

Thermochemically consistent realization of a stochastic autocatalytic exporter: chemical completion, inherited operating certificates, cycle affinities and a necessary initiation scale

14 September 2026

Abstract

A quantitative stochastic reaction model is physically informative only if its chemical completion preserves the propensities and the meaning of its output and resource counters. We exhibit such a completion for a maintained autocatalytic exporter whose finite-time operating, output and supply guarantees were proved in a companion paper as statements about a twenty-channel labelled count process. Six reversible chemical pairs on eight species admit positive elemental compositions and a single standard-potential assignment throughout a two-parameter rectangle of paired release and cleavage speeds; at unit fuel and waste activities their mass-action count propensities, stoichiometric updates, export marks and service marks coincide label by label with the certified source. Consequently every inherited estimate transfers without a change of constants: outward interval evaluation of the proved error budget gives success probabilities of at least 0.923849, 0.994201 and 0.999558 for one hundred observation windows at copy scales 10^8 , 2×10^8 and 3×10^8 , while deleting only the templated ligation pair, with every other chemical specification retained, bounds the probability of the same output schedule below 10^{-21699} at the first scale. We then extract three consequences of the completed chemistry. The internal stoichiometric cycles are exactly a passive cycle of zero affinity and one emergent fuel-to-waste cycle of affinity $\log(8 \times 10^{10})$; local detailed balance holds for the count propensities at neighbouring states with the falling-factorial convention; and, because the preparation contains food only, the first template can arise only through two channels, which yields the necessary initiation bound $V \geq (4 \times 10^8 / 401) \log(1/\varepsilon_0)$ for any operating guarantee with failure probability at most ε_0 , about 4.59×10^6 at 99%. Dimensional examples make the eight-hour startup allowance, the gross chemical service and the mixed-species character of the exported product explicit, and a conditional stock-sizing relation delimits the remaining extension to autonomous finite reservoirs. The algebraic and finite marked-kernel statements are verified in Lean 4; the identification with a nonexplosive count process and the initiation argument are conventional proofs, and every printed decimal is an interval-certified evaluation of a proved bound rather than a simulation.

1 Introduction

Structural theories of autocatalysis, from Kauffman’s autocatalytic sets [14] through the RAF formalism [12, 26] to the stoichiometric classification of autocatalytic cores [7], certify that a network can in principle make its own catalysts from a food source, and experimental template replicators [16, 25, 28] and kinetic design principles for minimal autocatalysts [22] show what such mechanisms look like chemically. A different kind of statement is an *operating certificate*: a proved lower bound on the probability that an explicit finite-copy stochastic reactor, prepared with food alone, establishes catalytic material by a deadline, keeps it against washout and

active destruction, and exports a prescribed amount of product in every observation window for a prescribed duration, within stated allowances of food and external driving. A companion paper [2] proves such a certificate for a six-species template reactor with a maintained-drive cleavage competitor, as a theorem about a twenty-channel labelled Markov jump process with mass-action propensities, and sharpens an earlier fixed-horizon result [3].

Such a theorem is stated for a labelled generator with effective rate coefficients. Whether that generator is the count process of *some* balanced chemistry with fixed standard potentials, maintained reservoirs and a consistent identical-reactant convention is a separate question, and it can fail in at least three ways. Reservoir species omitted from the generator can make two effective pairs demand incompatible equilibrium ratios, so that no single potential assignment exists. A different convention for a reaction between identical molecules, $2X \rightleftharpoons Z$ here, changes the propensity and hence the generator. And resource or output counters that are attached only after the dynamics has been projected to concentrations can lose the joint event whose probability is being certified. Open-network thermodynamics [9, 18–20] and the stochastic thermodynamics of reaction networks [23] say what a consistent completion must look like: elemental balance, a potential assignment satisfying local detailed balance, chemostatted species whose maintenance carries the affinity of emergent cycles. The question addressed here is whether the certified exporter admits such a completion *exactly*, with the same propensities, marks and preparation, so that the inherited probability bounds become statements about a specified open chemical model rather than about an abstract source.

Contributions. The answer is affirmative, and the completion has consequences of its own.

- (i) *Realization* (Theorem 3.1). Six reversible pairs on the species $U, W, X, C_1, C_2, Z, F, P$, with F, P maintained fuel and waste, balance three neutral elemental components and admit one standard-potential assignment, independent of the release speed $r \in [19, 21]$ and the cleavage speed $\delta \in [1/50, 1/25]$. At unit reservoir activities their count propensities, updates, export marks and service marks equal the certified source with ligation bias $K = 10$, label by label, including the falling-factorial convention for $2X \rightleftharpoons Z$.
- (ii) *Transfer* (Theorem 4.1, Proposition 4.3). The companion’s finite-duration error budget, its simple exponential bound, its output and service corollaries, and its matched-deletion comparison hold verbatim for the completed model. Deleting only $C_2 \rightleftharpoons Z$ preserves elemental balance, thermochemistry, the competing driven pair, the preparation and the observation protocol, so the comparison is between two consistently specified chemistries meeting the same output requirement.
- (iii) *Cycle structure and count-level local detailed balance* (Section 5). The kernel of the internal stoichiometric matrix is spanned by a passive cycle and a drive cycle; the full matrix kills only the passive one and sends the drive cycle to $e_P - e_F$. The affinity of $ac_p + bc_d$ is $b \log(8 \times 10^{10})$, so the six-pair subnetwork has exactly one independent fuel/waste driving direction. The forward and reverse propensities at neighbouring count states satisfy local detailed balance with an explicit count-level potential.
- (iv) *A necessary initiation scale* (Theorem 6.1). Starting from food only, the first template is created by basal ligation or by the reverse driven channel, with combined coefficient α_δ . A reference immigration–death construction and Jensen’s inequality give $\mathbb{P}(\tau_{\text{seed}} > t) \geq e^{-\alpha_\delta V t}$, hence any guarantee of the operating event with failure at most ε_0 requires $V \geq \log(1/\varepsilon_0)/(500\alpha_\delta) \geq (4 \times 10^8/401) \log(1/\varepsilon_0)$. Sufficient and necessary copy scales are therefore both logarithmic in the confidence, with a factor of about 39 between the exponent coefficients.
- (v) *Physical accounting and scope* (Sections 7–9). A dimensional reading at 1 mM and a one-minute time unit gives sub-picolitre volumes, an 8 h 20 min startup allowance,

whole-run food and driving allowances, a bound of about 0.775 nJ on the chemical work of the fuel/waste pair, and a mixed-species product. A proved activity corridor and conditional stock-sizing relation state precisely what a finite fuel/waste bath would additionally require, and a companion resident completion shows the same bookkeeping principle in a compartment model.

Verification and the three kinds of evidence. Three kinds of statement appear and are kept apart throughout. The algebraic identities (balance, potentials, coefficient binding, source projection, cycle classification, neighbouring-count ratios, initiation algebra, activity bounds) and the finite marked-kernel probability inequalities are verified in Lean 4 [8] with Mathlib [27]; Appendix B maps every such claim to a declaration and records the verification receipts. The identification of the finite stopped construction with the literal infinite-state count process, and the reference-process construction of Theorem 6.1, are conventional proofs written out in full (Appendix A and Section 6). Every printed decimal is an evaluation of a proved bound in outward-rounded interval arithmetic; no trajectory is simulated anywhere in the paper.

What is not claimed. The model is schematic chemistry: three formal neutral elemental components, ideal activities in the sense of [6, 19], and effective reactions, one of which, the reverse driven channel $U + W + P \rightarrow X + F$, is trimolecular and is not asserted to be elementary. No laboratory molecule is identified, no rate is fitted, and the maintained reservoirs are activities, not finite stocks. Varying r and δ changes both directions of a pair together and preserves every equilibrium ratio; nothing is claimed about robustness to independent changes of individual rate constants. The deletion comparison attributes the certified output schedule to the templated cycle within two specified generators; it does not prove that fuel consumption is necessary for export. The necessary initiation bound is necessary, not sufficient, and neither it nor the sufficient rule is claimed optimal in the observation duration. Finally, the probability estimates are inherited: the contribution is their exact chemical realization and its consequences, and the paper credits them accordingly.

Organization. Section 2 specifies the completed chemistry, the environment, the preparation and the observation protocol. Section 3 proves the realization theorem. Section 4 states the inherited guarantee, its consequences and the deletion comparison. Section 5 classifies cycles and proves count-level local detailed balance. Section 6 proves the necessary initiation scale. Section 7 gives the dimensional example, Section 8 the finite-reservoir relation, Section 9 the resident companion, and Section 10 discusses scope and open questions. Appendix A contains the conventional process identification and Appendix B the formal evidence.

2 The completed chemical model

2.1 Species, pairs and coefficients

The internal species, in order, are U, W, X, C_1, C_2, Z : two foods, the covalent product and template $X = UW$, the two substrate-binding complexes $C_1 = X:U$ and $C_2 = X:(U, W)$, and the duplex $Z = X:X$. Two further species, a fuel F and a waste P , are maintained at fixed activities a_F, a_P . Write

$$\varepsilon = \frac{1}{500\,000\,000}, \quad \eta = \frac{1}{8\,000\,000\,000}, \quad (r, \delta) \in [19, 21] \times \left[\frac{1}{50}, \frac{1}{25}\right]. \quad (1)$$

Table 1 lists the six reversible chemical pairs with their normalized mass-action coefficients; concentrations and time are dimensionless until Section 7. For copy scale $V \in \mathbb{N}_{>0}$ and counts

Table 1: The completed reversible chemical subnetwork. The forward orientation fixes all stoichiometric signs. Labels refer to the twenty-channel source of Table 2. The parameters r and δ multiply both directions of their pair, so every ratio k_j^+/k_j^- is fixed.

Pair	Reaction	Labels	k_j^+	k_j^-	k_j^+/k_j^-
0	$U + W \rightleftharpoons X$	0, 1	ε	$\varepsilon/10$	10
1	$X + U \rightleftharpoons C_1$	2, 3	20	20	1
2	$C_1 + W \rightleftharpoons C_2$	4, 5	20	20	1
3	$C_2 \rightleftharpoons Z$	6, 7	20	2	10
4	$Z \rightleftharpoons 2X$	8, 9	r	r	1
5	$X + F \rightleftharpoons U + W + P$	18, 19	δ	$\delta\eta$	8×10^9

$n = (u, w, x, c_1, c_2, z) \in \mathbb{N}^6$ the forward and reverse count propensities of the pairs are

$$\begin{aligned} a^+ &= \left(\varepsilon \frac{uw}{V}, 20 \frac{xu}{V}, 20 \frac{c_1 w}{V}, 20c_2, rz, \delta a_F x \right), \\ a^- &= \left(\frac{\varepsilon x}{10}, 20c_1, 20c_2, 2z, r \frac{x(x-1)}{V}, \delta \eta a_P \frac{uw}{V} \right). \end{aligned} \quad (2)$$

The identical-reactant propensity of $2X \rightarrow Z$ is the falling factorial $x(x-1)/V$ with no additional factor $1/2$; this convention is the one under which the potential assignment of Section 3 closes at the count level (Proposition 5.3). The reverse driven reaction $U + W + P \rightarrow X + F$ has three reactant species; it is an effective channel of the model, and an elementary implementation is not asserted.

2.2 Environment, preparation and labels

Eight further labelled operations describe the environment: separate feeds of U and W , each a Poisson stream of rate V , and washout of each internal species at rate equal to its count. Feed and washout are transfers between the reactor and external inventory; they are not reactions to or from an empty complex, and no thermodynamic reversibility is asserted for them. Together with the twelve chemical directions this gives the twenty labels of Table 2, numbered as in the companion paper so that pair 3 occupies labels 6 and 7 and the two driven directions are the distinct labels 18 and 19. The preparation is always the food-only state $n_0 = (V, V, 0, 0, 0, 0)$. Time is dimensionless with unit dilution rate; V divides every bimolecular propensity and equals each feed rate, so that n/V obeys the classical density scaling [5, 15]. Figure 1 shows the mechanism and its external services.

2.3 Observables and the operating event

Define the two food moieties, the weighted catalytic count and the covalent stock by

$$A = u + x + 2c_1 + 2c_2 + 2z, \quad B = w + x + c_1 + 2c_2 + 2z, \quad Y = x + \frac{9}{8}c_1 + \frac{7}{5}c_2 + \frac{9}{5}z. \quad (3)$$

Every chemical channel, including both driven directions, preserves A and B ; only feeds and washouts change them. Let $E(t)$ be the cumulative export with the marks of Table 2, so that $E/4$ counts *template equivalents* leaving the reactor in free product, in complexes and in duplexes; the marks do not separate free X from exported complexes. Let I_U, I_W count the two feeds and Q_F, Q_P the two driven directions; $Q_D = Q_F + Q_P$ is the *gross* driving service. Fix a post-startup duration $H = m + s$ with $m \in \mathbb{N}$ complete unit windows and a final fraction $0 \leq s < 1$, and put $T = 500 + H$.

Table 2: The twenty-channel labelled source at unit reservoir activities. A mark is the increment of the indicated cumulative counter; every unlisted mark is zero. Channels 6 and 7 are the only channels removed in the deletion comparison of Section 4.2. Channels 18 and 19 change the six counts exactly as channels 1 and 0 but carry distinct service marks and a different coefficient ratio.

Label	Transition	Propensity	Export mark	Service mark (Q_F, Q_P)
0	$U + W \rightarrow X$	$\varepsilon uw/V$		
1	$X \rightarrow U + W$	$\varepsilon x/10$		
2	$X + U \rightarrow C_1$	$20xu/V$		
3	$C_1 \rightarrow X + U$	$20c_1$		
4	$C_1 + W \rightarrow C_2$	$20c_1w/V$		
5	$C_2 \rightarrow C_1 + W$	$20c_2$		
6	$C_2 \rightarrow Z$	$20c_2$		
7	$Z \rightarrow C_2$	$2z$		
8	$Z \rightarrow 2X$	rz		
9	$2X \rightarrow Z$	$rx(x-1)/V$		
10, 11	feed of U ; feed of W	V ; V		
12, 13	washout of U ; of W	u ; w	0; 0	
14, 15	washout of X ; of C_1	x ; c_1	4; 4	
16, 17	washout of C_2 ; of Z	c_2 ; z	4; 8	
18	$X + F \rightarrow U + W + P$	δx		(1, 0)
19	$U + W + P \rightarrow X + F$	$\delta \eta uw/V$		(0, 1)

Definition 2.1 (Operating event). The event $G_{V,H}$ consists of the following simultaneous requirements.

- (a) *Resources*: $A(t), B(t) \in [0.9V, 1.1V]$ for all $t \leq T$.
- (b) *Entry and residence*: Y reaches $V/2500$ by time 500 and never subsequently reaches $Y \leq V/5000$ before T .
- (c) *Output*: on every complete window $[500 + i, 501 + i]$, $0 \leq i < m$, the export increment is at least $V/5000$ marks, i.e. $V/20000$ template equivalents.
- (d) *Supplies*: over $[0, T]$, $I_U, I_W \leq 2VT$ and $Q_D \leq VT/8$.

Window boundaries reset only the measurement counter; the chemical state, the entry/residence phase and all cumulative counters persist. The success event is a joint event of one law, not a conjunction of separately estimated marginals; the identification of the finite marked construction with this event on the literal count process is the subject of Appendix A.

3 Material balance, thermochemistry and the marked-source identity

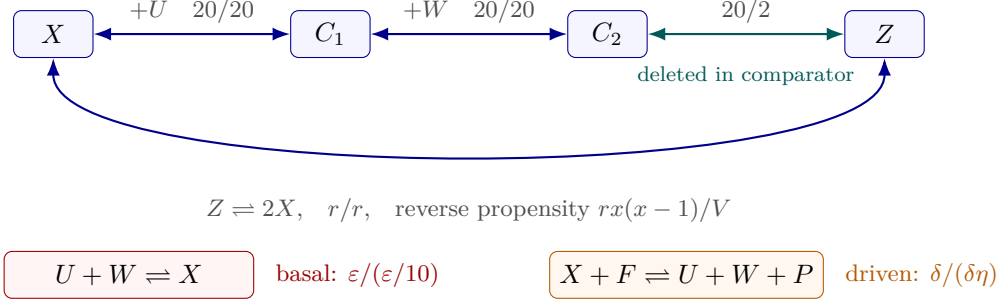
Assign the eight species the composition vectors, in three formal neutral elemental units,

$$U:(1, 0, 0), W:(0, 1, 0), X:(1, 1, 0), C_1:(2, 1, 0), C_2:(2, 2, 0), Z:(2, 2, 0), F:(0, 0, 1), P:(0, 0, 1), \quad (4)$$

and, in units of RT , the standard chemical potentials

$$g = (0, 0, -\log 10, -\log 10, -\log 10, -2\log 10, \log 10 + \log(8 \times 10^9), 0). \quad (5)$$

Let $\nu_j \in \mathbb{Z}^8$ be the product-minus-reactant column of pair j and let $S_{\text{full}} = (\nu_0, \dots, \nu_5)$; let S_{int} be its restriction to the six internal rows.



environment: feeds of U and W at rate V each; washout of all six internal species at unit rate;
 fuel F and waste P maintained at activities $a_F = a_P = 1$, affinity $\log(8 \times 10^{10})$

Figure 1: The completed source and its external services. Arrow annotations give forward/reverse normalized coefficients. The catalytic route $X \rightarrow C_1 \rightarrow C_2 \rightarrow Z \rightarrow 2X$ has net reaction $U + W \rightarrow X$ with the initiating template returned; the basal pair performs the same net reaction without a template. The driven forward reaction consumes fuel while cleaving X ; its reverse creates X from the two foods and the waste. Washout of template-containing species supplies the measured export.

Theorem 3.1 (Realization). *For every $(r, \delta) \in [19, 21] \times [1/50, 1/25]$ the completed chemistry of Section 2 has the following properties.*

- (a) Balance. *Every pair conserves each of the three elemental components of (4), and all charges are zero.*
- (b) Common thermochemistry. *All twelve coefficients are positive and*

$$\log \frac{k_j^+}{k_j^-} = -\nu_j \cdot g \quad (j = 0, \dots, 5), \quad (6)$$

with the single assignment (5), independent of r, δ . The fuel/waste potential difference is $g_F - g_P = \log(8 \times 10^{10})$.

- (c) Coefficient binding. *For every count vector, copy scale and activities a_F, a_P , the propensity of each direction equals its coefficient times the mass-action substrate factor of the displayed complex, with the falling factorial $x(x-1)/V$ for $2X \rightarrow Z$ and the activity factors a_F, a_P multiplying only the driven directions.*
- (d) Marked-source identity. *At $a_F = a_P = 1$ the twenty labelled count propensities, the stoichiometric updates, the export marks and the two service marks coincide label by label with the companion source at ligation bias $K = 10$ and the same (r, δ) .*
- (e) Family. *The rectangle $[19, 21] \times [1/50, 1/25]$ is compact with nonempty interior, contains the rational point $(20, 3/100)$, and is contained in the companion box $K \in [8, 12]$, $r \in [18, 22]$, $\delta \in [1/100, 1/20]$.*

Proof. (a) Each row of Table 1 is checked directly against (4): for instance $C_2 \rightarrow Z$ maps $(2, 2, 0)$ to $(2, 2, 0)$, $Z \rightarrow 2X$ maps $(2, 2, 0)$ to $2(1, 1, 0)$, and $X + F \rightarrow U + W + P$ maps $(1, 1, 0) + (0, 0, 1)$ to $(1, 0, 0) + (0, 1, 0) + (0, 0, 1)$. Reverse directions balance by symmetry.

(b) Positivity is immediate from $r, \delta > 0$. Substituting (5): for pair 0, $-\nu_0 \cdot g = g_U + g_W - g_X = \log 10 = \log(\varepsilon/(\varepsilon/10))$; for pairs 1 and 2, $g_X + g_U - g_{C_1} = 0$ and $g_{C_1} + g_W - g_{C_2} = 0$; for pair 3, $g_{C_2} - g_Z = \log 10 = \log(20/2)$; for pair 4, $g_Z - 2g_X = 0 = \log(r/r)$; for pair 5, $g_X + g_F - g_U - g_W - g_P = -\log 10 + \log 10 + \log(8 \times 10^9) = \log(1/\eta) = \log(\delta/(\delta\eta))$. The release and cleavage speeds cancel in their ratios, so the assignment does not depend on (r, δ) .

(c) is the definition (2) read direction by direction; the point of stating it is that the thermochemistry of (b) is attached to the very coefficients that multiply the count substrate factors, not to an unrelated rate array.

(d) The companion source with parameters $(\varepsilon, k, r, \delta)$ has basal reverse coefficient εk with $k = K^{-1}$, templated reverse coefficient $20k$ for $Z \rightarrow C_2$, duplex association $rx(x-1)/V$, driven coefficients δ and $\delta\eta$ with $\eta = \varepsilon/16$, feeds V , washouts equal to counts, export marks $(0, 0, 4, 4, 4, 8)$ on the six washouts and unit service marks on the two driven labels. At $K = 10$ these are exactly (2) at unit activities and Table 2; the reactant and product complexes of every direction, projected to the internal species, are the companion's arrays, so the natural-count updates agree. Note that the companion's low-level parameter is $k = 1/K$; using $k = 10$ would be a source error.

(e) is arithmetic. The formal statements are `material_balance`, `common_thermochemistry`, `coefficient_binding`, `rate_projection`, `marked_projection` and `rectangle_compact` (Appendix B). $\square$

Remark 3.2 (Paired variation is not rate robustness). Because r and δ scale both directions of their pairs, every point of the rectangle has the same equilibrium ratios and the same potentials. The rectangle is therefore a family of *kinetic* realizations of one thermochemistry. Nothing is claimed for perturbations that change k_j^+ and k_j^- independently; in particular, a change of the driven ratio $\eta a_P/a_F$ by varying reservoir activities is a change of thermochemistry and is treated separately in Section 8.

Remark 3.3 (A productive interior state). The rectangle is not an arbitrary sub-box of the companion's. At the concentration state $x = 10^{-4}$, $c_1 = \frac{3}{5}x$, $c_2 = \frac{3}{10}x$, $z = \frac{6}{25}x$ with unit food concentrations, the four catalytic drift components of the deterministic rate law, including washout and both driven directions, are affine in (r, δ) and bounded below on the whole rectangle by $381100249/(5 \times 10^{13})$, $7/50000$, $9/500000$ and $2421/10^8$, all positive (`productive_positive`). The x -component changes sign at $\delta^* = 581099999/4999993750 \approx 0.11622$ for $r = 19$ and is negative at $\delta = 1/5$ (`productive_boundary_failure`). This instantaneous witness is separate from the food-only preparation and is not an inoculum; it only illustrates that the admitted cleavage interval sits well inside the region where the mechanism is productive at this composition.

4 The inherited operating guarantee and matched deletion

4.1 The finite-duration error budget

Define the constants and error terms of the companion's finite-duration theorem,

$$c_E = \frac{8047}{312\,500\,000\,000}, \quad c_W = \frac{13}{1\,080\,000}, \quad (7)$$

$$\begin{aligned} E_E(V) &= e^{-c_E V} + e^{-3V/500}, \\ E_R(V, T) &= 4e^{-V/1000} + 12000 VT e^{1/50 - V/2000}, \\ E_Y(V, T) &= e^{-V/62500} + 3000 VT e^{18/125 - V/15625}, \\ E_S(V, T) &= 2e^{-(2 \log 2 - 1)VT} + e^{-VT/200}, \end{aligned} \quad (8)$$

$$\mathcal{B}(V, H) = E_E(V) + E_R(V, T) + E_Y(V, T) + m e^{1/8 - c_W V} + E_S(V, T), \quad T = 500 + H, \quad m = \lfloor H \rfloor. \quad (9)$$

The five groups bound, respectively, missed entry, resource exit, loss of residence, failure of some window, and violation of a supply budget. All terms are retained even when numerically negligible.

Theorem 4.1 (Transfer of the operating guarantee). *For every (r, δ) in the rectangle, every positive integer V and every $H \geq 0$, the finite marked law of the completed source satisfies*

$$\mathbb{P}(G_{V,H}) \geq \max\{0, 1 - \mathcal{B}(V, H)\}, \quad (10)$$

and if moreover $V \geq 10^8$ and $H \leq e^{V/10^6}$ then

$$\mathbb{P}(G_{V,H}) \geq 1 - 2e^{-V/(5 \times 10^7)}. \quad (11)$$

Under the conventional identification of Appendix A, both statements hold for the nonexplosive count process with the literal source of Section 2 started from n_0 .

Proof. By Theorem 3.1(d) the completed source at unit activities and the companion source at $K = 10$ are the same labelled marked generator, and by (e) every admitted (r, δ) lies in the companion box. The companion’s finite-duration theorem (Theorem 3.2 there; `competition_timed_supplied_probability`) is a statement about exactly this joint marked law with m complete windows and final fraction s : it bounds the expectation of the indicator of the good event below by $1 - \mathcal{B}(V, m + s)$. Nonnegativity of probabilities gives the maximum in (10). The simpler bound is the companion’s publication endpoint (`competition_publication_probability`), valid for $V \geq 10^8$ and $m + s \leq e^{V/10^6}$. The instantiation on the rectangle is `exporter_endpoint`; no comparison of different marginal laws enters. Appendix A identifies the finite good indicator with $G_{V,H}$ on the literal process. $\square$

Corollary 4.2 (Output, service and sizing). *On $G_{V,H}$ with $m \geq 1$, deterministically for each successful trajectory: the export over the m complete windows is at least $Vm/5000$ marks, i.e. at least $Vm/20000$ template equivalents; every interval $(s', s' + L] \subseteq [500, 500 + m]$ exports at least $(V/5000) \max\{0, \lfloor L \rfloor - 1\}$ marks; and the gross driving service per exported mark is at most $625T/m$, i.e. at most $2500T/m$ driven events per template equivalent. Consequently $\mathbb{E}[E(500 + m) - E(500)] \geq (Vm/5000) \max\{0, 1 - \mathcal{B}(V, H)\}$. For a target failure probability $\varepsilon_0 \in (0, 1)$ and duration $H \geq 1$, the integer rule*

$$V \geq \left\lceil \max\{10^8, 10^6 \log H, 5 \times 10^7 \log(2/\varepsilon_0)\} \right\rceil \quad (12)$$

is sufficient for $\mathbb{P}(G_{V,H}) \geq 1 - \varepsilon_0$.

Proof. The marks and the event are identical to the companion’s, so its Corollaries 7.1 and 7.2 apply verbatim: summing certified window increments gives the total, monotonicity of E gives the aligned-interval bound, and dividing the supply budget $Q_D \leq VT/8$ by the guaranteed output gives the ratios. The sizing rule is (11) with $2e^{-V/(5 \times 10^7)} \leq \varepsilon_0$ and $H \leq e^{V/10^6}$. $\square$

The ratios are quotients of a sufficient upper allowance by a guaranteed lower output, not estimates of turnover efficiency, and the sizing rule is sufficient, not minimal; direct evaluation of (9) certifies smaller sufficient scales (Section 7). Intervals shorter than two windows need not have positive certified output.

4.2 Matched deletion of the templated ligation pair

Delete pair 3, both directions together, and keep everything else: the five other pairs with their coefficients, the driven pair with its competing cleavage, the feeds, washouts and marks, the preparation and Definition 2.1. Every retained pair still balances (4) and still satisfies (6) with the same g (`retained_chemistry`); the driven propensities are unchanged (`disabled_preserves_driving`); and the retained source is exactly the companion’s disabled source, the only zeroed labels being 6 and 7 (`disabled_source_projection`).

Proposition 4.3 (Inherited deletion comparison). *For integer $H \geq 1$, the probability that the disabled model meets the output requirement (c) of Definition 2.1 on all H windows is at most*

$$D(V, H) = e^{-VH/25000} + E_R(V, 500 + H). \quad (13)$$

In particular this bounds its probability of $G_{V,H}$. At $V = 10^8$ and $H = 100$, outward interval evaluation certifies $D(V, H) < 10^{-21699}$.

Proof. Once ligation is disabled, covalent stock is created only by basal ligation (label 0) and by the reverse driven channel (label 19). Along every supported disabled trace from zero template stock, remaining covalent stock plus cumulative export marks is at most four times the number of creation events. Output on all H windows requires total export at least $VH/5000$, hence at least $VH/20000 > VH/21000$ creation events; counting creation during startup and the final fraction only favours the comparator. Stop the disabled law on resource exit alone, impose no entry or residence condition, and cap the creation counter strictly above $VH/21000$. The companion's counted-kernel estimate (Theorem 6.1 there; `competition_timed_disabled_comparison`) bounds the probability of the creation-counter event or a resource exit by (13): the creation intensity on the corridor is at most $\lambda_{\text{off}}V$ with $\lambda_{\text{off}} = 38841/(16 \times 10^{12})$, a unit-tilt Chernoff bound gives the first term, and the resource potential has the same generator as in the enabled model because the removed pair conserves A and B , so crediting every resource exit as a possible success costs at most E_R . The instantiation is `realized_disabled_comparison`. The labelled-clock identification and the counter-saturation argument of Appendix A transfer the bound to the literal disabled process. The decimal claim is a separate interval evaluation of the proved expression (Appendix B). $\square$

Within one consistently specified chemistry, the templated ligation pair therefore changes the certified probability of meeting the same output requirement from at least 0.9238 to at most 10^{-21699} at $V = 10^8$, $H = 100$. The comparison concerns two specified generators; it does not prove that fuel consumption is necessary for export, since the forward driven channel destroys X and both models retain it, and it does not assert that a laboratory intervention could remove one pair without changing others.

5 Cycle structure and count-level local detailed balance

5.1 Exact classification of the internal cycles

Proposition 5.1 (Cycle classification). *Over $\mathbb{R}$,*

$$\ker S_{\text{int}} = \text{span}\{c_p, c_d\}, \quad c_p = (-1, 1, 1, 1, 1, 0), \quad c_d = (1, 0, 0, 0, 0, 1), \quad (14)$$

so $\text{rank } S_{\text{int}} = 4$. Moreover $S_{\text{full}}c_p = 0$, $S_{\text{full}}c_d = e_P - e_F$, $\ker S_{\text{full}} = \text{span}\{c_p\}$ and $\text{rank } S_{\text{full}} = 5$. At unit maintained activities the affinity of $c = ac_p + bc_d$ is

$$\mathcal{A}(c) := \sum_{j=0}^5 c_j \log \frac{k_j^+}{k_j^-} = b \log(8 \times 10^{10}). \quad (15)$$

Proof. For $c = (c_0, \dots, c_5)$ the internal net vector $S_{\text{int}}c$ is

$$(-c_0 - c_1 + c_5, -c_0 - c_2 + c_5, c_0 - c_1 + 2c_4 - c_5, c_1 - c_2, c_2 - c_3, c_3 - c_4).$$

Vanishing of the last three coordinates forces $c_1 = c_2 = c_3 = c_4 =: a$, and the first coordinate then forces $c_0 = -a + c_5$; with $b := c_5$ the remaining coordinates vanish identically, and conversely these substitutions give zero. This is (14). The two additional coordinates of $S_{\text{full}}c$

are $-c_5$ and c_5 , which gives $S_{\text{full}}c_p = 0$, $S_{\text{full}}c_d = e_P - e_F$, the full kernel and the ranks. Finally the six log ratios are $(\log 10, 0, 0, \log 10, 0, \log(8 \times 10^9))$ by Table 1, and their scalar product with $ac_p + bc_d$ is $b(\log 10 + \log(8 \times 10^9))$. Formally: `internal_kernel_classification`, `full_kernel_classification`, `full_cycle_action`, `cycle_force`. $\square$

The passive cycle c_p says that the four templated pairs, traversed forward, perform the basal reaction $U + W \rightarrow X$: it is a closed stoichiometric cycle of the full network and has zero affinity, as it must under (6). The drive cycle c_d , basal ligation followed by driven cleavage, is a cycle of the internal matrix only; in the terminology of [18] it is the single *emergent cycle* created by chemostatting F and P , and its affinity is the maintained potential difference $g_F - g_P$. The six-pair chemical subnetwork therefore has exactly one independent fuel-to-waste driving direction; feeds, collection and any other apparatus services lie outside this classification. A cycle vector describes a linear flux combination and need not itself be an executable finite count trajectory.

Remark 5.2 (Cycle affinity at the count level). By Proposition 5.3 below, the product of forward-over-reverse propensity ratios around any closed loop of count transitions that traverses the pairs with multiplicities c equals $e^{A(c)}$: the count-level terms cancel around a loop and only the reservoir terms survive. Equation (15) is thus the Schnakenberg cycle affinity [24] of the count process, evaluated on the two generators of the cycle space.

5.2 Local detailed balance at neighbouring count states

Chemical consistency of a count process must relate the forward propensity at a state to the reverse propensity at the *neighbouring* state, not at the same state. Let $A_j, B_j \in \mathbb{N}^6$ be the internal reactant and product vectors of pair j , choose a surplus vector $q \in \mathbb{N}^6$, and set $n = q + A_j$, $n' = q + B_j$. With the falling-factorial notation $(n)_a = \prod_i n_i! / (n_i - a_i)!$ the mass-action convention of (2) reads $a_j^+(n) = k_j^+ V^{1-|A_j|} (n)_{A_j}$ and $a_j^-(n') = k_j^- V^{1-|B_j|} (n')_{B_j}$, where $|\cdot|$ is the internal molecularity and the reservoir activities are absorbed into $k_5^\pm$. The shared surplus factorials cancel:

$$\frac{a_j^+(n)}{a_j^-(n')} = \frac{k_j^+}{k_j^-} V^{|B_j|-|A_j|} \prod_{i \in \text{int}} \frac{n_i!}{n'_i!}. \quad (16)$$

Proposition 5.3 (Neighbouring-count ratios). *For every surplus vector $q = (u, w, x, c_1, c_2, z)$, every $V, r, \delta > 0$ and every pair j , the ratio (16) equals k_j^+/k_j^- times the factor in Table 3.*

Proof. Direct substitution of (2) at $n = q + A_j$ and $n' = q + B_j$; for pair 4 the reverse state has $x + 2$ free templates and the falling factorial gives $(x + 2)(x + 1)$. Formally `neighboring_rate_ratio`, which checks the six rational identities for arbitrary counts. $\square$

Corollary 5.4 (Local detailed balance). *Define the count-level potential*

$$\mathcal{G}_V(n) = \sum_{i \in \text{int}} (n_i g_i + \log n_i! - n_i \log V). \quad (17)$$

Then for every supported chemical transition of pair j from n to n' ,

$$\log \frac{a_j^+(n)}{a_j^-(n')} = -[\mathcal{G}_V(n') - \mathcal{G}_V(n)] - \nu_{j,\{F,P\}} \cdot g_{\{F,P\}}, \quad (18)$$

and for general positive activities the reservoir potentials are replaced by $g_i + \log a_i$.

Table 3: The neighbouring-state factors multiplying k_j^+/k_j^- in (16); $q = (u, w, x, c_1, c_2, z)$ is the surplus vector, not the initial state.

Pair	Factor	Pair	Factor
0	$\frac{(u+1)(w+1)}{V(x+1)}$	3	$\frac{c_2+1}{z+1}$
1	$\frac{(u+1)(x+1)}{V(c_1+1)}$	4	$\frac{V(z+1)}{(x+2)(x+1)}$
2	$\frac{(w+1)(c_1+1)}{V(c_2+1)}$	5	$\frac{V(x+1)}{(u+1)(w+1)}$

Proof. Take logarithms in (16) and insert (6). Since $n' - n$ restricted to the internal species is the internal part of ν_j ,

$$\log \frac{k_j^+}{k_j^-} = -\nu_j \cdot g = - \sum_{i \in \text{int}} (n'_i - n_i) g_i - \nu_{j, \{F, P\}} \cdot g_{\{F, P\}},$$

$$\log \left(V^{|B_j| - |A_j|} \prod_{i \in \text{int}} \frac{n_i!}{n'_i!} \right) = - \sum_{i \in \text{int}} [(\log n'_i! - n'_i \log V) - (\log n_i! - n_i \log V)],$$

and the two sums combine into $-\mathcal{G}_V(n') - \mathcal{G}_V(n)$. For the driven pair, a_F and a_P enter $k_5^\pm$ multiplicatively. $\square$

This is local detailed balance for the chemical transitions in the sense of [20, 23], at the level of the actual integer counts rather than of concentrations, and it fixes the identical-reactant convention: with x^2/V in place of $x(x-1)/V$ the factor for pair 4 would not be a ratio of factorials and (18) would fail. No detailed-balance assertion is made for the externally imposed feed or selective collection.

6 A necessary initiation scale

The food-only preparation creates a lower obstruction to high-confidence operation that is independent of the proof technique of the sufficient bound. Before the first template appears, every template-containing species is absent, so the only channels with positive propensity are the two feeds, the two food washouts, basal ligation and the reverse driven channel. The two creating channels have combined propensity

$$a_{\text{seed}}(u, w) = \alpha_\delta \frac{uw}{V}, \quad \alpha_\delta = \frac{1}{500\,000\,000} + \frac{\delta}{8\,000\,000\,000}, \quad (19)$$

(`seed_rate`), and on the rectangle $0 < \alpha_\delta \leq 401/(2 \times 10^{11})$ (`seed_coefficient_bound`). The companion paper noted the resulting Jensen bound as a remark on startup (Remark 2.2 there); here it is turned into a theorem with a complete construction and its consequence for copy-scale sizing.

Theorem 6.1 (No-initiation lower bound). *Let τ_{seed} be the first creation time of X in the literal count process started from $n_0 = (V, V, 0, 0, 0, 0)$. For every $t \geq 0$,*

$$\mathbb{P}(\tau_{\text{seed}} > t) \geq e^{-\alpha_\delta V t}. \quad (20)$$

Consequently, if $\mathbb{P}(G_{V,H}) \geq 1 - \varepsilon_0$ for some $0 < \varepsilon_0 < 1$ and some $H \geq 0$, then

$$V \geq \frac{\log(1/\varepsilon_0)}{500 \alpha_\delta} \geq \frac{400\,000\,000}{401} \log \frac{1}{\varepsilon_0}. \quad (21)$$

Proof. Reference processes. Construct two independent immigration–death processes U^0, W^0 , each started at V , with immigration rate V and unit per-particle death rate: give each initial particle an independent unit-exponential lifetime, let immigrants arrive in an independent rate- V Poisson stream, and give each immigrant an independent unit-exponential lifetime. Counts are bounded by initial counts plus arrivals, so there is no explosion. At a fixed time s , the survivors of the initial particles are $\text{Binomial}(V, e^{-s})$ and the surviving immigrants are $\text{Poisson}(V(1 - e^{-s}))$, independently, so $\mathbb{E}U^0(s) = \mathbb{E}W^0(s) = V$ for all s ; the two processes are independent as processes, hence $\mathbb{E}[U^0(s)W^0(s)] = V^2$. Only the constant mean is used; the deterministic initial state is not a stationary distribution.

Killing. Let E be a unit-exponential variable independent of (U^0, W^0) , and put

$$L_t = \frac{\alpha_\delta}{V} \int_0^t U^0(s)W^0(s) \, ds, \quad \sigma = \inf\{t : L_t \geq E\}.$$

Before σ use the reference paths as the food paths of the actual process; at σ let the seed be basal or driven-reverse creation with probabilities proportional to ε and $\delta\eta$, and continue with the full source. Since every other chemical propensity vanishes before the first creation, this construction has precisely the pre-initiation law of the actual process, and σ has the law of τ_{seed} ; it is the ordinary reaction-clock coupling up to the first seed. Conditional on the entire reference paths, survival through t has probability e^{-L_t} , so

$$\mathbb{P}(\tau_{\text{seed}} > t) = \mathbb{E} e^{-L_t}. \quad (22)$$

Jensen. By Tonelli’s theorem and the reference means, $\mathbb{E}L_t = \alpha_\delta Vt < \infty$. Convexity of $x \mapsto e^{-x}$ gives $\mathbb{E}e^{-L_t} \geq e^{-\mathbb{E}L_t}$, which is (20). Equivalently, integrate the supporting tangent $e^{-x} \geq e^{-\mu}(1 + \mu - x)$ at $\mu = \mathbb{E}L_t$ (`survival_tangent`); the right side is integrable and its expectation is $e^{-\mu}$. No independence of the actual food counts conditional on survival is assumed anywhere: independence is a property of the reference pair, and the bound is transferred through the identity (22).

Necessity. Requirement (b) of Definition 2.1 needs entry by time 500; with no seed, $Y \equiv 0$ and entry is impossible. Hence $\mathbb{P}(G_{V,H}^c) \geq \mathbb{P}(\tau_{\text{seed}} > 500) \geq e^{-500\alpha_\delta V}$. If $\mathbb{P}(G_{V,H}) \geq 1 - \varepsilon_0$ then $e^{-500\alpha_\delta V} \leq \varepsilon_0$, and taking logarithms gives the first inequality of (21) (`necessary_seed_size`); the bound on α_δ gives the second (`uniform_necessary_seed_size`). $\square$

Corollary 6.2 (Necessary and sufficient scales are both logarithmic). *At fixed $H \geq 1$ the sufficient rule (12) and the necessary bound (21) are both of order $\log(1/\varepsilon_0)$. For 99% success the uniform necessary scale is $\log(100) \cdot 4 \times 10^8 / 401 \approx 4\,593\,685.97$, so an integer copy scale must be at least 4 593 686, whereas the simple sufficient rule gives 264 915 869 and the full budget certifies 10^8 (Table 5). The largest necessary exponent coefficient $500\alpha_{1/25} = 1253125/32188 \cdot c_E \approx 38.93 c_E$.*

The factor 38.93 between the exponent coefficients is where the remaining room lies. It is not simply an inaccurate estimate of first creation: the sufficient bound must also pay for establishment of the seeded lineage against washout, reversible sequestration and driven cleavage before the deadline, and the necessary bound charges nothing for that. Neither bound establishes the optimal dependence on the observation duration. Figure 2 displays both.

7 Physical scales, output and whole-run supplies

7.1 Dimensional conversion

Choose a concentration scale c_* (moles per litre), a physical reactor volume Ω (litres) and a time unit τ (seconds), and set

$$V = N_A \Omega c_*, \quad t_{\text{phys}} = \tau t. \quad (23)$$

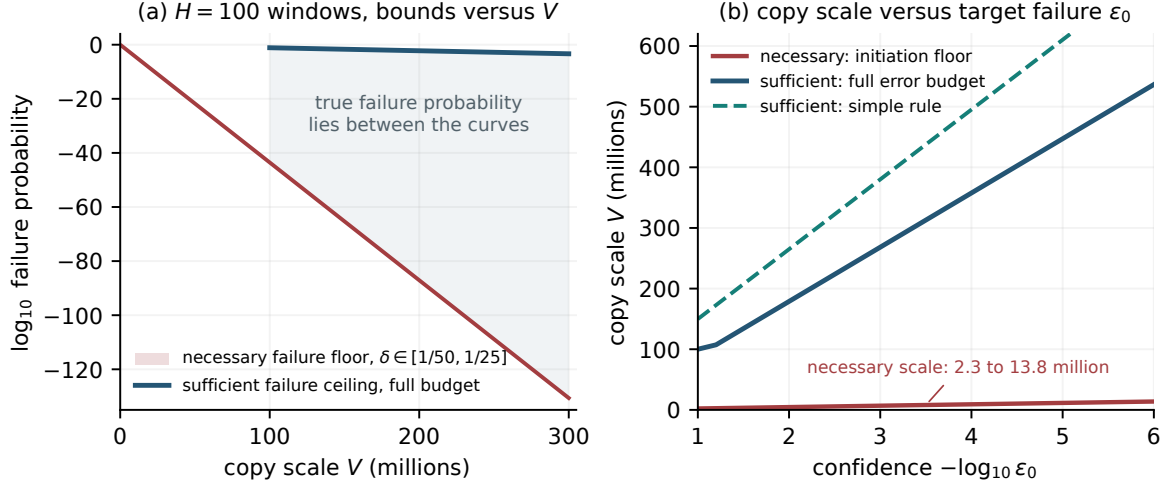


Figure 2: Analytic bounds at $H = 100$. (a) The necessary failure floor $-500\alpha_\delta V / \log 10$ from Theorem 6.1 (shaded band over $\delta \in [1/50, 1/25]$, solid line at $\delta = 1/25$) and the sufficient failure ceiling $\log_{10} \mathcal{B}(V, 100)$ from Theorem 4.1, shown for $V \geq 10^8$. The true failure probability lies between them; the shading is not a confidence band from simulation. (b) Copy scale against target failure probability: the necessary bound (21), the smallest V at which the full budget (9) certifies ϵ_0 , and the simple rule (12). All curves are evaluations of proved formulas.

Table 4: Dimensional coefficients at $c_* = 1$ mM and $\tau = 60$ s. The last row uses the full completed reactant complexes, including the maintained species.

Pair	Forward	Reverse
0	$3.3333 \times 10^{-8} \text{ M}^{-1}\text{s}^{-1}$	$3.3333 \times 10^{-12} \text{ s}^{-1}$
1, 2	$333.333 \text{ M}^{-1}\text{s}^{-1}$	0.333333 s^{-1}
3	0.333333 s^{-1}	0.0333333 s^{-1}
4	$[0.316667, 0.35] \text{ s}^{-1}$	$[316.667, 350] \text{ M}^{-1}\text{s}^{-1}$
5	$[0.333333, 0.666667] \text{ M}^{-1}\text{s}^{-1}$	$[4.16667, 8.33333] \times 10^{-8} \text{ M}^{-2}\text{s}^{-1}$

A reaction of full reactant molecularity κ with normalized coefficient k has dimensional concentration-law coefficient $k/(\tau c_*^{\kappa-1})$; for identical reactants this uses the same falling-factorial convention as the count model, with no additional combinatorial rescaling. Each feed corresponds to a concentration influx c_*/τ and washout to the rate $1/\tau$. Table 4 gives the coefficients at $c_* = 1$ mM and $\tau = 60$ s; at these scales the feed rate is $1.6667 \times 10^{-5} \text{ M s}^{-1}$ for each food and the washout rate is $1/60 \text{ s}^{-1}$. Maintained unit activities correspond to the reservoir species at their standard concentrations in the ideal-activity interpretation. After substituting them, the effective forward driven coefficient is $\delta/60 \in [3.33333, 6.66667] \times 10^{-4} \text{ s}^{-1}$ and the effective reverse coefficient $\delta\eta/(60 \text{ mM}) \in [4.16667, 8.33333] \times 10^{-11} \text{ M}^{-1} \text{ s}^{-1}$. Changing c_* or τ while holding physical constants fixed changes the dimensionless model; these are unit translations, not a calibrated protocol.

7.2 Certified worked examples

Table 5 evaluates Theorem 4.1 at $H = 100$ for three copy scales. Each printed success value is rounded downward and checked by outward interval arithmetic applied to (9): the upper endpoint of the interval for $\log \mathcal{B}$ is compared with the lower endpoint of the logarithm of the claimed failure allowance. The budget is dominated by the entry term $e^{-c_E V}$; at $V = 10^8$ it is $e^{-2.57504}$, i.e. $\log_{10}$ value -1.118326 , and every other term is below 10^{-500} . The failure bound

Table 5: Certified worked examples at $H = 100$ ($T = 600$). Success lower bounds are rounded downward and interval-certified evaluations of (9), not simulated frequencies. Volumes are illustrative, at $c_* = 1$ mM.

V	Volume (pL)	Success at least	Equivalents per window	Disabled, at most
10^8	0.166054	0.923849	5000	10^{-21699}
2×10^8	0.332108	0.994201	10000	10^{-43414}
3×10^8	0.498162	0.999558	15000	10^{-65128}

is 0.076150776706... at 10^8 , 0.005798940793... at 2×10^8 and 0.000441593845... at 3×10^8 .

At $\tau = 60$ s the startup allowance is 500 min (8 h 20 min), followed by 100 one-minute windows; the complete run lasts 10 h. In the first row each window exports at least 5000 template equivalents, which corresponds to 50 nM in the reactor volume, and the 100 windows give at least 5×10^5 equivalents, about 0.830 amol. The exported material is a mixture of X , C_1 , C_2 and Z ; a guarantee for isolated free product would require a recovery or conversion model with its own yields and costs, and until one is supplied the result is a quantified chemical-output benchmark.

7.3 Whole-run supplies and chemical work

On $G_{V,H}$ the whole-run allowance for *each* food is $2VT = 1.2 \times 10^{11}$ molecules at $V = 10^8$, about 199.26 fmol, in addition to the initial stocks V, V . The gross driving allowance is $VT/8 = 7.5 \times 10^9$ events, about 12.45 fmol of total forward-plus-reverse reservoir exchange; net fuel depletion cannot exceed this gross allowance, but the two are not equal in general. Both allowances scale linearly with V across the rows of Table 5. At $H = 100$ the sufficient service-to-output ratio of Corollary 4.2 is 3750 driven events per export mark, i.e. 15000 per template equivalent. Figure 3 collects these counts.

At 298.15 K the fixed fuel/waste potential difference of Theorem 3.1(b) is

$$\Delta\mu = RT \log(8 \times 10^{10}) \simeq 62.235 \text{ kJ mol}^{-1}, \quad (24)$$

and with unit activities the absolute chemical work exchanged through the two driven labels up to time T satisfies, on $G_{V,H}$,

$$|W_{F/P}| = \Delta\mu \frac{|Q_F - Q_P|}{N_A} \leq \Delta\mu \frac{VT}{8N_A} \quad (25)$$

(`chemical_work_bound`), about 0.775 nJ for the first row. This bounds the work of the specified reservoir pair only. Total entropy production, the chemical work of the food feeds, the cost of holding activities fixed, pumping, thermal control, collection and purification are not priced, following the accounting of [20]. The chosen scales show that a copy scale of 10^8 does not mean a macroscopic reactor; they do not establish that molecules with these coefficients and standard free energies exist.

8 What finite reservoirs would require

For varying reservoir activities, pair 5 has effective reverse-to-forward coefficient ratio η_{aP}/a_F (`paired_ratio_changed`). The admitted rectangle varies δ in both directions together and therefore cannot absorb a change in this ratio; a fluctuating bath changes the thermochemistry, not merely the kinetics.

Proposition 8.1 (Activity corridor and conditional stock size). *If $0 < \rho < 1$ and $1 - \rho \leq a_F, a_P \leq 1 + \rho$, then*

$$\eta \frac{1 - \rho}{1 + \rho} \leq \eta \frac{a_P}{a_F} \leq \eta \frac{1 + \rho}{1 - \rho}, \quad \left| \log \frac{a_F}{a_P} \right| \leq \log \frac{1 + \rho}{1 - \rho}. \quad (26)$$

For an ideal fixed-volume bath that initially contains $R > 0$ molecules of each reservoir species at unit activity and whose only changes are the driven exchanges, $|a_F - 1|, |a_P - 1| \leq B/R$ as long as the cumulative gross count is at most B ; hence $R \geq B/\rho$ suffices for the corridor on that event.

Proof. All denominators are positive; lowering a positive numerator and raising its denominator lowers a ratio, which gives the two-sided bound, and applying it to a_F/a_P and its reciprocal before taking logarithms gives the force bound. Each driven event changes each bath count by at most one, so the deviation of either count from R is at most the gross count B , and dividing by R gives the activity deviations. Formally `activity_ratio_bounds`, `activity_force_bound`, `bath_stock_sizing`. $\square$

For illustration, $B = 7.5 \times 10^9$ and $\rho = 0.01$ give $R \geq 7.5 \times 10^{11}$ molecules of each species, about 1.245 pmol, or 1.245 nL of bath at the same 1 mM reference; the force then varies by at most 0.0200 in units of RT . This is a conditional accounting statement. The maintained-model probability of the gross-count event cannot be asserted unchanged for a bath whose activities move, because that motion changes the generator.

The precise remaining implication is an operating estimate uniform over the changed ratio in (26): extend the source with the bath counts, prove the entry, residence, window-output and joint-count generator inequalities throughout the corridor, stop at the first corridor or operating failure, and control the exit with the gross-count estimate of the *extended* law. Using the maintained model's budget to justify its own replacement would be circular. If a uniform estimate fails, the failing entry, residence or output term must be identified. Closed-network limits of chemostatted dynamics have their own literature; in particular Remlein, Esposito

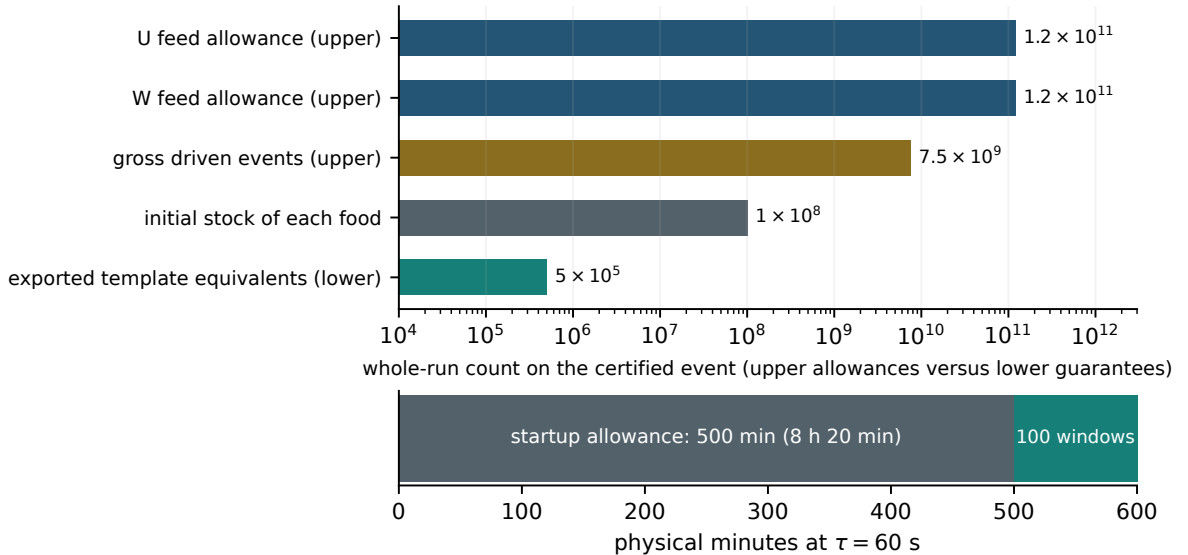


Figure 3: Whole-run accounting at $V = 10^8$, $H = 100$. Upper supply allowances and lower output guarantees are labelled separately; the logarithmic axis displays counts on the certified event, not measured consumption. The time bar shows that every supply allowance includes the startup period.

and Avanzini [21] obtain open-system thermodynamics from closed systems through time-scale and abundance separation, whereas the statement needed here is a finite-horizon, finite-stock marked-output estimate with explicit uniform control, which does not follow from an asymptotic reservoir interpretation.

9 A companion resident completion

The same bookkeeping principle applies to a compartment model used in a companion selection theorem [1], whose thirteen resident channels on (A, B, z, H) are effective open-system channels: the pairs $z \rightleftharpoons H$ and $H \rightleftharpoons 2z$ cannot both conserve a positive resident mass ($m_z = m_H$ and $m_H = 2m_z$ force $m_z = 0$; `Resident.closed_resident_obstruction`). Completing the list with a fuel G , a fuel F , two food reservoirs R_A, R_B and a collected waste W_H gives the reversible pairs



with forward coefficients $(1, 16, 1, 6, 27, 10^{-5})$ and reverse coefficients $(1, 1, 2, 1, 1, 10^{-5})$. The species $A, B, z, H, F, G, R_A, R_B, W_H$ receive, in this order, the positive elemental weights $(2, 1, 1, 2, 1, 3, 2, 1, 2)$ and the standard potentials $(0, 0, 0, -\log 2, 3 \log 2, 0, \log 6, \log 27, 0)$; direct substitution proves material balance and (6) for each pair (`Resident.balanced`, `Resident.thermo`), and unit maintained activities of F, G, R_A, R_B reproduce the original resident propensities exactly, with the same complexes (`Resident.resident_projection`, `Resident.resident_complex_projection`). Selective removal of H is an external transfer to W_H at rate $n_H/10000$; a growth event consumes one precursor Q and one z and retains their material in one added size unit (`Resident.growth_material`). These are separate lemmas about the resident source; importing the module into the exporter's terminal theorem would not make them part of its conclusion, and adding reverse sink or growth reactions would change the generator and is not done.

The completion also makes the maintenance bill visible. With N the newborn size scale and M the population scale, and on the companion's stopped domain where the aggregate resident counts s_z, s_H, s_A, s_B are at most $280NM$ and the aggregate size s_m at most $4NM$, the gross service intensities satisfy

$$16s_z + s_H \leq 4760NM, \quad 6s_m + s_A \leq 304NM, \quad 27s_m + s_B \leq 388NM,$$

and the identical-reactant term obeys $n(n-1)/m \leq 4900N$ for $m \geq N > 0$, $0 \leq n \leq 70N$ (`Resident.supply_bounds`, `Resident.bimolecular_supply`). Integrating over a duration T gives compensator budgets of order NMT ; at the companion's deadline $T = 8/\gamma$ the current sufficient resident-service allowance is $O(NM/\gamma)$, whereas the finite growth-precursor allowance is $O(NM)$. This compares sufficient allowances and is not a necessary $1/\gamma$ consumption law, which would need a lower bound on integrated intensities; it does show that a growth-precursor budget alone does not account for maintained resident chemistry.

10 Discussion

What the completion establishes. The certified exporter is the count process of one balanced chemistry with a single potential assignment, and the correspondence is an identity rather than a fit or a limit. Three consequences follow at once from the completed source. The cycle space is two-dimensional with exactly one emergent driving direction, so the reservoir affinity $\log(8 \times 10^{10})$ is the only thermodynamic force the chemistry needs; count-level local detailed balance fixes the identical-reactant convention; and the food-only preparation forces

a logarithmic necessary copy scale, so that the logarithmic sufficient rules of the companion are of the right order in the confidence, with a factor of about 39 left in the exponent. The deletion comparison becomes a statement about two consistently specified chemistries. And the dimensional reading shows the shape of the certificate: sub-picolitre volumes, an eight-hour startup, supply allowances of order VT against an inventory of order V , and a product that is a mixture of template-containing species.

Where the estimates are loose. The inherited certificate is a chain of Foster–Lyapunov generator inequalities on finite uniformized kernels [11, 13, 17] with explicit constants, and none of its constants is claimed sharp. The gap between the necessary initiation exponent and the sufficient entry exponent is the natural target: a sharper sufficient bound would have to estimate the probability that a seeded lineage establishes before washout under shared resources and reversible sequestration, by a branching or stopped-process comparison, since a parameter sweep of the full reactor cannot close a mathematical gap. Fixed-volume positive recurrence [4] and eventual catalyst loss remain compatible with every statement here; only finite-time operation is the correct object.

Remaining questions. Four extensions are distinct from, and not implied by, the present result. An autonomous fuel/waste bath needs the corridor-uniform estimates of Section 8. Recovery after a new perturbation or dilution protocol needs a new preparation-to-entry estimate, since the certificate is stated for the food-only preparation. Serial transfer and selection over repeated cycles need a specified transfer kernel and propagation of the required distributions and resource conditions; the resident companion of Section 9 supplies the chemistry, not that dynamics. Useful isolated product needs a conversion or recovery model. A stronger formal scope, in which the conventional identification of Appendix A and the reference-clock identity (22) are themselves machine-checked, is possible in principle but was not required for the statements made here.

Data and code availability. The Lean sources, verification receipts, the interval evaluator, the figure script and an independent replay of every printed number are described in Appendix B. The repository location will be supplied with the author information.

A The conventional process identification

This appendix gives the process-level argument invoked in Theorem 4.1 and Proposition 4.3. Its role is to identify an unrestricted labelled count process with the finite marked construction to which the Lean inequalities apply. It is conventional mathematics, in the style of [5], and not an exported Lean theorem about an infinite-state continuous-time Markov process.

A.1 Existence and nonexplosion

Use the twenty nonnegative rates of Table 2 with their stoichiometric jumps. Every positive chemical rate has sufficient reactants, including the $2X$ channel, so the process never leaves $\mathbb{N}^6$. Generate the two feeds with independent rate- V Poisson clocks. All chemical channels preserve A and B and washout decreases them, so pathwise up to any construction time

$$A(t) + B(t) \leq 2V + N_U(t) + N_W(t),$$

where N_U, N_W are the feed counts. Each internal species has positive weight in $A + B$, so a bound on $A + B$ bounds every count. On a finite interval the right side is finite almost surely; on each finite count set all rates are bounded, and the exponential-clock construction [10] has

finitely many jumps in finite time there. Localizing on successive finite resource sets gives a minimal process whose localization times exhaust every bounded interval almost surely: it is nonexplosive. Deleting labels 6 and 7 preserves the argument.

A.2 Stopping and uniformization

On the resource corridor each count is at most $1.1V$. The formal outer count box is larger, and its update agrees with the natural update for every transition out of a corridor state, including the first transition that exits the corridor; the construction retains that exiting state and then absorbs it. Thus no clipping alters a trajectory before the first declared resource failure. The formal total-rate bound on the stopped model is $q = 3000V$. A rate- q Poisson clock with label probabilities $a_j(n)/q$ and the remaining self-loop probability realizes exactly the stopped labelled jump law by thinning; the self-loop is not a reaction and carries no marks. The kernel powers and Poisson mixtures in Lean are the finite-dimensional expectations of this construction [11, 13], which identifies the stopped law on every deterministic time segment.

A.3 Entry, residence and measurement history

A phase variable records whether $Y \geq V/2500$ has been reached and whether a later state has $Y \leq V/5000$; small Y is allowed before entry. First resource exits and residence failures are retained as persistent failures. At time 500 an entry that has not occurred is marked failed, while the successful branch keeps its actual, randomly produced chemical state; no replacement by a fixed catalyst composition occurs. At the times $500 + i$ the measurement counter is compared with the integer threshold $\lceil V/5000 \rceil$, a history bit is conjoined with the comparison, and only the measurement counter is reset; the chemical state, the phase and the cumulative supply counters persist. A saturated export counter has the same threshold event as its natural count, so reaching a cap neither stops nor alters the chemistry. Induction over the finitely many window boundaries identifies the history bit with the conjunction of all completed-window output requirements, and the final fractional interval imposes the resource, residence and supply conditions without an additional complete-window requirement.

A.4 Joint supplies and coupling

The four cumulative counters record the two feeds and the two driven directions. Forgetting the counters commutes with the generators, the uniformized kernels, their powers, the Poisson mixtures and the startup/window/final-fraction composition, so operation and supplies are properties of one joint law. Each finite counter is $\min(C, N_j)$ for a cap C strictly above $2VT$, respectively $VT/8$; if $b < C$ then $\min(C, N_j) \leq b$ exactly when $N_j \leq b$, and for the driven sum a saturated summand already exceeds the budget while unsaturated summands are exact, so no single-counter or combined budget violation can be hidden. Counts are nondecreasing, so a terminal bound controls every earlier prefix. Couple the finite and unrestricted processes on the same reaction clocks up to the first declared failure: source, first exiting state, phase updates, measurements and marks agree throughout, and artificial absorption changes only paths that have already failed some requirement of $G_{V,H}$. Successful paths are unchanged through T , and conversely a good finite terminal record implies none of the failures and hence a successful unrestricted path with the same history. Therefore the finite good-indicator expectation equals $\mathbb{P}(G_{V,H})$, with no independence assumed between entry, residence, window output and supplies.

A.5 The disabled comparator

The disabled finite law is stopped on resource exit alone. Let C_t^{create} count basal and reverse-driven creation. From the initial absence of templates, along every supported disabled trace,

remaining covalent stock plus cumulative export marks is at most $4C_t^{\text{create}}$. Output on $H \in \mathbb{N}_{>0}$ complete windows therefore forces $C_T^{\text{create}} \geq VH/20000 > VH/21000$, and export during startup can only enlarge cumulative export. A creation-counter cap strictly above $VH/21000$ preserves this necessary event; coupling to the finite counted kernel until resource exit and granting every exit as a possible success bounds the output event by the union of the counter event and the resource exit, and the formal counted-kernel inequality bounds that union by (13) on one law.

B Formal evidence and reproducibility

The formal development is in Lean 4 with Mathlib, toolchain version v4.30.0, with a frozen manifest (SHA-256 prefix a8df0b60). The verification bridge compiles with warnings as errors and accepts no `sorry`; a successful receipt has exit code zero and empty compiler output, and the audited declarations depend only on the axioms `propext`, `Classical.choice` and `Quot.sound`. Two receipts support this paper. The realization endpoint `CommonPhysicalRealization.resolution` (source `proofs/CommonPhysicalRealization/Resolution.lean`, SHA-256 85b0f851...) verified in 91.2 s with 98 authenticated dependency modules; it assembles compactness of the rectangle, material balance, common thermochemistry, coefficient binding, the marked-source identity, the productive-state margins and the conservative worked bound $1/3$ at $V = 10^8$, $H = 100$. The publication audit `Publication.lean` (SHA-256 caf8e6f6...) verified in 123.2 s with 103 dependency modules and exports the six declarations `internal_kernel_classification`, `cycle_force`, `neighboring_rate_ratio`, `uniform_necessary_seed_size`, `activity_force_bound` and `realized_disabled_comparison`; it is an audit target, not an umbrella theorem with implicit conclusions. Table 6 maps the claims of the paper to declarations. Unqualified names are in namespace `CommonPhysicalRealization`; companion names are in `RandomViability.Binding`; resident names carry the further namespace `Resident`.

Table 6: Claims and the declarations that support them.
Companion declarations are inherited, not re-proved here.

Claim	Declarations
Theorem 3.1(a), (d)	<code>material_balance</code> , <code>internal_complexes</code> , <code>rate_projection</code> , <code>marked_projection</code> , <code>passive_cycle</code> , <code>driven_cycle</code>
Theorem 3.1(b), (c), (e)	<code>common_thermochemistry</code> , <code>coefficient_binding</code> , <code>finite_fuel_force</code> , <code>family_thermochemistry</code> , <code>rectangle_compact</code> , <code>rectangle_interior_point</code> , <code>donorParameters</code>
Remark 3.3	<code>productive_margin</code> , <code>productive_positive</code> , <code>productive_boundary_failure</code> , <code>family_productive</code>
Theorem 4.1	<code>exporter_endpoint</code> , <code>common_physical_family</code> , <code>worked_run</code> , <code>worked_supplies</code> ; inherited <code>competition_timed_supplied_probability</code> , <code>competition_publication_probability</code>
Corollary 4.2	inherited <code>competition_grid_sum</code> , <code>competition_arbitrary_interval</code> , <code>competition_service_ratio</code> , <code>competition_duration_sizing</code> , <code>competition_confidence_sizing</code>

Claim	Declarations
Proposition 4.3	disabled_source_projection, disabled_preserves_driving, retained_chemistry, realized_disabled_comparison; inherited competition_timed_disabled_comparison
Proposition 5.1	net_expansion, internal_kernel_classification, full_cycle_action, full_kernel_classification, cycle_force
Proposition 5.3	neighboring_rate_ratio
Theorem 6.1, algebra	seed_rate, seed_coefficient_bound, survival_tangent, necessary_seed_size, uniform_necessary_seed_size
Section 7, work bound	chemical_work_bound
Proposition 8.1	activity_ratio_bounds, activity_force_bound, bath_stock_sizing, stock_tolerance, paired_ratio_changed
Section 9	Resident.balanced, Resident.thermo, Resident.closed_resident_obstruction, Resident.resident_projection, Resident.resident_complex_projection, Resident.growth_material, Resident.sink_material, Resident.supply_bounds, Resident.bimolecular_supply

Conventional statements. The following are proved in the text and are not Lean declarations: the process identification of Appendix A; in Theorem 6.1, the reference immigration–death construction, the moment computation, the killing identity (22) and the integration of the tangent bound; Corollary 5.4; Remark 5.2; and the $O(NM/\gamma)$ comparison of Section 9.

Numerical certificates. Every decimal in Tables 5 and 4, Corollary 6.2 and Section 7 is an evaluation of a proved formula. The companion’s evaluator computes the nine components of (9) in logarithmic arithmetic and certifies a rounded success claim by comparing an outward-rounded upper endpoint for $\log \mathcal{B}$ with a lower endpoint for the logarithm of the claimed failure allowance; the disabled decimal is certified in the same way. These calculations are numerical consequences of symbolic bounds; they are neither Lean kernel evidence nor trajectory simulations. An independent script accompanying the manuscript re-derives, in exact rational and symbolic arithmetic, every balance, potential, cycle, neighbouring-ratio and initiation identity, and replays every printed constant of the paper including the interval certifications; the figure script evaluates the same formulas.

Reproduction. From the repository root, with the canonical interpreter and the frozen Lean workspace, the receipts are regenerated by

```
scripts/verify_proof.py proofs/CommonPhysicalRealization/Resolution.lean
  -declaration CommonPhysicalRealization.resolution
scripts/verify_proof.py proofs/CommonPhysicalRealization/Publication.lean
  -declaration CommonPhysicalRealization.internal_kernel_classification
  (and likewise for the five other audited declaration names)
```

and the numerical certificates and figures by the scripts named above. The build of this manuscript changes no proof source or dependency.

References

- [1] Anonymous. Finite-batch selection between inherited chemical states under a shared limiting resource. Companion manuscript, 2026.
- [2] Anonymous. Reliable covalent export from an autocatalytic reactor under driven cleavage: explicit finite-copy operating, output and supply certificates over an exponential horizon. Companion manuscript, 2026.
- [3] Anonymous. Finite-copy productive operation in a reversible autocatalytic binding network: an explicit probability guarantee with a machine-checked proof. Companion manuscript, 2026.
- [4] David F. Anderson and Jinsu Kim. Some network conditions for positive recurrence of stochastically modeled reaction networks. *SIAM Journal on Applied Mathematics*, 78(5): 2692–2713, 2018. doi: 10.1137/17M1161427.
- [5] David F. Anderson and Thomas G. Kurtz. *Stochastic Analysis of Biochemical Systems*, volume 1.2 of *Mathematical Biosciences Institute Lecture Series*. Springer, Cham, 2015. doi: 10.1007/978-3-319-16895-1.
- [6] Francesco Avanzini, Emanuele Penocchio, Riccardo Rao, and Massimiliano Esposito. Nonequilibrium thermodynamics of non-ideal chemical reaction networks. *The Journal of Chemical Physics*, 154(9):094114, 2021. doi: 10.1063/5.0041225.
- [7] Alex Blokhuis, David Lacoste, and Philippe Nghe. Universal motifs and the diversity of autocatalytic systems. *Proceedings of the National Academy of Sciences*, 117(41): 25230–25236, 2020. doi: 10.1073/pnas.2013527117.
- [8] Leonardo de Moura and Sebastian Ullrich. The Lean 4 theorem prover and programming language. In *Automated Deduction – CADE 28*, volume 12699 of *Lecture Notes in Computer Science*, pages 625–635. Springer, 2021. doi: 10.1007/978-3-030-79876-5_37.
- [9] Gianmaria Falasco and Massimiliano Esposito. Macroscopic stochastic thermodynamics. *Reviews of Modern Physics*, 97(1):015002, 2025. doi: 10.1103/RevModPhys.97.015002.
- [10] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977. doi: 10.1021/j100540a008.
- [11] Winfried K. Grassmann. Transient solutions in Markovian queueing systems. *Computers & Operations Research*, 4(1):47–53, 1977. doi: 10.1016/0305-0548(77)90007-7.
- [12] Wim Hordijk and Mike Steel. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227(4):451–461, 2004. doi: 10.1016/j.jtbi.2003.11.020.
- [13] Arne Jensen. Markoff chains as an aid in the study of Markoff processes. *Skandinavisk Aktuarietidskrift*, 36:87–91, 1953. doi: 10.1080/03461238.1953.10419459.
- [14] Stuart A. Kauffman. Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119(1):1–24, 1986. doi: 10.1016/S0022-5193(86)80047-9.
- [15] Thomas G. Kurtz. The relationship between stochastic and deterministic models for chemical reactions. *The Journal of Chemical Physics*, 57(7):2976–2978, 1972. doi: 10.1063/1.1678692.

- [16] Tracey A. Lincoln and Gerald F. Joyce. Self-sustained replication of an RNA enzyme. *Science*, 323(5918):1229–1232, 2009. doi: 10.1126/science.1167856.
- [17] Sean P. Meyn and Richard L. Tweedie. Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25(3): 518–548, 1993. doi: 10.2307/1427522.
- [18] Matteo Polettini and Massimiliano Esposito. Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws. *The Journal of Chemical Physics*, 141(2):024117, 2014. doi: 10.1063/1.4886396.
- [19] Riccardo Rao and Massimiliano Esposito. Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. *Physical Review X*, 6(4): 041064, 2016. doi: 10.1103/PhysRevX.6.041064.
- [20] Riccardo Rao and Massimiliano Esposito. Conservation laws and work fluctuation relations in chemical reaction networks. *The Journal of Chemical Physics*, 149(24):245101, 2018. doi: 10.1063/1.5042253.
- [21] Benedikt Remlein, Massimiliano Esposito, and Francesco Avanzini. Emergence of open chemical reaction network thermodynamics within closed systems, 2026. arXiv:2604.14848.
- [22] Yann Sakref and Olivier Rivoire. Design principles, growth laws, and competition of minimal autocatalysts. *Communications Chemistry*, 7:239, 2024. doi: 10.1038/s42004-024-01250-y.
- [23] Tim Schmiedl and Udo Seifert. Stochastic thermodynamics of chemical reaction networks. *The Journal of Chemical Physics*, 126(4):044101, 2007. doi: 10.1063/1.2428297.
- [24] Jürgen Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Reviews of Modern Physics*, 48(4):571–585, 1976. doi: 10.1103/RevModPhys.48.571.
- [25] Dirk Sievers and Günter von Kiedrowski. Self-replication of complementary nucleotide-based oligomers. *Nature*, 369:221–224, 1994. doi: 10.1038/369221a0.
- [26] Mike Steel, Wim Hordijk, and Joana C. Xavier. Autocatalytic networks in biology: structural theory and algorithms. *Journal of the Royal Society Interface*, 16(151):20180808, 2019. doi: 10.1098/rsif.2018.0808.
- [27] The mathlib Community. The Lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020)*, pages 367–381. ACM, 2020. doi: 10.1145/3372885.3373824.
- [28] Günter von Kiedrowski. A self-replicating hexadeoxynucleotide. *Angewandte Chemie International Edition in English*, 25(10):932–935, 1986. doi: 10.1002/anie.198609322.