exist $a_1, \ldots, a_k \in \mathbb{Q}$ such that $f(\mathbf{x}) = \sum_{i=1}^k a_i \mathbf{x}(i)$. A function $f : \mathbb{R}^k \to \mathbb{R}$ is *piecewise rational linear* if there is a finite set of partial rational linear functions $f_1, \ldots, f_p : \mathbb{R}^k \dashrightarrow \mathbb{R}$, with $\bigcup_{j=1}^p \operatorname{dom} f_j = \mathbb{R}^k$, such that, for all $j \in \{1, \ldots, p\}$ and all $\mathbf{x} \in \operatorname{dom} f_j$, $f(\mathbf{x}) = f_j(\mathbf{x})$. In this case, we say that $f_1, \ldots, f_p$ are the *components* of f.

The following is the main theorem of this paper.

THEOREM 3.1. *A function* $f : \mathbb{R}^k \to \mathbb{R}$ *is stably dual-computed by a CRN if and only if it is continuous and piecewise rational linear.*

We prove each direction of Theorem 3.1 separately via Lemmas 3.3 and 3.13.

3.1 Continuous piecewise rational linear functions are computable

We require the following theorem, due to Ovchinnikov [15], which characterizes continuous piecewise linear functions in a way that is conducive to computation by CRNs.

THEOREM 3.2 ([15]). *Let* $D \subseteq \mathbb{R}^k$ *be convex. For every continuous piecewise linear function* $f : D \to \mathbb{R}$ *with components* $f_1, \ldots, f_p$, *there exists a family* $S_1, \ldots, S_q \subseteq \{1, \ldots, p\}$ *with* $S_i \not\subseteq S_j$ *if* $i \neq j$, *such that, for all* $\mathbf{x} \in D$, $f(\mathbf{x}) = \max_{i \in \{1,\ldots,q\}} \min_{j \in S_i} f_j(\mathbf{x})$.

The following lemma shows that any continuous piecewise rational linear function is stably computable by a CRN (in fact, with monotonic production of outputs in a dual-rail representation).

LEMMA 3.3. *Let* $D \subseteq \mathbb{R}^k$ *be convex, and let* $f : D \to \mathbb{R}$ *be a continuous piecewise rational linear function. Then* f *is monotonically stably dual-computed by a CRN.*

PROOF. By Theorem 3.2, it suffices to show how to compute a dual-rail representation of any rational linear function, a dual-rail representation of the minimum function with two inputs, and a dual-rail representation of the maximum function with two inputs. The latter two can be composed in a tree of depth $\log l$ to compute the minimum or maximum functions with input arity l.

[11] We assume a canonical ordering of $\Sigma = \{X_1, \ldots, X_k\}$ so that a vector $\mathbf{x} \in \mathbb{R}^k_{\geq 0}$ (i.e., an input to f) can be viewed equivalently as a state $\mathbf{x} \in \mathbb{R}^{\Sigma}_{\geq 0}$ of $\mathcal{C}$ (i.e., an input to $\mathcal{C}$). Also, we have defined valid initial states to contain only the input species Σ; other species must have initial concentration 0. Our results would change slightly if we relaxed this assumption; however, considering any species $S \notin \Sigma$ as an "auxiliary input", the result would remain that what is computed is a (piecewise) linear function of those species present initially. Since the initial value of non-input species would be constant across all valid initial states, rather than linear functions $f(\mathbf{x}) = \sum_{i=1}^k a_i \mathbf{x}(i)$, we would instead deal with *affine* functions $f(\mathbf{x}) = c + \sum_{i=1}^k a_i \mathbf{x}(i)$, where the constant c would depend only on the initial concentrations of non-input species.

[12] Our results extend easily to functions $f : \mathbb{R}^k \to \mathbb{R}^l$, i.e., whose output is a vector of l real numbers. This is because such a function is equivalently l separate functions $f_i : \mathbb{R}^k \to \mathbb{R}$.

Let $g : \mathbb{R}^k \to \mathbb{R}$ be a rational linear function $g(\mathbf{x}) = \sum_{i=1}^k a_i \mathbf{x}(i)$. By appropriate integer arithmetic, there exist $n_1, \ldots, n_k \in \mathbb{Z}$ and $d \in \mathbb{Z}^+$ such that $g(\mathbf{x}) = \frac{1}{d} \sum_{i=1}^k n_i \mathbf{x}(i)$. The following reactions compute a dual-rail representation of g with input species $X_1^+, \ldots, X_k^+, X_1^-, \ldots, X_k^-$ and output species Y^+, Y^-. For each i such that $n_i > 0$, add the reactions

$$\begin{aligned} X_i^+ &\to n_i W^+ \\ X_i^- &\to n_i W^- \end{aligned}$$

For each i such that $n_i < 0$, add the reactions

$$\begin{aligned} X_i^+ &\to |n_i| W^- \\ X_i^- &\to |n_i| W^+ \end{aligned}$$

To divide the values of W^- and W^+ by d, add the reactions

$$\begin{aligned} dW^+ &\to Y^+ \\ dW^- &\to Y^- \end{aligned}$$

To see that this works, for all states $\mathbf{c}$, define $x_i(\mathbf{c}) = \mathbf{c}(X_i^+) - \mathbf{c}(X_i^-)$, $w(\mathbf{c}) = \mathbf{c}(W^+) - \mathbf{c}(W^-)$, and $y(\mathbf{c}) = \mathbf{c}(Y^+) - \mathbf{c}(Y^-)$. Let $\mathbf{i}$ be the initial state. It is routine to check that the reactions enforce that for any state $\mathbf{c}$ reachable from $\mathbf{i}$,

$$y(\mathbf{c}) + \frac{1}{d} w(\mathbf{c}) + \frac{1}{d} \sum_{i=1}^k n_i x_i(\mathbf{c}) = \frac{1}{d} \sum_{i=1}^k n_i x_i(\mathbf{i}). \qquad (3.1)$$

(The right side is a constant depending only on the initial state, and each reaction increases one term on the left side by the same amount it decreases another term.)

If $\mathbf{c}$ is output stable, then $\mathbf{c}(X_i^+) = \mathbf{c}(X_i^-) = \mathbf{c}(W^+) = \mathbf{c}(W^-) = 0$, whence by (3.1) $y(\mathbf{c}) = \frac{1}{d} \sum_{i=1}^k n_i x_i(\mathbf{i})$, i.e, the output value of the CRN is the desired rational linear function of the inputs. Furthermore, observe that from any reachable state, it is always possible to reach an output stable state by executing the reactions above to completion in the order in which they are listed. This shows that a dual-rail representation of any rational linear function can be monotonically stably dual-computed by a CRN.

The following reactions monotonically stably compute a dual-rail representation of min with input species X_1^+, X_2^+, X_1^-, X_2^- and output species Y^+, Y^-. Add the reactions

$$\begin{aligned} X_1^+ + X_2^+ &\to Y^+ && (3.2) \\ X_1^- &\to X_2^+ + Y^- && (3.3) \\ X_2^- &\to X_1^+ + Y^- && (3.4) \end{aligned}$$

To see that this works, for all states $\mathbf{c}$, define $x_1(\mathbf{c}) = \mathbf{c}(X_1^+) - \mathbf{c}(X_1^-)$, $x_2(\mathbf{c}) = \mathbf{c}(X_2^+) - \mathbf{c}(X_2^-)$, and $y(\mathbf{c}) = \mathbf{c}(Y^+) - \mathbf{c}(Y^-)$. Let $\mathbf{i}$ be the initial state, and let $\mathbf{c}$ be a state reachable from $\mathbf{i}$. Let c_1, c_2, and c_3 be the amount of reactions (3.2), (3.3) and (3.4), respectively, that have executed to get from state $\mathbf{i}$ to state $\mathbf{c}$. Then we have $x_1(\mathbf{i}) - c_1 + c_2 + c_3 = x_1(\mathbf{c}), x_2(\mathbf{i}) - c_1 + c_2 + c_3 = x_2(\mathbf{c}), c_1 - c_2 - c_3 = y(\mathbf{c})$, which implies that $x_1(\mathbf{i}) = x_1(\mathbf{c}) + y(\mathbf{c}), x_2(\mathbf{i}) = x_2(\mathbf{c}) + y(\mathbf{c})$. In a stable state $\mathbf{c}$, either $\mathbf{c}(X_1^+) = 0$ or $\mathbf{c}(X_2^+) = 0$, otherwise reaction (3.2) is applicable. Further, in a stable state $\mathbf{c}$, $x_1(\mathbf{c}) = \mathbf{c}(X_1^+)$ and $x_2(\mathbf{c}) = \mathbf{c}(X_2^+)$, because $\mathbf{c}(X_1^-) = \mathbf{c}(X_2^-) = 0$ for reactions (3.3) and (3.4) not to be applicable. Thus, $x_1(\mathbf{c}) = 0$ or $x_2(\mathbf{c}) = 0$.

Thus, $x_1(\mathbf{i}) = y(\mathbf{c})$ and $x_2(\mathbf{i}) = x_2(\mathbf{c}) + y(\mathbf{c})$, or $x_2(\mathbf{i}) = y(\mathbf{c})$ and $x_1(\mathbf{i}) = x_1(\mathbf{c}) + y(\mathbf{c})$. So if $x_1(\mathbf{i}) < x_2(\mathbf{i})$, then it must be the first case (since $x_2(\mathbf{c}) > 0$). Otherwise, the second case holds. This shows that if stable state is reached, then computation is correct. Finally, observe that from any reachable state, by executing to completion the last applicable reaction among (3.3) and (3.4), followed by executing to completion (3.2), we obtain concentration zero of X_1^-, X_2^-, and one of X_1^+ or X_2^+, which implies that no reaction is applicable and the state is output stable.

To monotonically stably compute a dual-rail representation of max, observe that it is equivalent to computing the min function with the roles of the "plus" and "minus" species reversed (which negates the value represented in dual-rail), since $\max(x_1, x_2) = -\min(-x_1, -x_2)$. In other words, add the reactions

$$\begin{aligned} X_1^- + X_2^- &\to Y^- \\ X_1^+ &\to X_2^- + Y^+ \\ X_2^+ &\to X_1^- + Y^+ \end{aligned}$$

By appropriate renaming of input and output species of the three types of CRNs described above, they can be composed to compute $f(\mathbf{x}) = \max_{i \in \{1,\ldots,q\}} \min_{j \in S_i} f_j(\mathbf{x})$ as in Theorem 3.2. □

Although not essential to understanding the results of this paper, we now discuss the behavior of our construction with regards to mass-action kinetics. The following lemma shows that our construction actually converges to the correct output value under mass-action no matter how an adversary might have pushed us initially:

LEMMA 3.4. *For any input state $\mathbf{x}$, for any state $\mathbf{z}$ reachable from $\mathbf{x}$, the mass-action trajectory of the CRN of Lemma 3.3 with any non-zero rate constants starting at $\mathbf{z}$ converges to the output stable state in the limit $t \to \infty$.*

PROOF SKETCH. Our proof strategy is as follows. Using a Lyapunov function, we show that starting in any state $\mathbf{z}$ (not necessarily even reachable from $\mathbf{x}$), the CRN following mass-action kinetics converges to a steady-state as $t \to \infty$. Further, the CRN has the property that for every input state $\mathbf{x}$, there is exactly one mass-action steady-state that is segment-reachable from $\mathbf{x}$, and this state happens to be the output stable state. Since mass-action reachability (even in the limit) implies segment-reachability by Lemma 2.8, the convergence to steady-state implies convergence to the correct output.

We now show how to construct the appropriate Lyapunov function. The trajectory followed by mass-action kinetics starting from any state $\mathbf{z}$ defines a differentiable curve $\rho_{\mathbf{z}} : [0, \infty) \to \mathbb{R}^n_{\geq 0}$. Let $\rho'_{\mathbf{z}} = \frac{d\rho_{\mathbf{z}}}{dt}$ be the corresponding mass-action differential equations. Let $f_j^{\mathbf{z}} : [0, \infty) \to \mathbb{R}_{\geq 0}$ be the "instantaneous flux" through the j'th reaction at time t. (In other words, $\rho'_{\mathbf{z}}(t) = \sum_j f_j^{\mathbf{z}}(t) \mathbf{v}_j$, where index j ranges over all reactions, and $\mathbf{v}_j$ is the j'th column of stoichiometry matrix $\mathbf{M}$.) Our construction has the property that $F_j^{\mathbf{z}} = \int_0^\infty f_j^{\mathbf{z}}(t) dt$ (i.e. the total flux through the j'th reaction) is finite for any starting state $\mathbf{z}$. This follows because the CRN conserves mass up to multiplicative constants (due to stoichiometry) and our CRN is "feedforward" in the sense that there is an ordering on reactions according to which mass can flow, but it cannot flow back. Thus we can construct the Lyapunov function $\mathcal{L}(\mathbf{z}) : \mathbb{R}^n_{\geq 0} \to \mathbb{R}_{\geq 0}$ as $\mathcal{L}(\mathbf{z}) = \sum_j F_j^{\mathbf{z}}$.

To apply the standard Lyapunov machinery [13], we need to show three properties of the Lyapunov function: (1) $\mathcal{L} > 0$

everywhere except at a steady-state where $\mathcal{L} = 0$, (2) $\mathcal{L}$ gets arbitrarily large with increasing $\mathbf{z}$, (3) $\frac{d}{dt}\mathcal{L}(\rho_{\mathbf{z}}(t)) \leq 0$, and strictly negative everywhere except when $\rho_{\mathbf{z}}(t)$ is a steady state. This would guarantee that no matter where we start, the CRN converges to a steady-state under mass-action. The first two properties of the Lyapunov function are easily checked. To show the last property, we observe that $\frac{d}{dt}\mathcal{L}(\rho_{\mathbf{z}}(t)) = \sum_j \frac{d}{dt}\int_0^\infty f_j^{\rho_{\mathbf{z}}(t)}(t')dt' = \sum_j \frac{d}{dt}\int_t^\infty f_j^{\mathbf{z}}(t')dt' = \sum_j \frac{d}{dt}\left(\int_0^\infty f_j^{\mathbf{z}}(t')dt' - \int_0^t f_j^{\mathbf{z}}(t')dt'\right) = -\sum_j f_j^{\mathbf{z}}(t) \leq 0$ and equal to zero at exactly a steady state. □

3.2 Computable functions are continuous piecewise rational linear

Let $\mathcal{C} = (\Lambda, R, \Sigma, Y)$ be a CRC. We call any siphon Ω such that $(\mathbf{c} \upharpoonright \Omega = \mathbf{0}) \implies (\mathbf{c}$ is output stable) an *output stable siphon*. Lemma 3.5 shows the underlying relationship between output stability and siphons.

LEMMA 3.5. *Either every state is output stable, or every state is output unstable, or there is a set of output stable siphons $\mathcal{S}$ such that a state $\mathbf{c}$ is output stable if and only if $\exists\Omega \in \mathcal{S}$ such that $\mathbf{c} \upharpoonright \Omega = \mathbf{0}$.*

We prove some facts about siphons first that will help us characterize output stability in terms of them.

LEMMA 3.6. *Let $\mathbf{c}, \mathbf{d}_1, \ldots, \mathbf{d}_l$ be states such that $\mathbf{c} \to \mathbf{d}_1$, ..., $\mathbf{c} \to \mathbf{d}_l$. Then there exists $\mathbf{d}$ such that $\mathbf{c} \to \mathbf{d}$ and $[\mathbf{d}] = \bigcup_{i=1}^l [\mathbf{d}_i]$.*

PROOF. Since $\mathbf{c} \to \mathbf{d}_1, \ldots, \mathbf{c} \to \mathbf{d}_l$, we have that for all $i \in \{1, \ldots, l\}$, $\frac{1}{l}\mathbf{c} \to \frac{1}{l}\mathbf{d}_i$. Hence $\mathbf{c} \to \frac{1}{l}\sum_{i=1}^l \mathbf{d}_i$. Letting $\mathbf{d} = \frac{1}{l}\sum_{i=1}^l \mathbf{d}_i$ completes the proof. □

The next lemma shows that if a set of species has even a single state from which none of the species can be produced, then it is a siphon.

LEMMA 3.7. *If $\mathbf{c}$ is a state and Ω' is a set of species such that for all $\mathbf{d}$ reachable from $\mathbf{c}$, $\mathbf{d} \upharpoonright \Omega' = \mathbf{0}$, then there exists a siphon $\Omega \supseteq \Omega'$ such that $\mathbf{c} \upharpoonright \Omega = \mathbf{0}$.*

PROOF. There is a unique largest set of species $\overline{\Omega}$ such that $\forall S \in \overline{\Omega}$, $\exists\mathbf{d}$ reachable from $\mathbf{c}$ and $\mathbf{d}(S) > 0$, i.e., $\overline{\Omega}$ is the set of species producible from $\mathbf{c}$. Let $\mathbf{d}$ be a state reachable from $\mathbf{c}$ such that, for all $S \in \overline{\Omega}$, $\mathbf{d}(S) > 0$; such a state exists by Lemma 3.6. Let $\Omega = \Lambda \setminus \overline{\Omega}$; we must show Ω is a siphon.

Suppose for the sake of contradiction that Ω is not a siphon. Then there exists some state $\mathbf{c}'$ with $\mathbf{c}' \upharpoonright \Omega = 0$ (implying $[\mathbf{c}'] \subseteq [\mathbf{d}]$), some state $\mathbf{d}'$ reachable from $\mathbf{c}'$, and $S \in \Omega$, such that $\mathbf{d}'(S) > 0$. Let $\epsilon > 0$ be sufficiently small that $\epsilon \cdot \mathbf{c}' \leq \mathbf{d}$; such an ϵ exists because $[\mathbf{c}'] \subseteq [\mathbf{d}]$. Then $\epsilon \cdot \mathbf{d}'$ is reachable from $\epsilon \cdot \mathbf{c}'$, and $\epsilon \cdot \mathbf{d}'(S) > 0$. Since $\epsilon \cdot \mathbf{c}' \leq \mathbf{d}$ and reachability respects addition, this implies that S is producible from $\mathbf{d}$ (hence from $\mathbf{c}$) as well, implying $S \in \overline{\Omega}$, a contradiction since we chose $S \in \Omega$ and $\Omega \cap \overline{\Omega} = \varnothing$. □

We are now ready to prove Lemma 3.5.

PROOF. (of Lemma 3.5) We create the set of siphons as follows. Let $\{r_j = (\mathbf{r}_j, \mathbf{p}_j)\}_j$ be the set of all reactions that change Y (consume or produce; i.e., $\mathbf{r}_j(Y) \neq \mathbf{p}_j(Y)$). We construct sets of species $\{\Omega'_k\}_k$ by taking one reactant from each reaction in every possible way (i.e., the set of subsets of species $\{\ \Omega'_k \subseteq \Lambda \mid (\forall j)\ \Omega'_k \cap \mathbf{r}_j \neq \varnothing\ \}$). For each possible siphon Ω such that $\Omega \supseteq \Omega'_k$, add Ω to $\mathcal{S}$. It is easy to see that if $\mathbf{c}$ is zero on some siphon $\Omega \in \mathcal{S}$ then it is output stable.

For the other direction: Let $O(\mathbf{c}) = \{\ \Omega'_k \mid \mathbf{c} \upharpoonright \Omega'_k = \mathbf{0}\ \}$. If in some state $\mathbf{c}$, we have $O(\mathbf{c}) = \varnothing$, then there is a reaction r_j that changes Y and is applicable in $\mathbf{c}$. Then Y can change by this reaction and $\mathbf{c}$ cannot be output stable. So for output stable $\mathbf{c}$, $O(\mathbf{c})$ is non-empty. Now let $\mathbf{d}_1, \ldots, \mathbf{d}_l$ be some states reachable from $\mathbf{c}$ such that if $\mathbf{c} \to \mathbf{d}$ then $\exists i, O(\mathbf{d}_i) = O(\mathbf{d})$ (i.e. these states cover the whole variety of $O(\mathbf{d}_i)$'s). Lemma 3.6 implies that if $\cap_{i=1}^l O(\mathbf{d}_i) = \varnothing$ then $\mathbf{c} \to \mathbf{d}$ such that $O(\mathbf{d}) = \varnothing$. Thus, for an output stable $\mathbf{c}$, at least one of the original sets $\Omega'_k \in O(\mathbf{c})$ stays zero on all states reachable from $\mathbf{c}$. This implies that there is some siphon $\Omega \in \mathcal{S}$ that includes Ω'_k and $\mathbf{c} \upharpoonright \Omega = \mathbf{0}$ (by Lemma 3.7). □

Recall that we write the set of reactions as $R = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$.

DEFINITION 3.8. *Let $\mathbf{s} = (R_1, \ldots, R_l) \in \mathcal{P}(R)^l$ denote a sequence of subsets of reactions. Say that $\mathbf{c} \to_{\mathbf{s}} \mathbf{d}$ if $\mathbf{c} = \mathbf{b}_0 \to^1_{\mathbf{u}_1} \mathbf{b}_1 \to^1_{\mathbf{u}_2} \mathbf{b}_2 \to^1_{\mathbf{u}_3} \ldots \to^1_{\mathbf{u}_l} \mathbf{b}_l = \mathbf{d}$, where, for each $i \in \{1, \ldots, l\}$ and $j \in \{1, \ldots, m\}$, $\mathbf{u}_i(j) > 0 \iff \alpha_j \in R_i$.*

In other words, $\mathbf{d}$ is reachable from $\mathbf{c}$ via l straight lines, in which the i'th straight line is defined by a strictly positive weighted sum of reactions in R_i.

Let $X_{\mathbf{s}}(\Omega) = \{\mathbf{x} \in \mathbb{R}^\Lambda_{\geq 0} \mid (\exists\mathbf{o})\ \mathbf{x} \to_{\mathbf{s}} \mathbf{o}$ and $\mathbf{o} \upharpoonright \Omega = \mathbf{0}\}$ denote those states from which Ω is drainable (thus the reached state is output stable since Ω is an output siphon) via straight lines respecting $\mathbf{s}$.

Note that the following lemma concerns *direct* stable computability.

LEMMA 3.9. *Let $f : \mathbb{R}^k_{\geq 0} \to \mathbb{R}_{\geq 0}$ be stably computed by a CRC $\mathcal{C} = (\Lambda, R, \Sigma, Y)$. Let Ω be an output stable siphon and let $\mathbf{s} = (R_1, \ldots, R_l) \in \mathcal{P}(R)^l$ for some $l \in \mathbb{N}$. Then f restricted to inputs in $X_{\mathbf{s}}(\Omega)$ is rational linear.*

PROOF. By the definition of output stable siphon, every state $\mathbf{o}$ such that $\mathbf{o} \upharpoonright \Omega = \mathbf{0}$ is output-stable. By the stability of $\mathcal{C}$, every such state $\mathbf{o}$ reachable from a particular initial state $\mathbf{x}$ must have the same value of $y = \mathbf{o}(Y) = f(\mathbf{x})$.

By the definition of $X_{\mathbf{s}}(\Omega)$, an output stable state $\mathbf{o}$ is reachable from $\mathbf{x}$, with $\mathbf{o} \upharpoonright \Omega = \mathbf{0}$, via straight-line segments respecting $\mathbf{s}$. Then by Definition 3.8, there are states $\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_l \in \mathbb{R}^n_{\geq 0}$ such that $\mathbf{x} = \mathbf{b}_0 \to^1_{\mathbf{u}_1} \mathbf{b}_1 \to^1_{\mathbf{u}_2} \mathbf{b}_2 \to^1_{\mathbf{u}_3} \ldots \to^1_{\mathbf{u}_l} \mathbf{b}_l = \mathbf{o}$, where, for each $i \in \{1, \ldots, l\}$ and $j \in R_i$, $\mathbf{u}_i(j) > 0 \iff \alpha_j \in R_i$.

We claim that the condition "$\mathbf{x} \in X_{\mathbf{s}}(\Omega)$" is equivalent to a conjunction of linear inequalities (some strict and some non-strict). The fact that each $\mathbf{b}_i \to^1_{\mathbf{u}_{i+1}} \mathbf{b}_{i+1}$ is captured by the inequalities $\mathbf{b}_i(S) > 0$ for each species S that appears as a reactant of some reaction in R_i, and to the linear equations $\mathbf{b}_i = \mathbf{b}_{i-1} + \mathbf{M}\mathbf{u}_i$ (where $\mathbf{M}$ is the reaction stoichiometry matrix from Definition 2.1). The fact that all concentrations must be nonnegative corresponds to the linear inequalities $\mathbf{b}_i \geq \mathbf{0}$. The fact that all reactions must happen in the forward direction (recall we represent reversible reactions such as $A \rightleftharpoons B$ as two irreversible reactions $A \to B$ and $B \to A$) corresponds to the linear inequalities $\mathbf{u}_i \geq \mathbf{0}$. The fact that the siphon Ω is drained corresponds to the linear inequality $\mathbf{o}(S) \leq 0$ for each $S \in \Omega$, together with the existing constraint $\mathbf{o} \geq \mathbf{0}$.

These equations and inequalities can be expressed as a single system of linear inequalities $\mathbf{Az} \quad \geq / > \quad \mathbf{0}$, (i.e., some strict and some non-strict inequalities) where $\mathbf{A}$ is a constant rational matrix depending only on the constants $\mathbf{M}$ and Ω, and $\mathbf{z}$ is a single vector of unknowns corresponding to all the unknowns above, namely $\mathbf{x}$ (the initial species concentrations subject to the constraint that all non-input species start at 0), $\mathbf{o}$ (the final species concentrations subject to the constraint that all siphon species end up at 0), and the various $\mathbf{u}_i$'s (the fluxes through each reaction to get from $\mathbf{b}_i$ to $\mathbf{b}_{i+1}$), and the $\mathbf{b}_i$'s.

These inequalities define a convex subset S of $\mathbb{R}^d$ for some constant d. Let $G \subseteq \mathbb{R}^{k+1}$ be the projection of S to the $(k+1)$-dimensional subspace corresponding to $\mathbf{x}$ and $y = \mathbf{o}(Y)$. G is the graph of the function $y = f(\mathbf{x})$ restricted to inputs $\mathbf{x} \in X_{\mathbf{s}}(\Omega)$. Since S is convex, G is also convex. We claim that G must be a subset of a k-dimensional rational hyperplane. For the sake of contradiction, suppose not. The matrix defining the set S was rational, so G cannot be a subset of a k-dimensional *irrational* hyperplane. Then there are $k+1$ non-coplanar points in G. Since G is convex, it contains the entire $(k+1)$-dimensional convex hull H of these points. Since H is a $(k+1)$-dimensional convex polytope, it contains two different values of y corresponding to the same value of $\mathbf{x}$, contradicting the fact that only a single y value exists in all output-stable states reachable from $\mathbf{x}$. This establishes the claim that G must be a subset of a k-dimensional rational hyperplane.

Since the graph of f is a subset of a k-dimensional rational hyperplane, it is a rational affine function. Since there are no reactions of the form $\varnothing \to \ldots$, Y cannot be produced from the initial state $\mathbf{x} = \mathbf{0}$ (nor can any other species), so $f(\mathbf{0}) = 0$. Therefore this hyperplane passes through the origin, so it defines a rational linear function. □

In Lemma 3.9, the reason that we restrict attention to a single output siphon Ω *and* a single sequence $\mathbf{s}$ of subsets of reactions is as follows. If different output siphons are drained, then different linear functions may be computed by the CRC. For example, $X_1 + X_2 \to Y$ computes $f(x_1, x_2) = x_1$ if siphon $\{X_2\}$ is drained and $f(x_1, x_2) = x_2$ if siphon $\{X_1\}$ is drained. Further our proof technique relies on fixing $\mathbf{s}$ to obtain convexity. If we do not force a particular set of reactions to occur with positive flux on each line segment, then to logically express the condition that Ω is drainable from state $\mathbf{x}$ would require that we specify, for each each line segment and each species S, that either S has positive concentration at the start, *or* S is not a reactant used in that line segment. This would imply that the condition of Ω being drainable from $\mathbf{x}$ is expressible as a conjunction of disjunctions of linear inequalities, which is not necessarily a convex set as required in the proof.

LEMMA 3.10. *Let $f : \mathbb{R}^k \to \mathbb{R}$ be stably dual computed by a CRC. Let Ω be an output stable siphon and let $\mathbf{s} = (R_1, \ldots, R_l) \in \mathcal{P}(R)^l$ for some $l \in \mathbb{N}$. Then f restricted to inputs that have a dual rail representation in $X_{\mathbf{s}}(\Omega)$ is rational linear.*

PROOF. A dual-rail computing CRC can be thought to *directly* compute two separate functions $\hat{f}^+, \hat{f}^- : \mathbb{R}^{2k}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\hat{f} = \hat{f}^+ - \hat{f}^-$ where $\hat{f}$ is a dual rail representation of f. By Lemma 3.9 we know that $\hat{f}^+$ and $\hat{f}^-$ are rational linear when restricted to $X_{\mathbf{s}}(\Omega)$. The lemma then follows because rational linearity is closed under subtraction. □

LEMMA 3.11. *Let Ω be an output siphon and $R' \subseteq R$ be a subset of reactions. Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \in X_{(R')}(\Omega)$ be a convergent sequence of states, where $\mathbf{a} = \lim_{i\to\infty} \mathbf{a}_i$, i.e., Ω is drainable from every $\mathbf{a}_i$, via a single straight line segment that uses precisely reactions from R'. Suppose $[\mathbf{a}] = \Lambda$. Then $\mathbf{a} \in X_{(R')}(\Omega)$.*

PROOF. As in the proof of Lemma 3.9, the condition $\mathbf{a}_i \in X_{(R')}(\Omega)$ is equivalent to the condition that $\mathbf{a}_i$ satisfies a set of linear inequalities. Some inequalities may be strict; out goal is to show we can remove the strict inequalities so that the remainder define a closed set.

Unlike Lemma 3.9, because $[\mathbf{a}] = \Lambda$, every reaction in R (including those in R') is applicable to $\mathbf{a}$. Thus all states reachable from $\mathbf{a}$ are reachable via a single straight line. For all $\epsilon > 0$, let $\mathcal{B}_\epsilon(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n_{\geq 0} \mid (\forall S \in \Lambda)\ |\mathbf{x}(S) - \mathbf{a}(S)| \leq \epsilon\}$ be the ball of (L_∞) radius ϵ around $\mathbf{a}$. For any $0 < \epsilon < \min_{S\in\Lambda} \mathbf{a}(S)$, $\mathbf{x} \in \mathcal{B}_\epsilon(\mathbf{a}) \implies [\mathbf{x}] = \Lambda$, since being within ϵ of $\mathbf{a}$ implies that all species are present. Furthermore, note that this condition is expressible using non-strict inequalities.

Thus the condition that $\mathbf{x} \in X_{(R')}(\Omega) \cap \mathcal{B}_\epsilon(\mathbf{a})$ is expressible as a conjunction of linear inequalities, *without* using the strict inequality $\mathbf{b}_i(k) > 0$ enforcing that reactants of reactions in R' are present. Hence we are left with only nonstrict inequalities, which define a closed set. Thus $X_{(R')}(\Omega) \cap \mathcal{B}_\epsilon(\mathbf{a})$ is a closed set. Since $\mathbf{a} = \lim_{i\to\infty} \mathbf{a}_i$, for all but finitely many i, $\mathbf{a}_i \in X_{(R')}(\Omega) \cap \mathcal{B}_\epsilon(\mathbf{a})$. Since this set is closed, it contains all its limit points, whence $\mathbf{a} \in X_{(R')}(\Omega)$. □

Note that the hypothesis $[\mathbf{a}] = \Lambda$ is necessary. Otherwise, consider the reactions $A \to C$, $A + B \to \varnothing$, and $F + C \to C$, with $\mathbf{a}_i(C) = 0$, $\mathbf{a}_i(F) = 1$, $\mathbf{a}_i(B) = 1$, and $\mathbf{a}_i(A)$ approaching 1 from above as $i \to \infty$ (whence $C \notin [\mathbf{a}]$). Then the siphon $\Omega = \{A, B, F\}$ is drainable from each $\mathbf{a}_i$ by running $A \to C$ until A and B have the same concentration, then running the other two reactions to completion. However, $\mathbf{a}(A) = \mathbf{a}(B)$, so running any amount of reaction $A \to C$ prevents reaction $A + B \to \varnothing$ from draining B. Therefore $\mathbf{a} \notin X_{(R')}(\Omega)$ but $\mathbf{a}_i \in X_{(R')}(\Omega)$ for all i.

The following lemma concerns *direct* computability and will be useful for both direct and dual computation by CRCs. A function $f : \mathbb{R}^k_{\geq 0} \to \mathbb{R}_{\geq 0}$ is *positive-continuous* if, for all $U \subseteq \{1, \ldots, k\}$, f is continuous on the domain $D_U = \left\{ \mathbf{x} \in \mathbb{R}^k_{\geq 0} \;\middle|\; (\forall i \in \{1, \ldots, k\})\ \mathbf{x}(i) > 0 \iff i \in U \right\}$. In other words, f is continuous on any subset $D \subset \mathbb{R}^k_{\geq 0}$ that does not have any coordinate $i \in \{1, \ldots, k\}$ that takes both zero and positive values in D.

LEMMA 3.12. *Let $f : \mathbb{R}^k_{\geq 0} \to \mathbb{R}_{\geq 0}$ be stably computed by a CRC. Then f is positive-continuous.*

PROOF. Let $U \subseteq \{1, \ldots, k\}$, let $\mathbf{x} \in D_U$, and let $\mathbf{x}_1$, $\mathbf{x}_2$, $\ldots \in D_U$ be an infinite sequence of points such that $\lim_{i\to\infty} \mathbf{x}_i = \mathbf{x}$. It suffices to show that $\lim_{i\to\infty} f(\mathbf{x}_i) = f(\mathbf{x})$ — i.e. that f is continuous on D_U. We take $\mathbf{x}_i$ and $\mathbf{x}$ equivalently to be the initial state of the CRC giving the concentrations of species in $\Sigma = \{X_1, \ldots, X_k\}$.

In analyzing the behavior of the CRC on states in D_U, it will help us to consider the functionally equivalent CRC in which we remove species that are not producible from states in D_U. For the purposes of this proof we consider this reduced CRC, and let Λ be the corresponding reduced set of species.

Let $\delta = \min_{i \in U} \mathbf{x}(i)$ be the smallest concentration of an initially positive species. Let $\mathbf{d} \in \mathbb{R}^k_{\geq 0}$ be the state $\mathbf{d}(X_i) = \delta$ for all $i \in U$ and $\mathbf{d}(S) = 0$ for all other species S. By an argument similar to the proof of Lemma 2.4, we can reach from $\mathbf{d}$ by at most m straight line segments to a state $\mathbf{e}$ such that $[\mathbf{e}] = \Lambda$. Because of the way we defined $\mathbf{d}$, $\mathbf{x}_i \geq \mathbf{d}$ and $\mathbf{x} \geq \mathbf{d}$. Thus, letting $\mathbf{a}_i = \mathbf{x}_i + \mathbf{e} - \mathbf{d}$, we have that $\mathbf{x}_i \to_{\mathbf{s}} \mathbf{a}_i$ for some fixed $\mathbf{s}$, and $[\mathbf{a}_i] = \Lambda$. Further, $\mathbf{a} = \lim_{i \to \infty} \mathbf{a}_i = (\lim_{i \to \infty} \mathbf{x}_i) + \mathbf{e} - \mathbf{d} = \mathbf{x} + \mathbf{e} - \mathbf{d}$. Thus, $\mathbf{x} \to_{\mathbf{s}} \mathbf{a}$ and $[\mathbf{a}] = \Lambda$.

Since there are a finite number of output siphons, by the pigeonhole principle, one such output siphon Ω must be drainable from infinitely many such states $\mathbf{a}_i$ (via a straight line as observed previously since $[\mathbf{a}_i] = \Lambda$). Let $\mathbf{o}_i$ be an output-stable state such that $\mathbf{a}_i \to^1 \mathbf{o}_i$.

There is a finite number of subsets of reactions that may occur in the final straight line segment from $\mathbf{a}_i$ to $\mathbf{o}_i$ with strictly positive flux, so one (call it R') must repeat infinitely often. We used an identical prefix of straight line segments $\mathbf{s}$ for all i to reach from $\mathbf{x}_i$ to $\mathbf{a}_i$. Letting $\mathbf{s}'$ be $\mathbf{s}$ concatenated with R', we have that $\mathbf{s}'$ is infinitely often the sequence of subsets of reactions with strictly positive flux, i.e., for infinitely many i, $\mathbf{x}_i \to_{\mathbf{s}} \mathbf{a}_i \to^1_{(R')} \mathbf{o}_i$ and the same siphon Ω is drained in all these $\mathbf{o}_i$.

Let this infinite subsequence of $\mathbf{a}_i$ states agreeing on Ω and R' (and thus $\mathbf{s}'$) be denoted $\mathbf{a}'_1, \mathbf{a}'_2, \ldots \in \mathbb{R}^n_{\geq 0}$. Let $\mathbf{x}'_1, \mathbf{x}'_2, \ldots \in \mathbb{R}^k$ be the initial states corresponding to the subsequence $(\mathbf{a}'_i)$ (i.e., $\mathbf{x}'_i \to_{\mathbf{s}} \mathbf{a}'_i$ as described before).

Clearly, $\mathbf{x}'_i \in X_{\mathbf{s}'}(\Omega)$. Further, observe that the limit $\lim_{i \to \infty} \mathbf{a}'_i = \lim_{i \to \infty} \mathbf{a}_i = \mathbf{a}$, and recall $[\mathbf{a}] = \Lambda$. Since $\mathbf{a}'_i \in X_{(R')}(\Omega)$ and $[\mathbf{a}] = \Lambda$, this implies by Lemma 3.11, $\mathbf{a} \in X_{(R')}(\Omega)$. Since $(\mathbf{x}'_i)$ is a infinite subsequence of a sequence converging to $\mathbf{x}$, $\mathbf{x} = \lim_{i \to \infty} \mathbf{x}'_i$. Since $\mathbf{x} \to_{\mathbf{s}} \mathbf{a}$ and $\mathbf{a} \in X_{(R')}(\Omega)$, it follows that $\mathbf{x} \in X_{\mathbf{s}'}(\Omega)$.

By Lemma 3.9, f is rational linear – hence continuous – on inputs in $X_{\mathbf{s}'}(\Omega)$. Because $\mathbf{x}, \mathbf{x}'_1, \mathbf{x}'_2, \ldots \in X_{\mathbf{s}'}(\Omega)$, the continuity on inputs in $X_{\mathbf{s}'}(\Omega)$ implies that $f(\mathbf{x}) = \lim_{i \to \infty} f(\mathbf{x}'_i)$, which equals $\lim_{i \to \infty} f(\mathbf{x}_i)$ since $\mathbf{x}'_i$ is a subsequence of $\mathbf{x}_i$. Since $\mathbf{x}$ was arbitrary and $\mathbf{x}_1, \mathbf{x}_2, \ldots$ was an arbitrary sequence converging to $\mathbf{x}$, it follows that f is continuous. □

LEMMA 3.13. *Let $f : \mathbb{R}^k \to \mathbb{R}$ be stably dual-computable by a CRC. Then f is continuous and piecewise rational linear.*

PROOF. Let $\mathcal{C}$ be the CRC stably computing a dual-rail representation $\hat{f}$ of f, with input species $X_1^+, \ldots, X_k^+, X_1^-, \ldots, X_k^-$ and output species Y^+, Y^-.

Similarly to the proof of Lemma 3.10, a dual-rail computing CRC can be thought to *directly* compute two separate functions $\hat{f}^+, \hat{f}^- : \mathbb{R}^{2k}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\hat{f} = \hat{f}^+ - \hat{f}^-$ where $\hat{f}$ is a dual rail representation of f. Let $\delta > 0$. Then for *any* input $\mathbf{x}' \in \mathbb{R}^k$ to f, there is an initial state $\mathbf{x} \in \mathbb{R}^\Sigma_{\geq \delta}$ representing $\mathbf{x}'$. Any sequences inputs $\mathbf{x}'_1, \mathbf{x}'_2, \ldots \in \mathbb{R}^k$ to f such that $\lim_{i \to \infty} \mathbf{x}'_i = \mathbf{x}'$ are similarly represented by a sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots \in \mathbb{R}^\Sigma_{\geq 0}$ of initial states of the CRC such that $\lim_{i \to \infty} \mathbf{x}_i = \mathbf{x}$ has the property that all but finitely many $\mathbf{x}_i$ obey $[\mathbf{x}_i] = \Sigma$. By Lemma 3.12, f is continuous on the domain in which all input species are positive, which includes the input represented by $\mathbf{x}$ and the inputs represented by all but finitely many of $\mathbf{x}_i$. Therefore, $f(\mathbf{x}') = \lim_{i \to \infty} f(\mathbf{x}'_i)$, so f is continuous.

It remains to show that f is piecewise rational linear. If $\mathbf{x}$ is an initial state and $\mathbf{x} \to \mathbf{o}$ where $\mathbf{o}$ is output stable, then by Lemma 2.4, $\mathbf{x} \rightsquigarrow^{m+1} \mathbf{o}$. Let Ω be an output siphon such that $\mathbf{o} \upharpoonright \Omega = \mathbf{0}$, which exists by Lemma 3.5. Let $\mathbf{s} = (R_1, \ldots, R_l) \in \mathcal{P}(R)^l$, where $\mathbf{x} \to_{\mathbf{s}} \mathbf{o}$, where by Lemma 2.4 we may assume $l \leq m + 1$. By Lemma 3.10, this value of y is a rational linear function of $\mathbf{x}$. Since there are a finite number of output siphons and sequences in $\mathcal{P}(R)^{\leq m+1}$, this establishes the piecewise rational linearity of f. □

4. FUNCTIONS OVER NONNEGATIVE REALS USING DIRECT REPRESENTATION

In this section, we show an analogous result to Theorem 3.1 for CRNs that compute functions *directly*, instead of computing a dual-rail representation. As a result, since concentrations are nonnegative, we may only compute functions $f : \mathbb{R}^k_{\geq 0} \to \mathbb{R}_{\geq 0}$. Interestingly, the class of computable functions expands slightly to include some discontinuous functions, specifically the *positive-continuous* functions defined in Section 3.

THEOREM 4.1. *A function $f : \mathbb{R}^k_{\geq 0} \to \mathbb{R}_{\geq 0}$ is stably computable by a CRN if and only if f is positive-continuous and piecewise rational linear.*

We prove each direction of Theorem 4.1 separately via Lemmas 4.2 and 4.4.

4.1 Positive-continuous piecewise rational linear functions are computable

LEMMA 4.2. *Every positive-continuous piecewise linear function $f : \mathbb{R}^k_{\geq 0} \to \mathbb{R}_{\geq 0}$ is stably computable by a CRN.*

PROOF. The CRN will have input species $X_1, \ldots, X_k$ and output species Y^+. (While it will be helpful to think of a Y^+ and Y^- species, and during the computation the output will be encoded in their difference, the output of the CRC is only the Y^+ species as per direct computability.)

Let $f : \mathbb{R}^k_{\geq 0} \to \mathbb{R}_{\geq 0}$ be a positive-continuous piecewise linear function. Since it is positive-continuous, there exist 2^k domains

$$D_U = \{\mathbf{x} \in \mathbb{R}^k_{\geq 0} \mid (\forall i \in \{1, \ldots, k\})\ \mathbf{x}(i) > 0 \iff i \in U\},$$

one for each subset $U \subseteq \{1, \ldots, k\}$, such that $f \upharpoonright D_U$ is continuous on D_U. Define $f_U = f \upharpoonright D_U$.

Since each D_U is convex, by Lemma 3.3, there is a CRN $\mathcal{C}_U$ monotonically computing a dual-rail representation $\hat{f}_U : \mathbb{R}^k_{\geq 0} \times \mathbb{R}^k_{\geq 0} \to \mathbb{R} \times \mathbb{R}$ of f_U. By letting the initial concentration of the "minus" version of the i'th input species X_i^- be 0, we convert $\mathcal{C}_U$ into a CRN that directly computes an output dual-rail representation of f_U.

The intuition of the proof is as follows. The case $U = \varnothing$ is trivial, as we will have no reactions of the form $\varnothing \to A$ for any species A, so if no species are initially present, no species (including Y^+) will ever be produced. For each nonempty U, we compute f_U independently in parallel by CRN $\mathcal{C}_U$, modifying each reaction producing Y^+ to produce an equivalent amount of species Y_U, which is specific to U. For each such U there are inactive and active "indicator" species J_U and I_U. In parallel, there are reactions that will activate indicator species I_U (i.e. convert J_U to I_U) if and only if all species X_i are present initially for each $i \in U$. These I_U species will then counteract the effect of any CRN computing $f_{U'}$ for $U' \subset U$ by catalytically converting all $Y_{U'}$ to Y^-. If

U is the complete set of indices of non-zero inputs, then only CRNs computing $f_{U'}$ for subsets $U' \subset U$ have produced any amount of Y^+, so eventually all of these will be counteracted by I_U.

Formally, construct the CRN as follows. For each $i \in \{1,\ldots,k\}$, add the reaction $X_i \to I_{\{i\}} + J_{\{i\}} + J_{U_1} + X_i^{U_1} + J_{U_2} + X_i^{U_2} + \ldots + J_{U_l} + X_i^{U_l}$, where $U_1, U_2, \ldots, U_l$ are all subsets of $\{1,\ldots,k\}$ that are strict supersets of $\{i\}$. The extra versions of X_i are used as inputs to the parallel computation of each f_U. We generate the inactive indicator species from the input species in this manner, because the CRN is not allowed to start with anything other than the input.

The indicator species are activated as follows. For each nonempty $U, U' \subseteq \{1,\ldots,k\}$ such that $U \neq U'$, add the reaction $I_U + I_{U'} + J_{U\cup U'} \to I_U + I_{U'} + I_{U\cup U'}$.

For each nonempty $U \subseteq \{1,\ldots,k\}$, let $\mathcal{C}_U$ be the CRN computing an output dual-rail representation of f_U (i.e. dual rail on the output). Modify $\mathcal{C}_U$ as follows. Rename the output species of $\mathcal{C}_U$ to Y^+ and Y^-, i.e., all parallel CRNs share the same output species. For each reaction producing the output species Y^+, add the product Y_U^+ (which is a species specific to $\mathcal{C}_U$) with the same net stoichimetry. Similarly, for each reaction producing the output species Y^-, add the product Y_U^- with the same net stoichimetry. For instance, replace the reaction $A + B \to Y^+$ by the reaction $A + B \to Y^+ + Y_U^+$, and replace the reaction $A + Y^+ \to B + 4Y^+$ by the reaction $A + Y^+ \to B + 4Y^+ + 3Y_U^+$. Therefore the eventual amount of Y_U^+ is equal to the total amount of Y^+ produced by $\mathcal{C}_U$, and similarly for Y_U^- and Y^-. For each $U' \subset U$, add the reactions $I_U + Y_{U'}^+ \to I_U + Y^-$, $I_U + Y_{U'}^- \to I_U + Y^+$. Also, for each reaction in $\mathcal{C}_U$, add I_U as a catalyst. This ensures that $\mathcal{C}_U$ cannot execute any reactions (and therefore cannot produce any amount of Y^+ or Y^-) unless all species X_i for $i \in U$ are present. Finally, add the reaction $Y^+ + Y^- \to \varnothing$. □

The following lemma shows that our construction actually converges to the correct output value under mass-action kinetics; it follows similarly to Lemma 3.4.

LEMMA 4.3. *For any input state* $\mathbf{x}$*, for any state* $\mathbf{z}$ *reachable from* $\mathbf{x}$*, the mass-action trajectory of the CRN of Lemma 4.2 with any non-zero rate constants starting at* $\mathbf{z}$ *converges to the output stable state in the limit* $t \to \infty$*.*

4.2 Computable functions are positive-continuous piecewise rational linear

In this section we prove that stable CRNs can compute only the positive-continuous piecewise linear functions.

LEMMA 4.4. *Let* $f : \mathbb{R}^k_{\geq 0} \to \mathbb{R}_{\geq 0}$ *be stably computable by a CRC. Then* f *is positive-continuous and piecewise rational linear.*

PROOF. Let $\mathcal{C}$ be the CRC stably computing f, with input species $X_1, \ldots, X_k$ and output species Y. By Lemma 3.12, f is positive-continuous.

It remains to show that f is piecewise rational linear. If $\mathbf{x}$ is an initial state and $\mathbf{x} \to \mathbf{o}$ where $\mathbf{o}$ is output stable, then by Lemma 2.4, $\mathbf{x} \rightsquigarrow^{m+1} \mathbf{o}$. Let Ω be an output siphon such that $\mathbf{o} \restriction \Omega = \mathbf{0}$, which exists by Lemma 3.5. Let $\mathbf{s} = (R_1, \ldots, R_l) \in \mathcal{P}(R)^l$, where $\mathbf{x} \to_{\mathbf{s}} \mathbf{o}$, where by Lemma 2.4 we may assume $l \leq m + 1$. By Lemma 3.9, this value of y is a rational linear function of $\mathbf{x}$. Since there are a finite number of output siphons and sequences in $\mathcal{P}(R)^{\leq m+1}$, this establishes the piecewise rational linearity of f. □

5. ACKNOWLEDGEMENTS

We thank Manoj Gopalkrishnan, Elisa Franco, Damien Woods, and the organizers and participants of the American Mathematical Institute workshop on Mathematical Problems Arising from Biochemical Reaction Networks for insightful discussions. We are grateful to anonymous reviewers for insightful comments and suggestions that have improved this paper. DD was supported by the Molecular Programming Project under NSF grants 0832824 and 1317694 and by NSF grants CCF-1219274 and CCF-1162589. DS was supported by NIGMS Systems Biology Center grant P50 GM081879.

6. REFERENCES

[1] U. Alon. *Introduction to Systems Biology: Design Principles of Biological Circuits*. CRC press, 2006.

[2] D. Angeli, P. De Leenheer, and E. D. Sontag. On the structural monotonicity of chemical reaction networks. In *45th IEEE Conference on Decision and Control*, pages 7–12. IEEE, 2006.

[3] D. Angeli, P. De Leenheer, and E. D. Sontag. A Petri net approach to the study of persistence in chemical reaction networks. *Mathematical Biosciences*, 210(2):598–618, 2007.

[4] D. Angluin, J. Aspnes, and D. Eisenstat. Stably computable predicates are semilinear. In *PODC 2006: Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*, pages 292–299, New York, NY, USA, 2006. ACM Press.

[5] N. Barkai and S. Leibler. Robustness in simple biochemical networks. *Nature*, 387(6636):913–917, 1997.

[6] O. Bournez, P. Chassaing, J. Cohen, L. Gerin, and X. Koegler. On the convergence of population protocols when population goes to infinity. *Applied Mathematics and Computation*, 215(4):1340–1350, 2009.

[7] L. Cardelli. Strand algebras for DNA computing. *Natural Computing*, 10(1):407–428, 2011.

[8] L. Cardelli and A. Csikász-Nagy. The cell cycle switch computes approximate majority. *Scientific Reports*, 2, 2012.

[9] H.-L. Chen, D. Doty, and D. Soloveichik. Deterministic function computation with chemical reaction networks. *Natural Computing*, 2013. to appear. Preliminary version appeared in DNA 2012.

[10] Y.-J. Chen, N. Dalchau, N. Srinivas, A. Phillips, L. Cardelli, D. Soloveichik, and G. Seelig. Programmable chemical controllers made from DNA. *Nature Nanotechnology*, 8(10):755–762, 2013.

[11] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[12] M. Gopalkrishnan, E. Miller, and A. Shiu. A projection argument for differential inclusions, with applications to persistence of mass-action kinetics. *Symmetry, Integrability and Geometry: Methods and Applications*, 9(0):25–25, 2013.

[13] H. K. Khalil. *Nonlinear systems*. Prentice Hall, 2002.

[14] T. G. Kurtz. The relationship between stochastic and deterministic models for chemical reactions. *The Journal of Chemical Physics*, 57(7):2976–2978, 1972.

[15] S. Ovchinnikov. Max-min representation of piecewise linear functions. *Contributions to Algebra and Geometry*, 43(1):297–302, 2002.

[16] L. Qian and E. Winfree. Scaling up digital circuit computation with DNA strand displacement cascades. *Science*, 332(6034):1196, 2011.

[17] M. S. Samoilov and A. P. Arkin. Deviant effects in molecular reaction pathways. *Nature biotechnology*, 24(10):1235–1240, 2006.

[18] G. Seelig, D. Soloveichik, D. Y. Zhang, and E. Winfree. Enzyme-free nucleic acid logic circuits. *Science*, 314(5805):1585–1588, 2006.

[19] D. Soloveichik, M. Cook, E. Winfree, and J. Bruck. Computation with finite stochastic chemical reaction networks. *Natural Computing*, 7(4):615–633, 2008.

[20] D. Soloveichik, G. Seelig, and E. Winfree. DNA as a universal substrate for chemical kinetics. *Proceedings of the National Academy of Sciences*, 107(12):5393, 2010. Preliminary version appeared in DNA 2008.