Timing in chemical reaction networks

David Doty[*]

Abstract

Chemical reaction networks (CRNs) formally model chemistry in a well-mixed solution. CRNs are widely used to describe information processing occurring in natural cellular regulatory networks, and with upcoming advances in synthetic biology, CRNs are a promising programming language for the design of artificial molecular control circuitry. Due to a formal equivalence between CRNs and a model of distributed computing known as *population protocols*, results transfer readily between the two models.

We show that if a CRN respects *finite density* (at most $O(n)$ additional molecules can be produced from n initial molecules), then starting from any *dense* initial configuration (all molecular species initially present have initial count $\Omega(n)$, where n is the initial molecular count and volume), *every* producible species is produced in constant time with high probability.

This implies that no CRN obeying the stated constraints can function as a *timer*, able to produce a molecule, but doing so only after a time that is an unbounded function of the input size. This has consequences regarding an open question of Anglin, Aspnes, and Eisenstat concerning the ability of population protocols to perform fast, reliable leader election and to simulate arbitrary algorithms from a uniform initial state.

1 Introduction

1.1 Background of the field The engineering of complex artificial molecular systems will require a sophisticated understanding of how to *program chemistry*. A natural language for describing the interactions of molecular species in a well-mixed solution is that of (finite) chemical reaction networks (CRNs), i.e., finite sets of chemical reactions such as $A + B \rightarrow A + C$. When the behavior of individual molecules is modeled, CRNs are assigned semantics through *stochastic chemical kinetics* [15], in which reactions occur probabilistically with rate proportional to the product of the molecular count of their reactants and inversely proportional to the volume of the reaction vessel. The kinetic model of CRNs is based on the physical assumption of well-mixedness that is valid in a dilute solution. Thus, we assume the *finite density constraint*, which stipulates that a volume required to execute a CRN must be proportional

to the maximum molecular count obtained during execution [24]. In other words, the total concentration (molecular count per volume) is bounded. This realistically constrains the speed of the computation achievable by CRNs.

Traditionally CRNs have been used as a descriptive language to analyze naturally occurring chemical reactions, as well as numerous other systems with a large number of interacting components such as gene regulatory networks and animal populations. However, recent investigations have viewed CRNs as a programming language for engineering artificial systems. These works have shown CRNs to have eclectic algorithmic abilities. Researchers have investigated the power of CRNs to simulate Boolean circuits [20], neural networks [16], and digital signal processing [17]. Other work has shown that bounded-space Turing machines can be simulated with an arbitrarily small, non-zero probability of error by a CRN with only a polynomial slowdown [4].[1] It is EXPSPACE-hard to predict whether a particular species is producible [19]. Space- and energy-efficient simulation of space-bounded Turing machines can be done by CRNs implementable by logically and thermodynamically reversible DNA strand displacement reactions, and as a consequence this prediction problem is PSPACE-hard even when restricted to such CRNs [26]. Even Turing-universal computation is possible with an arbitrarily small probability of error over all time [24]. The computational power of CRNs also provides insight on why it can be computationally difficult to simulate them [23], and why certain questions are frustratingly difficult to answer (e.g. undecidable) [13, 28]. The programming approach to CRNs has also, in turn, resulted in novel insights regarding natural cellular regulatory networks [7].

1.2 Focus of this paper Informally, our main theorem shows that CRNs "respecting finite density" (meaning that the total molecular count obtainable is bounded by a constant times the initial molecular count) start-

---

[1]This is surprising since finite CRNs necessarily must represent binary data strings in a unary encoding, since they lack positional information to tell the difference between two molecules of the same species.

ing from any *dense* initial configuration (meaning that any species present initially has count $\Omega(n)$, where n is the initial molecular count and volume) will create $\Omega(n)$ copies of *every* producible species in constant time with high probability. We now explain the significance of this result in the context of an open question regarding the ability of stochastic CRNs to do fast, reliable computation.

Many of the fundamental algorithmic abilities and limitations of CRNs were elucidated under the guise of a related model of distributed computing known as *population protocols*. Population protocols were introduced by Angluin, Aspnes, Diamadi, Fischer, and Peralta [1] as a model of resource-limited mobile sensors; see Aspnes and Ruppert [5] for an excellent survey of the model. A population protocol consists of a set of n agents with finite state set Λ (where we imagine $|\Lambda|$ is constant with respect to n), together with a transition function $\delta : \Lambda^2 \to \Lambda^2$, with $(r_1, r_2) = \delta(q_1, q_2)$ indicating that if two agents in states q_1 and q_2 interact, then they change to states r_1 and r_2, respectively. Multiple semantic models may be overlaid on this syntactic definition, but the randomized model studied by Angluin, Aspnes, and Eisenstat [4], in which the next two agents to interact are selected uniformly at random, coincides precisely with the model of stochastic chemical kinetics [15], so long as the CRN's reactions each have two reactants and two products,[2] i.e., $q_1 + q_2 \to r_1 + r_2$. In fact, every population protocol, when interpreted as a CRN, respects finite density, because the total molecular count is constant over time. Therefore, although we state our result in the language of CRNs, the results apply *a fortiori* to population protocols.

Angluin, Aspnes, and Eisenstat [4], and independently Soloveichik, Cook, Winfree, and Bruck [24] showed that arbitrary Turing machines may be simulated by randomized population protocols/CRNs with only a polynomial-time slowdown. Because a binary input $x \in \{0,1\}^*$ to a Turing machine must be specified in a CRN by a unary molecular count $n \in \mathbb{N}$, where $n \approx 2^{|x|}$, this means that each construction is able to simulate each step of the Turing machine in time $O(\text{polylog}(n))$, i.e. polynomial in $|x|$. Both constructions make essential use of the ability to specify *arbitrary* initial configurations, where a configuration is a vector $\mathbf{c} \in \mathbb{N}^\Lambda$ specifying the initial count $\mathbf{c}(S)$ of each species $S \in \Lambda$. In particular, both constructions require a certain species to be present with initial count 1, a so-called "leader" (hence the title of the former paper).

A test tube with a single copy of a certain type of molecule is experimentally difficult to prepare. The major open question of [4] asks whether this restriction may be removed: whether there is a CRN starting from a *uniform* initial configuration (i.e., with only a single species present, whose count is equal to the volume n) that can simulate a Turing machine.

Angluin, Aspnes, and Eisenstat [4] observe that a leader may be elected from a uniform initial configuration of n copies of L by the reaction $L + L \to L + N$; in other words, when two candidate leaders encounter each other, one drops out. However, this scheme has problems:

1. The expected time until a single leader remains is $\Theta(n)$, i.e., exponential in m, where $m \approx \log n$ is the number of bits required to represent the input n. In contrast, if a leader is assumed present from the start, the computation takes time $O(t(\log n) \cdot \log^5 n)$, where $t(m)$ is the running time of the Turing machine on an m-bit input, i.e., the time is polynomial in t. Therefore for polynomial-time computations, electing a leader in this manner incurs an exponential slowdown.

2. There is no way for the leader to "know" when it has been elected. Therefore, even if a single leader is eventually elected, and even if the CRN works properly under the assumption of a single initial leader, prior to the conclusion of the election, multiple leaders will be competing and introducing unwanted behavior.

These problems motivate the need for a *timer* CRN, a CRN that, from a uniform initial configuration of size n, is able to produce some species S, but the expected time before the first copy of S is produced is $t(n)$, for some unbounded function $t : \mathbb{N} \to \mathbb{N}$. If a CRN exists that can produce a copy of S after $\Omega(n)$ time, then by adding the reaction $S + L \to L_{\text{active}}$, a leader election can take place among "inactive" molecules of type L, and the remaining species of the leader-driven CRNs of [4] or [24] can be made to work only with the "active" species L_{active}, which will not appear (with high probability) until there is only a single copy of L to be activated. This shows how a timer CRN could alleviate the second problem mentioned above. In fact, even the first problem, the slowness of the naïve leader election algorithm, is obviated in the presence of a timer CRN. If a timer CRN produces S after $t(n) = \Omega(\log n)$ time from a uniform initial configuration of n copies of X, then this can be used to construct a CRN that elects a leader in time $O(t(n))$ [10]. In other words, the problem of constructing a timer CRN is "leader-election-hard".

---

[2]This turns out not to be a significant restriction on CRNs when considering their computational ability; see the introduction of [8] for a discussion of the issue.

Unfortunately, timer CRNs cannot be constructed under realistic conditions. The main theorem of this paper, stated informally, shows the following (the formal statement is Theorem 3.1 in Section 3): Let $\mathcal{C}$ be a CRN with volume and initial molecular count n that *respects finite density* (no sequence of reactions can produce more than $O(n)$ copies of any species) with an initial configuration that is *dense* (all species initially present have initial count $\Omega(n)$). Then with probability $\geq 1 - 2^{-\Omega(n)}$, every producible species is produced (in fact, attains count $\Omega(n)$) in time $O(1)$. Since a uniform initial configuration is dense, timer CRNs that respect finite density cannot be constructed. It should be noted that the condition of respecting finite density is physically realistic; it is obeyed, for instance, by every CRN that obeys the law of conservation of mass. As noted previously, all population protocols respect finite density, which implies that, unconditionally, *no* population protocol can function as a timer.

This theorem has the following consequences for CRNs respecting finite density that start from a dense initial configuration with n molecules:

1. No such CRN can decide a predicate "monotonically". By this we mean that if the CRN solves a decision problem (a yes/no question about an input encoded in its initial configuration) by producing species Y if the answer is yes and species N if the answer is no, then the CRN cannot be guaranteed to produce *only* the correct species.[3] It necessarily must produce *both* but eventually dispose of the incorrect species. This implies that composing such a CRN with a downstream CRN that uses the answer necessarily requires the downstream CRN to account for the possibility that the upstream CRN will "change its mind" before converging on the correct answer.

2. No CRN leader election algorithm can "avoid war": any CRN that elects a unique leader (in any amount of time) must necessarily have at least 2 (in fact, at least $\Omega(n)$) copies of the leader species for some time before the unique leader is elected. This implies that any downstream CRN requiring a leader must be designed to work in the face of multiple leaders being present simultaneously for some amount of time before a unique leader is finally elected.

We emphasize that we are *not* claiming "$O(1)$-time leader election is impossible" as a consequence of our main theorem.[4] The two consequences above merely outline what sort of behavior one can expect from a CRN with a uniform initial configuration, due to our main theorem.

Angluin, Aspnes, Diamadi, Fischer, and Peralta, in an unpublished technical report [2, Lemma 5], sketched an informal argument for a similar result in a related model called *urn automata*. However, this was a limiting argument without bounds on convergence, whereas we give (and formally prove) probability bounds that scale exponentially in the size of the system. It may also be possible to derive a different proof of our main result using Kurtz's theorem [18] that every stochastic CRN, in the limit of large volume and molecular count scaling, for some amount of time will likely track closely the behavior of a deterministic mass-action system whose behavior is given by differential equations. A similar, more recent, result known as Wormald's theorem [27] might also be useful in this sense. However, these are also limit arguments. Furthermore, we believe that a direct proof in the stochastic model, rather than appealing to approximations by differential equations, gives insight as to *why* the result is true and enhances our understanding of the behaviors stochastic CRNs allow.

Another line of research that has illustrated a striking difference between "chemical computation" and standard computational models has been the role of randomness. The model of stochastic chemical kinetics is one way to impose semantics on CRNs, but another reasonable alternative is to require that all possible sequences of reactions – and not just those likely to occur by the model of chemical kinetics – should allow the CRN to carry out a correct computation. Under this more restrictive deterministic model of CRN computation, Angluin, Aspnes, and Eisenstat [3] showed that precisely the *semilinear predicates* $\phi : \mathbb{N}^k \to \{0,1\}$ (those definable in the first-order theory of Presburger arithmetic [22]) are deterministically computed by CRNs. This was subsequently extended to *function* computation by Chen, Doty, and Soloveichik [8], who showed that precisely the functions $f : \mathbb{N}^k \to \mathbb{N}^l$ whose graph is a semilinear set are deterministically computed by CRNs. Since semilinear sets are compu-

---

[3]This assumes the incorrect species is producible at all. This is not a strong assumption. From any initial state $\mathbf{x}$ in which the correct answer is yes (for example), if any substate $\mathbf{x}' < \mathbf{x}$ has correct answer no, then N is producible from $\mathbf{x}$ because the presence of additional molecules cannot prevent a sequence of reactions from being possible to execute.

[4]This claim is easy to show using a simpler argument: $\Omega(\log n)$ time is required for every agent to interact with every other. Furthermore, the claim does not even follow directly from our main theorem, since (assuming ignorance of the $\Omega(\log n)$ argument), *a priori*, it is conceivable that all species are produced in the first second, then the next four seconds all leaders but one are eliminated.

tationally very weak (one characterization is that they are finite unions of "periodic" sets, suitably generalizing the definition of periodic to multi-dimensional spaces), this shows that introducing randomization adds massive computational ability to CRNs.

This strongly contrasts other computational models. Turing machines decide the decidable languages whether they are required to be deterministic or randomized (or even nondeterministic). In the case of polynomial-time Turing machines, it is widely conjectured [21] that $\mathsf{P} = \mathsf{BPP}$, i.e., that randomization adds at most a polynomial speedup to any predicate computation. Since it is known that $\mathsf{P} \subseteq \mathsf{BPP} \subseteq \mathsf{EXP}$, even the potential exponential gap between deterministic and randomized polynomial-time computation is dwarfed by the gap between deterministic CRN computation (semilinear sets, which are all decidable in linear time by a Turing machine) and randomized CRN computation (which can decide any decidable problem).

Along this line of thinking (*"What power does the stochastic CRN model have over the deterministic model?"*), our main theorem may also be considered a stochastic extension of work on deterministic CRNs by Condon, Hu, Maňuch, and Thachuk [11] and Condon, Kirkpatrick, and Maňuch [12]. Both of those papers showed results of the following form: every CRN in a certain restricted class (the class being different in each paper, but in each case is a subset of the class of CRNs respecting finite density) with k species and initial configuration $\mathbf{i} \in \mathbb{N}^k$ has the property that every producible species is producible from initial configuration $m \cdot \mathbf{i}$ through at most t *reactions*, where m and t are bounded by a polynomial in k. In other words, if the goal of the CRN is to delay the production of some species until $\omega(\text{poly}(k))$ reactions have occurred (such as the "Gray code counter" CRN of [11], which iterates through 2^k states before producing the first copy of a certain species), then the CRN cannot be *multicopy tolerant*, since m copies of the CRN running in parallel can "short-circuit" and produce every species in $\text{poly}(k)$ reactions, if the reactions are carefully chosen. Our main theorem extends these deterministic impossibility results to the stochastic model, showing that not only is there a short sequence of reactions that will produce every species, but furthermore that the laws of chemical kinetics will force such a sequence to actually happen in constant time with high probability.

The paper is organized as follows. Section 2 defines the model of stochastic chemical reaction networks. Section 3 states and proves the main theorem, Theorem 3.1, that every CRN respecting finite density, starting from a dense initial configuration, likely produces every producible species in constant time. Appendix A proves

a Chernoff bound on continuous-time stochastic exponential decay processes, Lemma A.1, used in the proof of Theorem 3.1. Appendix B proves a Chernoff bound on continuous-time biased random walks, Lemma B.6, a messy-to-state consequence of which (Lemma B.7) is used in the proof of Theorem 3.1. Section 4 discusses questions for future work.

2 Preliminaries

2.1 Chemical reaction networks If Λ is a finite set of chemical species, we write $\mathbb{N}^\Lambda$ to denote the set of functions $f : \Lambda \to \mathbb{N}$. Equivalently, we view an element $\mathbf{c} \in \mathbb{N}^\Lambda$ as a vector of $|\Lambda|$ nonnegative integers, with each coordinate "labeled" by an element of Λ. Given $S \in \Lambda$ and $\mathbf{c} \in \mathbb{N}^\Lambda$, we refer to $\mathbf{c}(S)$ as the *count of S in $\mathbf{c}$*. Let $\|\mathbf{c}\| = \|\mathbf{c}\|_1 = \sum_{S \in \Lambda} \mathbf{c}(S)$ represent the total count of species in $\mathbf{c}$. We write $\mathbf{c} \le \mathbf{c}'$ to denote that $\mathbf{c}(S) \le \mathbf{c}'(S)$ for all $S \in \Lambda$. Given $\mathbf{c}, \mathbf{c}' \in \mathbb{N}^\Lambda$, we define the vector component-wise operations of addition $\mathbf{c} + \mathbf{c}'$, subtraction $\mathbf{c} - \mathbf{c}'$, and scalar multiplication $n\mathbf{c}$ for $n \in \mathbb{N}$. If $\Delta \subset \Lambda$, we view a vector $\mathbf{c} \in \mathbb{N}^\Delta$ equivalently as a vector $\mathbf{c} \in \mathbb{N}^\Lambda$ by assuming $\mathbf{c}(S) = 0$ for all $S \in \Lambda \setminus \Delta$. For all $\mathbf{c} \in \mathbb{N}^\Lambda$, let $[\mathbf{c}] = \{ S \in \Lambda \mid \mathbf{c}(S) > 0 \}$ be the set of species with positive counts in $\mathbf{c}$, so that $S \in [\mathbf{c}] \iff \mathbf{c}(S) > 0$.

Given a finite set of chemical species Λ, a *reaction* over Λ is a triple $\beta = (\mathbf{r}, \mathbf{p}, k) \in \mathbb{N}^\Lambda \times \mathbb{N}^\Lambda \times \mathbb{R}^+$, specifying the stoichiometry of the reactants and products, respectively, and the *rate constant* k. For instance, given $\Lambda = \{A, B, C\}$, the reaction $A + 2B \xrightarrow{4.7} A + 3C$ is the triple $((1, 2, 0), (1, 0, 3), 4.7)$. 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 *configuration* of a CRN $\mathcal{C} = (\Lambda, R)$ is a vector $\mathbf{c} \in \mathbb{N}^\Lambda$. When the configuration $\mathbf{c}$ is clear from context, we write $\#X$ to denote $\mathbf{c}(X)$.

Given a configuration $\mathbf{c}$ and reaction $\beta = (\mathbf{r}, \mathbf{p}, k)$, we say that β is *applicable* to $\mathbf{c}$ if $\mathbf{r} \le \mathbf{c}$ (i.e., $\mathbf{c}$ contains enough of each of the reactants for the reaction to occur). If β is applicable to $\mathbf{c}$, then write $\beta(\mathbf{c})$ to denote the configuration $\mathbf{c} + \mathbf{p} - \mathbf{r}$ (i.e., the configuration that results from applying reaction β to $\mathbf{c}$). If $\mathbf{c}' = \beta(\mathbf{c})$ for some reaction $\beta \in R$, we write $\mathbf{c} \to_\mathcal{C} \mathbf{c}'$, or merely $\mathbf{c} \to \mathbf{c}'$ when $\mathcal{C}$ is clear from context. An *execution* (a.k.a., *execution sequence*) $\mathcal{E}$ is a finite or infinite sequence of one or more configurations $\mathcal{E} = (\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2, \dots)$ such that, for all $i \in \{1, \dots, |\mathcal{E}|\}$, $\mathbf{c}_{i-1} \to \mathbf{c}_i$. If a finite execution sequence starts with $\mathbf{c}$ and ends with $\mathbf{c}'$, we write $\mathbf{c} \to_\mathcal{C}^* \mathbf{c}'$, or merely $\mathbf{c} \to^* \mathbf{c}'$ when the CRN $\mathcal{C}$ is clear from context. In this case, we say that $\mathbf{c}'$ is *reachable* from $\mathbf{c}$.

We say that a reaction $\beta = (\mathbf{r}, \mathbf{p}, k)$ *produces* $S \in \Lambda$ if $\mathbf{p}(S) - \mathbf{r}(S) > 0$, and that β *consumes* S if

$\mathbf{r}(S) - \mathbf{p}(S) > 0$. When a CRN $\mathcal{C} = (\Lambda, R)$ and an initial configuration $\mathbf{i} \in \mathbb{N}^\Lambda$ are clear from context, we say a species $S \in \Lambda$ is *producible* if there exists $\mathbf{c}$ such that $\mathbf{i} \to_{\mathcal{C}}^* \mathbf{c}$ and $\mathbf{c}(S) > 0$.

2.2 Kinetic model The following model of stochastic chemical kinetics is widely used in quantitative biology and other fields dealing with chemical reactions in which stochastic effects are significant [15]. It ascribes probabilities to execution sequences, and also defines the time of reactions, allowing us to study the running time of CRNs in Section 3.

A reaction is *unimolecular* if it has one reactant and *bimolecular* if it has two reactants. We assume no higher-order reactions occur.[5] The kinetic behavior of a CRN is described by a continuous-time Markov process, whose states are configurations of the CRN, as follows. Given a fixed volume $v > 0$ and current configuration $\mathbf{c}$, the *propensity* of a unimolecular reaction $\beta : X \xrightarrow{k} \dots$ in configuration $\mathbf{c}$ is $\rho(\mathbf{c}, \beta) = k\mathbf{c}(X)$. The propensity of a bimolecular reaction $\beta : X + Y \xrightarrow{k} \dots$, where $X \neq Y$, is $\rho(\mathbf{c}, \beta) = \frac{k}{v}\mathbf{c}(X)\mathbf{c}(Y)$. The propensity of a bimolecular reaction $\beta : X + X \xrightarrow{k} \dots$ is $\rho(\mathbf{c}, \beta) = \frac{k}{v}\frac{\mathbf{c}(X)(\mathbf{c}(X)-1)}{2}$.[6] The propensity function determines the evolution of the system as follows. The time until the next reaction occurs is an exponential random variable with rate $\rho(\mathbf{c}) = \sum_{\beta \in R} \rho(\mathbf{c}, \beta)$. The probability that next reaction will be a particular β_{next} is $\frac{\rho(\mathbf{c}, \beta_{\text{next}})}{\rho(\mathbf{c})}$.

3 CRNs that produce all species in constant time

We say a CRN *respects finite density* if there is a constant $\hat{c}$ such that, for any initial configuration $\mathbf{i}$ and any configuration $\mathbf{c}$ such that $\mathbf{i} \to^* \mathbf{c}$, $\|\mathbf{c}\| \leq \hat{c}\|\mathbf{i}\|$. That is, the maximum molecular count attainable is bounded by a constant multiplicative factor of the initial counts, implying that if the volume is sufficiently large to contain the initial molecules, then within a constant factor, it is sufficiently large to contain all the molecules ever produced. We therefore safely assume that for any CRN respecting finite density with initial configuration $\mathbf{i}$, the volume is $\|\mathbf{i}\|$.

CRNs respecting finite density constitute a wide class of CRNs that includes, for instance, all mass-conserving CRNs: CRNs for which there exists a mass

---

function $m : \Lambda \to \mathbb{R}^+$ such that $\sum_{S \in \mathbf{r}} m(S) = \sum_{S \in \mathbf{p}} m(S)$ for all reactions $(\mathbf{r}, \mathbf{p}, k)$. This includes all population protocols [5].

In this section we prove that no CRN respecting finite density, starting from an initial configuration with "large" species counts, can delay the production of any producible species for more than a constant amount of time. All producible species are produced in time $O(1)$ with high probability.

Let $\alpha > 0$. We say that a configuration $\mathbf{c}$ is α-*dense* if for all $S \in \Lambda$, $S \in [\mathbf{c}] \implies \mathbf{c}(S) \geq \alpha\|\mathbf{c}\|$, i.e., all species present initially occupy a constant (α) fraction of the total volume. Given initial configuration $\mathbf{i} \in \mathbb{N}^\Lambda$, let $\Lambda_{\mathbf{i}}^* = \{ S \in \Lambda \mid (\exists \mathbf{c}) \; \mathbf{i} \to^* \mathbf{c}$ and $\mathbf{c}(S) > 0 \}$ denote the set of species producible from $\mathbf{i}$. For all $t > 0$ and $S \in \Lambda$, let $\#_t S$ be the random variable representing the count of S after t seconds, if $\mathbf{i}$ is the initial configuration at time 0.

The following is the main theorem of this paper.

THEOREM 3.1. *Let $\alpha > 0$ and $\mathcal{C} = (\Lambda, R)$ be a CRN respecting finite density. Then there are constants $\epsilon, \delta, t > 0$ such that, for all sufficiently large n (how large depending only on α and $\mathcal{C}$), for all α-dense configurations $\mathbf{i}$ with $\|\mathbf{i}\| = n$, $\Pr[(\forall S \in \Lambda_{\mathbf{i}}^*) \; \#_t S \geq \delta n] \geq 1 - 2^{-\epsilon n}$.*

Proof. The handwaving intuition of the proof is as follows. We first show that every species initially present, because they have count $\Omega(n)$ and decay at an at most exponential rate, remain at count $\Omega(n)$ (for some smaller constant fraction of n) for a constant amount of time. Because of this, during this entire time, the reactions for which they are reactants are running at rate $\Omega(n)$. Therefore the products of those reactions are produced quickly enough to get to count $\Omega(n)$ in time $O(1)$. Because *those* species (the products) decay at an at most exponential rate, they have count $\Omega(n)$ for a constant amount of time. Therefore the reactions for which they are *reactants* are running at rate $\Omega(n)$. Since there are a constant number of reactions, this will show that all producible species are produced in constant time with high probability. The tricky part is to show that although species may be consumed and go to lower counts, every species produced has count $\Omega(n)$ for $\Omega(1)$ time, sufficient for it to execute $\Omega(n)$ reactions of which it is a reactant and produce $\Omega(n)$ of the products of that reaction.

On to the gritty details. Because $\mathcal{C}$ respects finite density, the maximum count obtained by any species is $O(n)$. Let $\hat{c}$ be a constant such that all species have count at most $\hat{c}n$ in any configuration reachable from $\mathbf{i}$. Let $\hat{K} = \sum_{(\mathbf{r}, \mathbf{p}, k) \in R} k$ be the sum of the rate constants of every reaction. Let $\hat{k} = \min_{(\mathbf{r}, \mathbf{p}, k) \in R} k$

---

be the minimum rate constant of any reaction. For any $\Delta \subseteq \Lambda$, define $\text{PROD}(\Delta) = \{S \in \Lambda \mid (\exists \beta = (\mathbf{r}, \mathbf{p}, k) \in R) \ \beta$ produces S and $[\mathbf{r}] \subseteq \Delta\}$ to be the set of species producible by a single reaction assuming that only species in Δ are present, and further assuming that those species have sufficient counts to execute the reaction.

Define subsets $\Lambda_0 \subset \Lambda_1 \subset \ldots \subset \Lambda_{m-1} \subset \Lambda_m \subseteq \Lambda$ as follows. Let $\Lambda_0 = [\mathbf{i}]$. For all $i \in \mathbb{Z}^+$, define $\Lambda_i = \Lambda_{i-1} \cup \text{PROD}(\Lambda_{i-1})$. Let m be the smallest integer such that $\text{PROD}(\Lambda_m) \subseteq \Lambda_m$. Therefore for all $i \in \{1, \ldots, m\}$, $|\Lambda_i| > |\Lambda_{i-1}|$, whence $m < |\Lambda|$.

By the hypothesis of the theorem, all $S \in \Lambda_0$ satisfy $\mathbf{i}(S) \geq \alpha n$. S may be produced and consumed. To establish that $\#S$ remains high for a constant amount of time, in the worst case we assume that S is only consumed. Let $\lambda = \hat{c}\hat{K}$. We will assume that $\lambda \geq 1$ since we can choose $\hat{c} \geq \frac{1}{\hat{K}}$ if it is not already greater than $\frac{1}{\hat{K}}$. This assumption is required in Lemma B.7, which is employed later in the proof. We also assume that $\hat{c} \geq 1$, implying $\lambda \geq \hat{K}$.

For all $S \in \Lambda$, the rate of any unimolecular reaction consuming S is at most $\hat{K}\#S \leq \lambda\#S$, and the rate of any bimolecular reaction consuming S is at most $\frac{\hat{K}}{n}(\hat{c}n)\#S = \lambda\#S$, whence the expected time for any reaction consuming S is at least $\frac{1}{\lambda\#S}$. We can thus upper bound the overall consumption of S by an exponential decay process with rate λ, i.e., this process will consume copies of S at least as quickly as the actual CRN consumes copies of S.

Let $c = 4e^{\lambda(m+1)}$. Let $\delta_0 = \frac{\alpha}{c}$. Fix a particular $S \in \Lambda_0$. Let $N = \alpha n$ and $\delta = \frac{1}{c}$ in Lemma A.1. Then by the fact that the rate of decay of S is bounded by an exponential decay process with rate λ and initial value αn (the process $\mathbf{D}_\lambda^N$ as defined in Section A) and Lemma A.1,

$$
\begin{aligned}
&\Pr[(\exists t \in [0, m+1]) \ \#_t S < \delta_0 n] \\
&\leq \ \Pr\left[\mathbf{D}_\lambda^{\alpha n}(m+1) < \frac{1}{c}\alpha n\right] \\
&< \ \left(2\frac{1}{c}e^{\lambda(m+1)}\right)^{\alpha n/c - 1} = 2^{-\delta_0 n + 1}.
\end{aligned}
$$

By the union bound,
(3.1)
$$\Pr[(\exists S \in \Lambda_0)(\exists t \in [0, m+1]) \ \#_t S < \delta_0 n] < |\Lambda_0| 2^{-\delta_0 n + 1}.$$

That is, with high probability, all species in Λ_0 have "high" count (at least $\delta_0 n$) for the entire first $m + 1$ seconds. Call this event $H(\Lambda_0)$ (i.e., the complement of the event in (3.1)).

We complete the proof by a "probabilistic induc-

tion"[7] on $i \in \{0, 1, \ldots, m\}$ as follows. Inductively assume that for all $X \in \Lambda_i$ and all $t \in [i, m+1]$, $\#_t X \geq \delta_i n$ for some $\delta_i > 0$. Call this event $H(\Lambda_i)$. Then we will show that for all $S \in \Lambda_{i+1}$ and all $t \in [i+1, m+1]$, assuming $H(\Lambda_i)$ is true, with high probability $\#_t S \geq \delta_{i+1} n$ for some $\delta_{i+1} > 0$. We will use Lemma B.7 to choose particular values for the δ_i's, and these values will not depend on n.

The base case is established by (3.1) for $\delta_0 = \frac{\alpha}{c}$. Fix a particular species $S \in \Lambda_{i+1}$. By the definition of Λ_{i+1}, it is produced by either at least one reaction of the form $X \to S + \ldots$ for some $X \in \Lambda_i$ or by $X + Y \to S + \ldots$ for some $X, Y \in \Lambda_i$. By the induction hypothesis $H(\Lambda_i)$, for all $t \in [i, m+1]$, $\#_t X \geq \delta_i n$ and $\#_t Y \geq \delta_i n$ for some $\delta_i > 0$. In the case of a unimolecular reaction, the propensity of this reaction is at least $\hat{k}\delta_i n$. In the case of a bimolecular reaction, the propensity is at least $\frac{\hat{k}}{n}(\delta_i n)^2 = \hat{k}\delta_i^2 n$ if $X \neq Y$ or $\frac{\hat{k}}{n}\frac{\delta_i n(\delta_i n - 1)}{2} = \frac{\hat{k}}{2}(\delta_i^2 n - 1)$ if $X = Y$.

The last of these three possibilities is the worst case (i.e., the smallest). Because δ_i is a constant independent of n, for sufficiently large n ($n \geq 2/\delta_i^2$), $\delta_i^2 n - 1 \geq \delta_i^2 n/2$, whence the calculated rate is at least $\frac{\hat{k}}{4}\delta_i^2 n$.

Let $\delta_f = \frac{\hat{k}\delta_i^2}{4}$ and $\delta_r = \frac{\delta_f}{4\lambda}$. Similar to the argument above concerning the maximum rate of decay of any $S \in \Lambda_0$, the rate at which reactions consuming $S \in \Lambda_{i+1}$ occur is at most $\lambda\#S$. Also, by the above arguments, the minimum rate at which reactions producing S proceed is at least $\delta_f n$. Therefore we may lower bound the *net* production of S (i.e., its total count) by a continuous-time random walk on $\mathbb{N}$ that starts at 0, has constant forward rate $\delta_f n$ from every state i to $i + 1$, and has reverse rate λi from i to $i - 1$, defined as the process $\mathbf{W}_{\delta_f, \lambda}^n$ in Section B, Lemma B.7. By this lower bound and Lemma B.7[8] (recall that we have assumed $\lambda \geq 1$),

$$
\begin{aligned}
&\Pr\left[\max_{\hat{t} \in [i, i+1]} \#_{\hat{t}} S < \delta_r n\right] \\
(3.2) \quad &\leq \ \Pr\left[\max_{\hat{t} \in [0,1]} \mathbf{W}_{\delta_f, \lambda}^N(\hat{t}) < \delta_r n\right] \\
(3.3) \quad &< \ 2^{-\delta_f n/22 + 1} \\
(3.4) \quad &= \ 2^{-\hat{k}\delta_i^2 n/88 + 1}.
\end{aligned}
$$

---

[7] The author made up this phrase. It means the following. We show that if the induction hypothesis fails with low probability p, and if the induction step fails with low probability q, given that the induction hypothesis holds, then by the union bound, the induction step fails with probability at most $p + q$.

[8] Although Lemma B.7 is stated for times between 0 and 1, here we use it for times between i and $i + 1$ (just shift all times down by i).

Therefore with high probability $\#S$ reaches count at least $\delta_r n$ at some time $\hat{t} \in [i, i+1]$. Let $\delta_{i+1} = \frac{\delta_r}{c}$. As before, we can bound the decay rate of S from time $\hat{t}$ until time $m+1$ by an exponential decay process with rate λ and initial value $\delta_r n$, proceeding for $m+1$ seconds (since $\hat{t} \geq 0$). By Lemma A.1,

$$\Pr[(\exists t \in [\hat{t}, m+1]) \ \#_t S < \delta_{i+1} n]$$
$$(3.5) \quad \leq \quad \Pr\left[\mathbf{D}_\lambda^{\delta_r n}(m+1) < \frac{\lambda}{c}\delta_r n\right]$$
$$(3.6) \quad < \quad \left(2\frac{1}{c}e^{\lambda(m+1)}\right)^{\lambda\delta_r n/c-1}$$
$$(3.7) \quad = \quad 2^{-\delta_{i+1}n+1}.$$

Fix a particular $S \in \Lambda_{i+1}$. By the union bound applied to (3.4) and (3.7) and the fact that $\hat{t} \leq i+1$, the probability that $\#_t S < \delta_{i+1} n$ at any time $t \in [i+1, m+1]$, given that the induction hypothesis $H(\Lambda_i)$ holds (i.e., given that $\#_t X \geq \delta_i n$ for all $X \in \Lambda_i$ and all $t \in [i, m+1]$), is at most $2^{-\hat{k}\delta_i^2 n/88+1} + 2^{-\delta_{i+1}n+1}$. Define $H(\Lambda_{i+1})$ to be the event that this does not happen for *any* $S \in \Lambda_{i+1}$, i.e., that $(\forall S \in \Lambda_{i+1})(\forall t \in [i, m+1]) \ \#_t S \geq \delta_{i+1} n$. By the union bound over all $S \in \Lambda_{i+1}$,

$$\Pr[\neg H(\Lambda_{i+1})|H(\Lambda_i)]$$
$$(3.8) \quad < \quad |\Lambda_{i+1}|\left(2^{-\hat{k}\delta_i^2 n/88+1} + 2^{-\delta_{i+1}n+1}\right)$$

By the union bound applied to (3.1) and (3.8), and the fact that $\delta_m < \delta_i$ for all $i \in \{0, \ldots, m-1\}$, the probability that any step of the induction fails is at most

$$\Pr[\neg H(\Lambda_0)] + \sum_{i=1}^m \Pr[\neg H(\Lambda_i)|H(\Lambda_{i-1})]$$
$$(3.9) \quad < \quad |\Lambda_0|2^{-\delta_0 n+1}$$
$$+ \sum_{i=1}^m |\Lambda_i|\left(2^{-\hat{k}\delta_{i-1}^2 n/88+1} + 2^{-\delta_i n+1}\right)$$
$$(3.10) \quad < \quad |\Lambda|\left(2^{-\delta_m n+1} + 2^{-\hat{k}\delta_m^2 n/88+1}\right)$$
$$(3.11) \quad < \quad |\Lambda|\left(2^{-\hat{k}\delta_m^2 n/88+2}\right).$$

At this point we are essentially done. The remainder of the proof justifies that the various constants involved can be combined into a single constant ϵ that does not depend on n (although it depends on $\mathcal{C}$ and α) as in the statement of the theorem.

By our choice of δ_{i+1}, we have

$$\delta_{i+1} = \frac{\delta_r}{c} = \frac{\delta_f}{4\lambda c} = \frac{\delta_i^2\hat{k}}{16\lambda c} > \left(\frac{\delta_i\hat{k}}{16\lambda c}\right)^2.$$

Recall that $\delta_0 = \frac{\alpha}{c}$. Therefore, for all $i \in \{1, \ldots, m\}$, $\delta_i > \left(\frac{\alpha\hat{k}}{16\lambda c^2}\right)^{2^i}$. So

$$(3.12) \quad \delta_m > \left(\frac{\alpha\hat{k}}{16\lambda c^2}\right)^{2^m} \geq \left(\frac{\alpha\hat{k}}{16\lambda c^2}\right)^{2^{|\Lambda|-1}}.$$

Therefore, if $n > \left(\frac{\alpha\hat{k}}{16\lambda c^2}\right)^{2^{|\Lambda|-1}}$, then by (3.11) and (3.12), the failure probability in (3.11) is at most
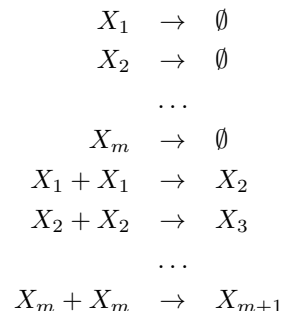
$$|\Lambda|2^{-\hat{k}\delta_m^2 n/88+2} < |\Lambda|2^{-\hat{k}\left(\left(\frac{\alpha\hat{k}}{16\lambda c^2}\right)^{2^{|\Lambda|-1}}\right)^2 n/88+2}$$
$$= |\Lambda|2^{-\hat{k}\left(\frac{\alpha\hat{k}}{16\lambda c^2}\right)^{2^{|\Lambda|}} n/88+2}$$
$$= |\Lambda|2^{-\hat{k}\left(\frac{\alpha\hat{k}}{16\lambda\left(4e^{\lambda(m+1)}\right)^2}\right)^{2^{|\Lambda|}} n/88+2}$$
$$= |\Lambda|2^{-\hat{k}\left(\frac{\alpha\hat{k}}{256\lambda e^{2\lambda(m+1)}}\right)^{2^{|\Lambda|}} n/88+2}$$
$$\leq |\Lambda|2^{-\hat{k}\left(\frac{\alpha\hat{k}}{256\hat{K}\hat{c}e^{2\hat{K}\hat{c}|\Lambda|}}\right)^{2^{|\Lambda|}} n/88+2}.$$

Letting $\epsilon' = \frac{\hat{k}}{88}\left(\frac{\alpha\hat{k}}{256\hat{K}\hat{c}e^{2\hat{K}\hat{c}|\Lambda|}}\right)^{2^{|\Lambda|}}$ implies failure probability at most $|\Lambda|2^{-\epsilon'n+2}$. For $n \geq \frac{2+2\log|\Lambda|}{\epsilon'}$, $|\Lambda|2^{-\epsilon'n+2} \leq 2^{-(\epsilon'/2)n}$. Letting $t = m+1$, $\delta = \delta_m$, and $\epsilon = \epsilon'/2$ completes the proof.

We have chosen $t = m+1$ (i.e., t depends on $\mathcal{C}$); however, the same proof technique works if we choose $t = 1$ (or any other constant independent of $\mathcal{C}$), although this results in smaller values for δ and ϵ.

Although the choice of ϵ is very small, the analysis used many very loose bounds for the sake of simplifying the argument. A more careful analysis would show that a much larger value of ϵ could be chosen, but for our purposes it suffices that ϵ depends on $\mathcal{C}$ and α but not on the volume n.

However, it does seem fundamental to the analysis that $\epsilon \leq \gamma^{2^{|\Lambda|}}$ for some $0 < \gamma < 1$. Consider the CRN with n initial copies of X_1 in volume n and reactions

$$\begin{aligned}
X_1 &\rightarrow \emptyset \\
X_2 &\rightarrow \emptyset \\
&\cdots \\
X_m &\rightarrow \emptyset \\
X_1 + X_1 &\rightarrow X_2 \\
X_2 + X_2 &\rightarrow X_3 \\
&\cdots \\
X_m + X_m &\rightarrow X_{m+1}
\end{aligned}$$

This is a CRN in which the number of stages m is actually equal to its worst-case value $|\Lambda| - 1$, and in which each species is being produced at the minimum rate possible – by a bimolecular reaction – and consumed at the fastest rate possible – by a unimolecular reaction (in addition to a bimolecular reaction). Therefore extremely large values of n are required for the "constant fraction of n" counts in the proof to be large enough to work, i.e., for the production reactions to outrun the consumption reactions. This appears to be confirmed in stochastic simulations as well.

4 Conclusion

The reason we restrict attention to CRNs respecting finite density is that the proof of Theorem 3.1 relies on bounding the rate of consumption of any species by an exponential decay process, i.e., no species S is consumed faster than rate $\lambda \# S$, where $\lambda > 0$ is a constant. This is not true if the CRN is allowed to violate finite density, because a reaction consuming S such as $X + S \to X$ proceeds at rate $\#X\#S/n$ in volume n, and if $\#X$ is an arbitrarily large function of the volume, then for any constant λ, this rate will eventually exceed $\lambda \# S$.
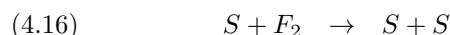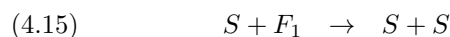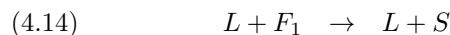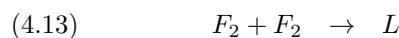
In fact, for the proof to work, the CRN need not respect finite density for all time, but only for some constant time after $t = 0$. This requirement is fulfilled even if the CRN has non-mass-conserving reactions such as $X \to 2X$. Although such a CRN eventually violates finite density, for all constant times $t > 0$, there is a constant $c > 0$ such that $\#_t X \leq c\#_0 X$ with high probability.

The assumption that the CRN respects finite density is a perfectly realistic constraint satisfied by all real chemicals,[9] and as noted, many non-mass-conserving CRNs respect finite density for a constant amount of initial time. Nevertheless, it is interesting to note that there are syntactically correct CRNs that violate even this seemingly mild constraint. Consider the reaction $2X \to 3X$. This reaction has the property that with high probability, $\#X$ goes to infinity in less than 1 second (assuming a sufficiently large rate constant). Therefore it is conceivable that such a reaction could be used to consume a species Y via $X + Y \to X$ at such a fast rate that, although species S is producible via $Y \to S$, with high probability S is never produced, because copies of Y are consumed quickly by X before they can undergo the unimolecular reaction that pro-

duces S. We have not explored this issue further since such CRNs are not physically implementable by chemicals. However, it would be interesting to know whether such a CRN violating finite density could be used to construct a counterexample to Theorem 3.1.

An open problem is to precisely characterize the class of CRNs obeying Theorem 3.1; as noted, those respecting finite density with dense initial configurations are a strict subset of this class.

One may ask whether a dense initial configuration is a realistic assumption for a wet lab,[10] although as we have pointed out, the language of CRNs can be used to describe other distributed computing systems [1]. The strongest possible violation of this assumption is a leader: a species with initial count 1. A weaker violation is a large but sublinear initial count. If one starts in an initial configuration of volume n that has n copies of F_1 and $\frac{\sqrt{n}}{\log n}$ copies of F_2, the following CRN acts as both a timer and a leader election algorithm, producing a single copy of L with high probability in time $O(\log^2 n)$:

$$(4.13) \qquad F_2 + F_2 \quad \to \quad L$$

$$(4.14) \qquad L + F_1 \quad \to \quad L + S$$

$$(4.15) \qquad S + F_1 \quad \to \quad S + S$$

$$(4.16) \qquad S + F_2 \quad \to \quad S + S$$

Reaction (4.13) takes expected time $n \left(\frac{\log n}{\sqrt{n}} \right)^2 = \log^2 n$ (hence is a timer that produces L only after $\log^2 n$ seconds). Reaction (4.14) takes expected time $O(1)$ and initiates an epidemic (see [4] for a discussion of these ideas) through reactions (4.15) and (4.16) to convert all F_1 and F_2 to S in expected time $O(\log n)$, preventing the production of a second copy of L from a second occurrence of reaction (4.13). Hence the CRN is also a leader election algorithm. An interesting question is how large initial counts may be while allowing a timer to be constructed; for example, does $\frac{n}{\log n}$ suffice?

References

---

[9]The reader may notice that recent work has proposed concrete chemical implementations of arbitrary CRNs using DNA strand displacement [6, 25]. However, these implementations assume a large supply of "fuel" DNA complexes supplying mass and energy for reactions such as $X \to 2X$ that violate the laws of conservation of mass and energy.

[10]Theoretical computer science emphasizes asymptotic bounds as the input size $n \to \infty$, whereas in physical science, *in vitro* volumes are generally limited to a milliliter or less. To rephrase the question in language more familiar to physical scientists, "What are the consequences of a CRN when initial concentrations are sufficiently large that bimolecular reactions among initial reagents are no longer a rate-limiting step?" Our proof of the main theorem shows that *every* bimolecular reaction becomes possible almost immediately and is not rate-limiting either.

[1] Dana Angluin, James Aspnes, Zoë Diamadi, Michael Fischer, and René Peralta. Computation in networks of passively mobile finite-state sensors. *Distributed Computing*, 18:235–253, 2006. Preliminary version appeared in PODC 2004.

[2] Dana Angluin, James Aspnes, Zo e Diamadi, Michael J. Fischer, and René Peralta. Urn automata. Technical Report YALEU/DCS/TR-1280, Yale University, November 2003.

[3] Dana Angluin, James Aspnes, and David 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.

[4] Dana Angluin, James Aspnes, and David Eisenstat. Fast computation by population protocols with a leader. *Distributed Computing*, 21(3):183–199, September 2008. Preliminary version appeared in DISC 2006.

[5] James Aspnes and Eric Ruppert. An introduction to population protocols. *Bulletin of the European Association for Theoretical Computer Science*, 93:98–117, 2007.

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

[7] Luca Cardelli and Attila Csikász-Nagy. The cell cycle switch computes approximate majority. *Scientific Reports*, 2, 2012.

[8] Ho-Lin Chen, David Doty, and David Soloveichik. Deterministic function computation with chemical reaction networks. In *DNA 18: Proceedings of The 18th International Meeting on DNA Computing and Molecular Programming*, volume 7433 of *Lecture Notes in Computer Science*, pages 25–42. Springer, 2012.

[9] Ranjit H. Chitale. *Random Processes*. Technical Publications, 2008.

[10] Anne Condon, David Doty, and Chris Thacuk. Unpublished.

[11] Anne Condon, Alan Hu, Ján Maňuch, and Chris Thachuk. Less haste, less waste: On recycling and its limits in strand displacement systems. *Journal of the Royal Society Interface*, 2:512–521, 2012. Preliminary version appeared in DNA 2011.

[12] Anne Condon, Bonnie Kirkpatrick, and Ján Maňuch. Reachability bounds for chemical reaction networks and strand displacement systems. In *DNA 18: 18th International Meeting on DNA Computing and Molecular Programming*, volume 7433, pages 43–57. Springer, 2012.

[13] Matthew Cook, David Soloveichik, Erik Winfree, and Jehoshua Bruck. Programmability of chemical reaction networks. In Anne Condon, David Harel, Joost N. Kok, Arto Salomaa, and Erik Winfree, editors, *Algorithmic Bioprocesses*, pages 543–584. Springer Berlin Heidelberg, 2009.

[14] Massimo Franceschetti, Olivier Dousse, David N. C. Tse, and Patrick Thiran. Closing the gap in the capacity of wireless networks via percolation theory. *IEEE Transactions on Information Theory*, 53(3):1009–1018, 2007.

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

[16] Allen Hjelmfelt, Edward D. Weinberger, and John Ross. Chemical implementation of neural networks and Turing machines. *Proceedings of the National Academy of Sciences*, 88(24):10983–10987, 1991.

[17] Hua Jiang, Marc Riedel, and Keshab Parhi. Digital signal processing with molecular reactions. *IEEE Design and Test of Computers*, 29(3):21–31, 2012.

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

[19] Richard J. Lipton. The reachability problem requires exponential space. Technical report, Yale University, 1976.

[20] Marcelo O. Magnasco. Chemical kinetics is Turing universal. *Physical Review Letters*, 78(6):1190–1193, 1997.

[21] Noam Nisan and Avi Wigderson. Hardness vs randomness. *Journal of Computer and System Sciences*, 49(2):149 – 167, 1994.

[22] Mojżesz Presburger. Uber die vollständigkeit eines gewissen systems der arithmetik ganzer zahlen. In *welchem die Addition als einzige Operation hervortritt. Compte Rendus du I. Congrks des Mathematiciens des pays Slavs, Warsaw*, pages 92–101, 1930.

[23] David Soloveichik. Robust stochastic chemical reaction networks and bounded tau-leaping. *Journal of Computational Biology*, 16(3):501–522, 2009.

[24] David Soloveichik, Matthew Cook, Erik Winfree, and Jehoshua Bruck. Computation with finite stochastic chemical reaction networks. *Natural Computing*, 7(4):615–633, 2008.

[25] David Soloveichik, Georg Seelig, and Erik 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.

[26] Chris Thachuk and Anne Condon. Space and energy efficient computation with DNA strand displacement systems. In *DNA 18: Proceedings of The 18th International Meeting on DNA Computing and Molecular Programming*, pages 135–149, 2012.

[27] Nicholas C. Wormald. Differential equations for random processes and random graphs. *Annals of Applied Probability*, 5(4):1217–1235, 1995.

[28] Gianluigi Zavattaro and Luca Cardelli. Termination problems in chemical kinetics. *CONCUR 2008-Concurrency Theory*, pages 477–491, 2008.

Technical appendix

A   Chernoff bound for exponential decay

In this section, we prove a Chernoff bound on the probability that an exponential decay process reduces to a constant fraction of its initial value after some amount of time.

For all $N \in \mathbb{Z}^+$ and $\lambda > 0$, let $\mathbf{D}_\lambda^N(t)$ be a Markov process on $\{0, 1, \ldots, N\}$ governed by exponential decay, with initial value N and decay constant λ; i.e., $\mathbf{D}_\lambda^N(t) = N$ for all $t \in [0, \mathbf{T}_1)$, $\mathbf{D}_\lambda^N(t) = N-1$ for all $t \in [\mathbf{T}_1, \mathbf{T}_1 + \mathbf{T}_2)$, $\mathbf{D}_\lambda^N(t) = N-2$ for all $t \in [\mathbf{T}_1 + \mathbf{T}_2, \mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3)$, etc., where for $i \in \{1, \ldots, N\}$, $\mathbf{T}_i$ is an exponential random variable with rate $\lambda(N - i + 1)$.

LEMMA A.1. *Let* $N \in \mathbb{Z}^+$, $\lambda, t > 0$, *and* $0 < \delta < 1$. *Then*

$$\Pr[\mathbf{D}_\lambda^N(t) < \delta N] < (2\delta e^{\lambda t})^{\delta N - 1}.$$

Proof. Let ϵ be the largest number such that $\epsilon < \delta$ and $\epsilon N \in \mathbb{N}$. Then $\epsilon N < \delta N \leq \epsilon N + 1$.

Define $\mathbf{T}^\delta = \sum_{i=1}^{N - \epsilon N} \mathbf{T}_i$ as the random variable representing the time required for $N - \epsilon N$ decay events to happen, i.e., for $\mathbf{D}_\lambda^N$ to decay to strictly less than δN from its initial value N.

If $\mathbf{X}$ is an exponential random variable with rate λ', then the moment-generating function $M_\mathbf{X} : \mathbb{R} \to \mathbb{R}$ of $\mathbf{X}$ is $M_\mathbf{X}(\theta) = \mathrm{E}[e^{\theta \mathbf{X}}] = \frac{\lambda'}{\lambda' - \theta}$, defined whenever $|\theta| < \lambda'$ [9]. Therefore

$$M_{\mathbf{T}_i}(\theta) = \frac{\lambda(N - i + 1)}{\lambda(N - i + 1) - \theta} = \frac{N - i + 1}{N - i + 1 - \frac{\theta}{\lambda}},$$

defined whenever $|\theta| < \lambda(N - i + 1)$. In particular, the smallest such value is $\lambda(N - (N - \epsilon N) + 1) = \lambda(\epsilon N + 1)$, so we must choose $|\theta| < \lambda(\epsilon N + 1)$.

Because each $\mathbf{T}_i$ is independent, the moment-generating function of $\mathbf{T}^\delta$ is

$$
\begin{aligned}
M_{\mathbf{T}^\delta}(\theta) &= \mathrm{E}[e^{\theta \mathbf{T}^\delta}] \\
&= \mathrm{E}[e^{\theta \sum_{i=1}^{N - \epsilon N} \mathbf{T}_i}] \\
&= \mathrm{E}\left[\prod_{i=1}^{N - \epsilon N} e^{\theta \mathbf{T}_i}\right] \\
&= \prod_{i=1}^{N - \epsilon N} \mathrm{E}[e^{\theta \mathbf{T}_i}] \\
&= \prod_{i=1}^{N - \epsilon N} \frac{N - i + 1}{N - i + 1 - \frac{\theta}{\lambda}} \\
&= \prod_{i=\epsilon N+1}^{N} \frac{i}{i - \frac{\theta}{\lambda}}.
\end{aligned}
$$

For any $t > 0$, the event that $\mathbf{T}^\delta < t$ (it takes fewer than t seconds for $N - \epsilon N$ decay events to happen) is equivalent to the event that $\mathbf{D}_\lambda^N(t) < \delta N$ ($\mathbf{D}_\lambda^N$, which started at $\mathbf{D}_\lambda^N(0) = N$, has experienced at least $N - \epsilon N$ decay events after t seconds to arrive at $\mathbf{D}_\lambda^N(t) < \delta N$).

Using Markov's inequality, for all $\theta < 0$,

$$
\begin{aligned}
\Pr[\mathbf{D}_\lambda^N(t) < \delta N] &= \Pr[\mathbf{T}^\delta < t] \\
&= \Pr[e^{\theta \mathbf{T}^\delta} > e^{\theta t}] \\
&\leq \frac{\mathrm{E}[e^{\theta \mathbf{T}^\delta}]}{e^{\theta t}} \\
&= \frac{1}{e^{\theta t}} \prod_{i=\epsilon N+1}^{N} \frac{i}{i - \frac{\theta}{\lambda}}.
\end{aligned}
$$

Let $\theta = -\lambda \epsilon N$. This implies $|\theta| = \lambda \epsilon N < \lambda(\epsilon N + 1)$ as required for $M_{\mathbf{T}_i}(\theta)$ to be defined for all $i \in \{1, \ldots, N - \epsilon N\}$.

Then

$$\Pr[\mathbf{D}_\lambda^N(t) < \delta N] \leq e^{\lambda \epsilon N t} \prod_{i=\epsilon N+1}^{N} \frac{i}{i + \epsilon N}$$

$$= e^{\lambda \epsilon N t} \left(\frac{\epsilon N + 1}{\epsilon N + \epsilon N + 1}\right) \left(\frac{\epsilon N + 2}{\epsilon N + \epsilon N + 2}\right)$$

$$\cdots \left(\frac{N - 1}{N + \epsilon N - 1}\right) \left(\frac{N}{N + \epsilon N}\right)$$

$$= e^{\lambda \epsilon N t} \frac{(\epsilon N + 1)(\epsilon N + 2) \ldots (\epsilon N + \epsilon N - 1)(\epsilon N + \epsilon N)}{(N + 1)(N + 2) \ldots (N + \epsilon N - 1)(N + \epsilon N)}$$

$$\text{cancel terms}$$

$$< e^{\lambda \epsilon N t} \frac{(2\epsilon N)^{\epsilon N}}{N^{\epsilon N}} = \left(2\epsilon e^{\lambda t}\right)^{\epsilon N} < \left(2\delta e^{\lambda t}\right)^{\delta N - 1},$$

which completes the proof.

B   Chernoff bounds for biased random walks

This section proves Chernoff bounds for biased random walks that are used in Section 3. Lemma B.6 shows that a random walk on $\mathbb{Z}$ with forward rate $\hat{f}$ and reverse rate $\hat{r}$ has a high probability to take at least $\Omega((\hat{f} - \hat{r})t)$ net forward steps after t seconds. Lemma B.7 uses Lemma B.6 to show that a random walk on $\mathbb{N}$ (i.e., with a reflecting barrier at state 0) with reverse rate in state i proportional to i, has a high probability to reach state j in time t, where j is sufficiently small based on the backward rate and t.
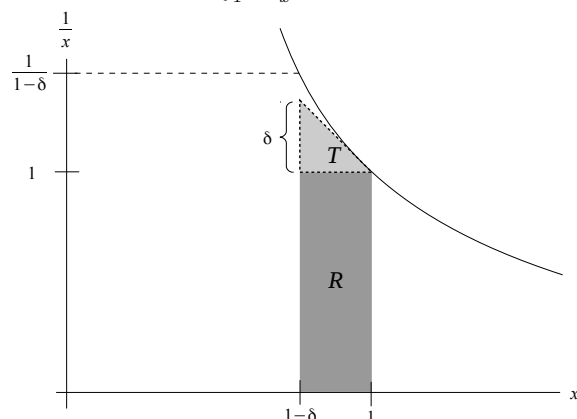
We require the following Chernoff bound on Poisson distributions, due to Franceschetti, Dousse, Tse, and Thiran [14].

THEOREM B.1. ([14]) *Let* $\mathbf{P}(\lambda)$ *be a Poisson random variable with rate* λ. *Then for all* $n \in \mathbb{N}$

$$\Pr[\mathbf{P}(\lambda) \geq n] \leq e^{-\lambda} \left(\frac{e\lambda}{n}\right)^n$$
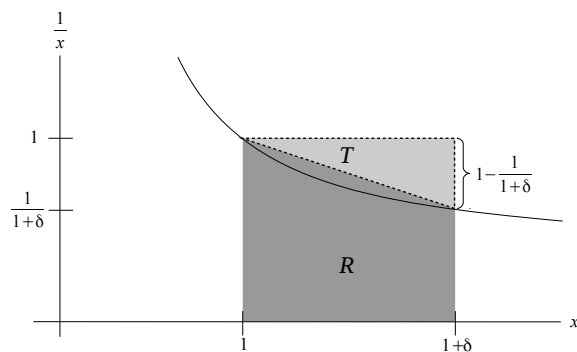
(a) Lemma B.4: $\int_1^{1+\delta} \frac{1}{x}\,dx < \text{area}(R) - \text{area}(T)$.

(b) Lemma B.5: $\int_{1-\delta}^1 \frac{1}{x}\,dx > \text{area}(R) + \text{area}(T)$.

Figure 1: Illustration of Lemmas B.4 and B.5 (not to scale).

if $n > \lambda$ *and*

$$\Pr\left[\mathbf{P}(\lambda) \leq n\right] \leq e^{-\lambda}\left(\frac{e\lambda}{n}\right)^n$$

if $n < \lambda$.

The following corollaries are used in the proof of Lemma B.6.

COROLLARY B.2. *Let* $0 < \gamma < 1$. *Then* $\Pr[\mathbf{P}(\lambda) \leq \gamma\lambda] \leq e^{-\lambda}\left(\frac{e\lambda}{\gamma\lambda}\right)^{\gamma\lambda} = \left(\frac{e^{1-\frac{1}{\gamma}}}{\gamma}\right)^{\gamma\lambda}$.

COROLLARY B.3. *Let* $\gamma > 1$. *Then* $\Pr[\mathbf{P}(\lambda) \geq \gamma\lambda] \leq e^{-\lambda}\left(\frac{e\lambda}{\gamma\lambda}\right)^{\gamma\lambda} = \left(\frac{e^{1-\frac{1}{\gamma}}}{\gamma}\right)^{\gamma\lambda}$.

We require the following bound on the natural logarithm function.

LEMMA B.4. *Let* $\delta > 0$. *Then*

$$\ln(1+\delta) < \frac{\delta}{2} + \frac{\delta}{2(1+\delta)}.$$

Proof. See Figure 1a for an illustration of the geometric intuition. Recall that for $a \geq 1$, $\ln a = \int_1^a \frac{1}{x}\,dx$. Since $\frac{1}{x}$ is convex, the area defined by this integral is at most the area of R, the rectangle of width δ and height 1, minus the area of T, the right triangle of width δ and height $1 - \frac{1}{1+\delta}$. Therefore

$$\begin{aligned}\ln(1+\delta) &= \int_1^{1+\delta} \frac{1}{x}\,dx < \delta - \frac{1}{2}\delta\left(1 - \frac{1}{1+\delta}\right) \\ &= \frac{\delta}{2} + \frac{\delta}{2(1+\delta)}.\end{aligned}$$

A often-useful upper bound is $\ln(1+\delta) < \delta$, i.e., using the area of R as an upper bound. Interestingly, shaving δ down to $\frac{\delta}{2} + \frac{\delta}{2(1+\delta)}$ by subtracting area(T) is crucial for proving Lemma B.6; using the weaker bound $\ln(1+\delta) < \delta$ in the proof of Lemma B.6 gives only the trivial upper bound of 1 for the probability in that proof. The same holds true for the next lemma (i.e., the term $\frac{\delta^2}{2}$ is crucial), which is employed similarly in the proof of Lemma B.6.

LEMMA B.5. *Let* $0 < \delta < 1$. *Then*

$$\ln(1-\delta) < -\delta - \frac{\delta^2}{2}.$$

Proof. See Figure 1b for an illustration of the geometric intuition. Recall that for $0 < a \leq 1$, $-\ln a = \int_a^1 \frac{1}{x}\,dx$. The area defined by this integral is at least the area of R, the rectangle of width δ and height 1, plus the area of T, the right triangle of width δ and height δ. This is because $\frac{d}{dx}\frac{1}{x} = -1$ at the value $x = 1$ and $\frac{d}{dx}\frac{1}{x} < -1$ for all $0 < x < 1$, so the hypotenuse of T touches the curve at $x = 1$ and lies strictly underneath the curve for all $1 - \delta \leq x < 1$. Therefore

$$-\ln(1-\delta) = \int_{1-\delta}^1 \frac{1}{x}\,dx > \delta + \frac{\delta^2}{2}.$$

B.1 Chernoff bound for biased random walk on $\mathbb{Z}$ Let $\hat{f}, \hat{r} > 0$. Let $\mathbf{U}_{\hat{f},\hat{r}}(t)$ be a continuous-time Markov process with state set $\mathbb{Z}$ in which $\mathbf{U}_{\hat{f},\hat{r}}(0) = 0$, with transitions from state i to $i+1$ with rate $\hat{f}$ for all $i \in \mathbb{Z}$, and with transitions from state $i+1$ to i with rate $\hat{r}$ for all $i \in \mathbb{Z}$. In other words, $\mathbf{U}_{\hat{f},\hat{r}}(t)$ is a continuous time biased random walk on $\mathbb{Z}$, with forward bias $\hat{f}$ and reverse bias $\hat{r}$.

LEMMA B.6. *For all* $\hat{f} > \hat{r} > 0$ *and all* $t, \hat{\epsilon} > 0$,

$$\Pr\left[\mathbf{U}_{\hat{f},\hat{r}}(t) < (1-\hat{\epsilon})(\hat{f}-\hat{r})t\right] < 2e^{-\frac{\hat{\epsilon}^2(\hat{f}-\hat{r})^2}{8\hat{f}}t}$$

Proof. Since the forward and reverse rates of $\mathbf{U}_{\hat{f},\hat{r}}$ are constants independent of the state, the total number of forward transitions $\mathbf{F}$ and the total number of reverse transitions $\mathbf{R}$ in the time interval $[0, t]$ are independent random variables, such that $\mathbf{F} - \mathbf{R} = \mathbf{U}_{\hat{f},\hat{r}}(t)$. $\mathbf{F}$ is a Poisson distribution with rate $f = \hat{f}t$ (hence $\mathrm{E}[\mathbf{F}] = f$) and $\mathbf{R}$ is a Poisson distribution with rate $r = \hat{r}t$.

Let $\epsilon = \hat{\epsilon}/2$. Let $d = f - r > 0$. Let $\gamma = \frac{f - \epsilon d}{f} = \frac{(1-\epsilon)f + \epsilon r}{f}$. Let $\lambda = f$. Let $\delta = \frac{\epsilon(f-r)}{(1-\epsilon)f+\epsilon r}$, and note that $1 + \delta = \frac{f}{(1-\epsilon)f+\epsilon r}$. By Corollary B.2,

$$
\begin{aligned}
\Pr[\mathbf{F} \leq \gamma\lambda] &\leq \left(\frac{e^{1-\frac{1}{\gamma}}}{\gamma}\right)^{\gamma\lambda} \\
&= \left(\frac{f\exp\left(1 - \frac{f}{(1-\epsilon)f+\epsilon r}\right)}{(1-\epsilon)f+\epsilon r}\right)^{(1-\epsilon)f+\epsilon r} \\
&= ((1+\delta)\exp(1-(1+\delta)))^{(1-\epsilon)f+\epsilon r} \\
&= \exp(\ln(1+\delta)-\delta))^{(1-\epsilon)f+\epsilon r} \\
&< \exp\left(\frac{\delta}{2} + \frac{\delta}{2(1+\delta)} - \delta\right)^{(1-\epsilon)f+\epsilon r}
\end{aligned}
$$

by Lemma B.4

$$
\begin{aligned}
&= \exp\left(\frac{\delta - \delta(1+\delta)}{2(1+\delta)}\right)^{(1-\epsilon)f+\epsilon r} \\
&= \exp\left(\frac{-\delta^2}{2(1+\delta)}\right)^{(1-\epsilon)f+\epsilon r} \\
&= \exp\left(\frac{-\left(\frac{\epsilon(f-r)}{(1-\epsilon)f+\epsilon r}\right)^2}{2\left(1+\frac{\epsilon(f-r)}{(1-\epsilon)f+\epsilon r}\right)}\right)^{(1-\epsilon)f+\epsilon r} \\
&= \exp\left(\frac{-\left(\frac{\epsilon(f-r)}{(1-\epsilon)f+\epsilon r}\right)^2}{2\left(\frac{(1-\epsilon)f+\epsilon r+\epsilon(f-r)}{(1-\epsilon)f+\epsilon r}\right)}\right)^{(1-\epsilon)f+\epsilon r} \\
&= \exp\left(\frac{-\left(\frac{\epsilon(f-r)}{(1-\epsilon)f+\epsilon r}\right)^2}{\frac{2f}{(1-\epsilon)f+\epsilon r}}\right)^{(1-\epsilon)f+\epsilon r} \\
&= \exp\left(\frac{-(\epsilon(f-r))^2}{2f((1-\epsilon)f+\epsilon r)}\right)^{(1-\epsilon)f+\epsilon r}
\end{aligned}
$$

$$
(\mathrm{B}.1) \qquad = \exp\left(\frac{-(\epsilon(f-r))^2}{2f}\right).
$$

Let $\gamma' = \frac{r+\epsilon d}{r} = \frac{\epsilon f + (1-\epsilon)r}{r}$ and $\lambda' = r$. Let $\delta = \frac{\epsilon(f-r)}{\epsilon f + (1-\epsilon)r}$, and note that $1 - \delta = \frac{r}{\epsilon f + (1-\epsilon)r}$. By

Corollary B.3,

$$
\begin{aligned}
\Pr[\mathbf{R} \geq \gamma'\lambda'] &\leq \left(\frac{e^{1-\frac{1}{\gamma'}}}{\gamma'}\right)^{\gamma'\lambda'} \\
&= \left(\frac{r\exp\left(1 - \frac{r}{\epsilon f + (1-\epsilon)r}\right)}{\epsilon f + (1-\epsilon)r}\right)^{\epsilon f + (1-\epsilon)r} \\
&= ((1-\delta)\exp(1-(1-\delta)))^{\epsilon f + (1-\epsilon)r} \\
&= \exp(\ln(1-\delta)+\delta)^{\epsilon f + (1-\epsilon)r} \\
&< \exp\left(-\delta - \frac{\delta^2}{2} + \delta\right)^{\epsilon f + (1-\epsilon)r}
\end{aligned}
$$

by Lemma B.5

$$
\begin{aligned}
&= \exp\left(\frac{-(\epsilon(f-r))^2}{2(\epsilon f + (1-\epsilon)r)^2}\right)^{\epsilon f + (1-\epsilon)r} \\
&= \exp\left(\frac{-(\epsilon(f-r))^2}{2(\epsilon f + (1-\epsilon)r)}\right) \\
&< \exp\left(\frac{-(\epsilon(f-r))^2}{2(\epsilon f + (1-\epsilon)f)}\right)
\end{aligned}
$$

$$
(\mathrm{B}.2) \qquad = \exp\left(\frac{-(\epsilon(f-r))^2}{2f}\right).
$$

Observe that $\gamma\lambda - \gamma'\lambda' = f - r - 2\epsilon(f - r) = (1 - 2\epsilon)(f - r)$. By (B.1), (B.2), and the union bound, $\Pr[\mathbf{F} - \mathbf{R} \leq (1 - 2\epsilon)(f - r)] < 2 \cdot \exp\left(-\frac{\hat{\epsilon}^2(f-r)^2}{2f}\right)$. Substituting the definitions $f = \hat{f}t$, $r = \hat{r}t$, and $\epsilon = \hat{\epsilon}/2$ completes the proof. $\qquad \blacksquare$

B.2 Chernoff bound for random walk on $\mathbb{N}$ with state-dependent reverse bias Let $N \in \mathbb{Z}^+$, let $\delta_f, \lambda_r > 0$, and let $\mathbf{W}^N_{\delta_f,\lambda_r}(t)$ for $t > 0$ be a continuous-time Markov process with state set $\mathbb{N}$ in which $\mathbf{W}^N_{\delta_f,\lambda_r}(0) = 0$, with transitions from state i to state $i + 1$ with rate $\delta_f N$ for all $i \in \mathbb{N}$, and with transitions from state $i + 1$ to i with rate $\lambda_r i$ for all $i \in \mathbb{N}$. In other words, $\mathbf{W}^N_{\delta_f,\lambda_r}$ is a continuous time random walk on $\mathbb{N}$ with a reflecting barrier at 0, in which the rate of going forward is a constant $\delta_f N$, and the rate of going in reverse from state i is proportional to i (with constant of proportionality λ_r).

LEMMA B.7. *Let $\lambda_r \geq 1$, let $\delta_f, \delta_r > 0$ such that $\delta_r \leq \frac{\delta_f}{4\lambda_r}$. Then for all $N \in \mathbb{Z}^+$ such that $N \geq 6/(\delta_f)$,*

$$
\Pr\left[\max_{\hat{t}\in[0,1]} \mathbf{W}^N_{\delta_f,\lambda_r}(\hat{t}) < \delta_r N\right] < 2^{-\delta_f N/22+1}.
$$

Proof. Consider the event that $(\forall \hat{t} \in [0,1])\, \mathbf{W}^N_{\delta_f,\lambda_r}(\hat{t}) < \delta_r N$. Then in this case, the maximum reverse transition rate is at most $\lambda_r \delta_r N \leq \delta_f N/4$.

For $\hat{t} \in [0,1]$, consider the random walk $\mathbf{U}_{\delta_f N, \delta_f N/4}(\hat{t})$ defined as in Lemma B.6 as a Markov process on $\mathbb{Z}$ (i.e., the states are allowed to go negative) in which the forward rate from any state $i \in \mathbb{Z}$ is $\delta_f N$ as in $\mathbf{W}_{\delta_f, \lambda_r}^N(\hat{t})$, but the reverse rate from any state $i \in \mathbb{Z}$ is $\delta_f N/4$, which is an upper bound on the reverse rate of $\mathbf{W}_{\delta_f, \lambda_r}^N(\hat{t})$ from any state $i \in \{1, \dots, \delta_r N\}$.

Therefore $\Pr[(\forall \hat{t} \in [0,1]) \ \mathbf{W}_{\delta_f, \lambda_r}^N(\hat{t}) < \delta_r N] < \Pr[(\forall \hat{t} \in [0,1]) \ \mathbf{U}_{\delta_f N, \delta_f N/4}(\hat{t}) < \delta_r N]$, since $\mathbf{U}_{\delta_f N, \delta_f N/4}$ has a strictly higher reverse rate and does not have a reflecting barrier at state $i = 0$, hence is strictly less likely never to reach the state $\delta_r N$ at any time $\hat{t} \in [0,1]$.

We prove the theorem by bounding $\Pr[\mathbf{U}_{\delta_f, \delta_f N/4}(1) < \delta_r N]$.

Lemma B.6, with $\hat{\epsilon} = \frac{2}{3}$, $t = 1$, $\hat{f} = \delta_f N$ and $\hat{r} = \delta_f N/4$ implies that

$$
\begin{aligned}
& \Pr\left[\mathbf{U}_{\delta_f N, \delta_f N/4}(1) < \frac{\delta_f N}{4}\right] \\
= \ & \Pr\left[\mathbf{U}_{\delta_f N, \delta_f N/4}(1) < \left(1 - \frac{2}{3}\right)(\delta_f N - \delta_f N/4)\right] \\
< \ & 2e^{-\frac{\left(\frac{2}{3}\right)^2 (\delta_f N - \delta_f N/4)^2}{8\delta_f N}} \qquad \text{by Lemma B.6} \\
= \ & 2e^{-\frac{(3\delta_f/4)^2}{18\delta_f}N} = 2e^{-\delta_f N/32}.
\end{aligned}
$$

Note that $e^{-n/32} = 2^{-n/(32\ln 2)} < 2^{-n/22}$ for all $n > 0$. Therefore,

$$
\Pr\left[\mathbf{U}_{\delta_f N, \delta_f N/4}(1) < \frac{\delta_f N}{4}\right] < 2 \cdot 2^{-\delta_f N/22} = 2^{-\delta_f N/22 + 1}.
$$

Since $\delta_f \geq 4\lambda_r \delta_r$ and $\lambda_r \geq 1$, $\delta_f N/4 \geq \lambda_r \delta_r N \geq \delta_r N$, so

$$
\Pr\left[\mathbf{U}_{\delta_f N, \delta_f N/4}(1) < \delta_r N\right] \leq 2^{-\delta_f N/22 + 1}.
$$

By our observation that $\Pr[(\forall \hat{t} \in [0,1]) \ \mathbf{W}_{\delta_f, \lambda_r}^N(\hat{t}) < \delta_r N] < \Pr[(\forall \hat{t} \in [0,1]) \ \mathbf{U}_{\delta_f N, \delta_f N/4}(\hat{t}) < \delta_r N]$, and since $\mathbf{U}_{\delta_f N, \delta_f N/4}(1) \geq \delta_r N$ is a counterexample to the latter event with $\hat{t} = 1$,

$$
\Pr\left[\max_{\hat{t} \in [0,1]} \mathbf{W}_{\delta_f, \lambda_r}^N(\hat{t}) < \delta_r N\right] \leq 2^{-\delta_f N/22 + 1},
$$

completing the proof.