From enzyme activity to functional recovery: kinetic guarantees, sharp population bounds and measurement requirements

16 September 2026

Abstract

An enzyme assay reports a rate. A recovery claim concerns a trajectory. A population guarantee concerns how many trajectories complete a task. We treat these as three distinct objects and compute what each observation certifies about the next. For a prescribed, G6PD-inspired kinetic family we show that a finite nonnegative dual certificate on observed rate bands implies literal solutions and a *uniform* arrival deadline for every observation-compatible parameter vector, and we exhibit a parameter sequence along which every member eventually recovers while no common deadline exists. For finite weighted populations that differ only in catalytic capacity, with mean capacity one and support $[5/8, 9/14] \cup [1, 11/8]$, the set of fractions reaching the target by time 120 is exactly $[20/41, 1]$, although every such population has the *same entire* pooled clamped initial-rate response surface; a calibrated classification readout narrows the exact set to $[33/49, 4/5]$, and one three-capacity construction realizes every value in both sets. These claims, including solution existence, are verified in Lean 4, as are the resulting minimax constants, the decision rule, and the weighting conversion. Ordinary analysis then quantifies what the modelling assumptions contribute: removing the support gap replaces the lower endpoint by an unattained infimum near .1164; allowing a fraction η of off-band capacity gives a sharp interpolating infimum; equal mean and equal variance still permit recovery fractions $1/4$ and $3/4$; readout error and cell-weight normalization draw on one exactly computable margin; and the deadline-free limit of the guarantee is $197/407 < 1/2$, approached at the exponential rate $9/8470$ per unit model time with an explicitly identified prefactor. A row-span criterion decides when a further assay can help at all, and a monotone comparison theorem, together with a positive two-state counterexample, delimits transfer of the scalar threshold to multidimensional models. All numerical inputs are designed model values. The results specify measurement requirements; they do not supply a clinically calibrated recovery predictor.

Keywords. partial identification; enzyme kinetics; population heterogeneity; first passage time; sharp bounds; measurement design; formal verification.

1 Introduction

An average activity assay summarizes rates. A recovery question counts cells that complete a task under a stated load and deadline. A population can hold enough high-capacity cells to maintain its mean while many other cells miss that deadline, so the inferentially relevant object is not a point estimate but the full set of recovery fractions compatible with the observations and the kinetic assumptions. Writing $\mathcal{C}(y)$ for the class of models, populations and nuisance parameters compatible with an observation y, and $J(M)$ for the functional outcome of a member M, that object is the

identified set

$$\mathcal{I}(y) = \{J(M) : M \in \mathcal{C}(y)\}. \tag{1}$$

A guarantee must hold on all of $\mathcal{I}(y)$; a bound is sharp when it is attained, or approached, by compatible witnesses; and a further measurement is useful exactly when it cuts the class in a direction that moves the endpoints. This is standard partial-identification reasoning [9, 16, 3], and we claim no novelty for the framework. What we supply is a complete worked chain for one declared kinetic law: from rate observations to actual solutions, from actual solutions to a population identified set with explicit extremal witnesses, and from there to a measurement specification, with the central implications machine-checked.

Glucose-6-phosphate dehydrogenase (G6PD) supplies the motivating vocabulary. G6PD deficiency is common and clinically consequential [8], its activity is measured both in bulk and cell by cell [13, 7], and the expression mosaic in heterozygotes makes population heterogeneity concrete rather than hypothetical. A kinetic study of redox imbalance after oxidative crisis in G6PD-deficient erythrocytes [14] supplies the functional form of the rate law we use. None of these studies calibrates the particular task defined below, and we do not claim that they do. The general warning that population averages can conceal heterogeneity is also long established in cell biology [1]; our contribution is quantitative, not admonitory.

What this paper establishes

Section 2 fixes the rate law, the scalar task, the population class, the pooled observation model and the unit conventions, and separates three meanings of "reserve" that must not be conflated.

Section 3 treats a single cell whose six kinetic coefficients are constrained by rate bands. In reciprocal coordinates the class is a polyhedron, and a finite nonnegative dual certificate on the target state yields a lower bound on the target rate. Theorem 3.2 turns such a certificate into literal solutions of the original rate law and a uniform arrival deadline valid for every compatible coefficient vector. Example 3.3 works out an observation design in exact rational arithmetic; Remark 3.4 shows that pointwise eventual recovery of every class member does not give a common deadline, so a positive uniform margin is not a technical convenience but the content of the statement.

Section 4 contains the principal population result. Theorem 4.1 determines two exact identified sets for the recovery fraction under a mean activity constraint and a designed separated support: $[20/41, 1]$ from bulk information, and $[33/49, 4/5]$ after a calibrated binary readout, with one three-capacity family realizing every value in each. The pooled obstruction is stronger than a coincidence of one assay number: all admitted populations share their entire clamped initial-rate response surface.

Section 5 converts the sets into decision statements: the unavoidable minimax error, the readout precision needed for a designed two-thirds performance requirement, the conversion between weighted and number fractions, and—Proposition 5.3—an exact frontier showing that readout error and weight normalization consume one shared margin.

Section 6 isolates what the modelling assumptions contribute. Proposition 6.1 removes the support gap and shows the lower endpoint becomes an unattained infimum, bracketed by exact rational rectangle sums; Proposition 6.2 gives the deadline dependence and the deadline-free limit $197/407$; Proposition 6.3 identifies the exponential rate and the exact constant governing approach to that limit; and Proposition 6.4 interpolates between the designed gap and connected support through a sharp off-band sensitivity parameter.

Section 7 answers "would another assay help?" with a row-span criterion (Proposition 7.1) and an explicit case where a second, nonlinear, correctly measured assay changes nothing. Section 8 states conditions under which a scalar capacity threshold survives in a multidimensional model, and

a positive two-state counterexample showing that cooperativity in the state alone does not suffice. Section 9 states the resulting measurement requirement, and Section 10 reports evidence scopes, formal declarations and reproduction.

What this paper does not establish

The principal theorem is deliberately small: one varying parameter, fixed remaining kinetic inputs, finite nonnegative weights, no distributional fit, no independence assumption and no numerical ODE solver in its proof. The support, the load, the deadline and the readout calibration are designed mathematical inputs, not estimates from a measured preparation. The scalar state is not a full NADPH balance, and nothing here establishes simultaneous energy service, peroxide control and whole-cell restart; erythrocyte models of that scope exist and have different provenance [11, 5]. Machine verification certifies the stated implications, not the biological adequacy of their hypotheses.

2 Model, task and observation semantics

2.1 Rate law and scalar specialization

Consider the source rate

$$v = \frac{V(N/K_n)(S/K_g)}{1 + (N/K_n)(1 + S/K_g) + H/K_h + A/K_a + B/K_b}, \tag{2}$$

whose functional form we take from the kinetic study [14]: V is a catalytic capacity, N and S are substrate variables, and H, A, B are inhibitory species with affinities K_h, K_a, K_b. All kinetic parameters are positive. For the population results we prescribe

$$K_n = 3, \quad K_g = S = 7, \quad K_h = 56, \quad K_a = 125, \quad K_b = 520, \qquad N = 56 - g,\ H = g,\ A = B = 0, \tag{3}$$

so that a single reduced-carrier coordinate g both supplies the substrate N and inhibits through H. With a constant demand q the task is

$$g' = Vk(g) - q, \qquad k(g) = \frac{56 - g}{115 - (109/56)g}, \qquad q = \frac{3}{10}, \qquad g(0) = 10. \tag{4}$$

The pool is $0 \le g \le 56$, the target is $g = 28$, and the deadline is $T = 120$. Success means $g(t) \ge 28$ for some t in the *closed* interval $[0, 120]$: arrival exactly at the deadline counts.

2.2 Population class and the success functional

A population consists of finitely many capacities V_i with weights w_i,

$$w_i \ge 0, \quad \sum_i w_i = 1, \quad \sum_i w_i V_i = 1, \qquad V_i \in \mathcal{S} := [5/8, 9/14] \cup [1, 11/8], \tag{5}$$

and for any family of actual solutions of (4) on $[0, 120]$ we write

$$p = \sum_i w_i \mathbf{1}\{\exists t \in [0, 120] : g_i(t) \ge 28\}. \tag{6}$$

Real weights permit arbitrary convex mixtures; rational weights become cell counts after clearing denominators, and in the worked interpretation of Section 9 the weights are cell numbers. Mean capacity one is a reference normalization chosen for convenience, not a clinical category of normal activity. Table 1 collects the notation used throughout. Only V varies across the population: the cells are uncoupled and obey a common protocol, so no independence assumption is involved anywhere except in the explicitly labelled sampling illustration of Section 5.

2.3 Pooled clamped observations

Under common clamped assay inputs $z = (N, S, H, A, B)$ the fixed affinities in (2) make the initial rate proportional to capacity, $a(V; z) = V\psi(z)$ with $\psi(z) = a(1; z)$, so the pooled observation is

$$\bar{a}(z) = \sum_i w_i a(V_i; z) = \Big(\sum_i w_i V_i\Big)\psi(z) = \psi(z). \tag{7}$$

This is an identity between functions of z, not the coincidence of one number. Its interpretation is restricted to nonnegative input concentrations and positive fixed kinetic parameters, and it presumes common clamping, additive initial rates and consistent normalization. An intact-cell time course with heterogeneous internal substrates is a different observation model, and (7) says nothing about it.

2.4 Units

The constants in (3)–(4) are model units. For a pool P and a reference rate V_{ref} on the same amount basis, writing $x = g/P$, $\theta = V_{\mathrm{ref}}t/P$, $\nu = V/V_{\mathrm{ref}}$ and $\ell = q/V_{\mathrm{ref}}$ gives the dimensionless form $dx/d\theta = \nu k(Px) - \ell$, with $x_0 = 5/28$, $x_R = 1/2$ and $\theta_T = 15/7$ at reference rate one. Restoring physical units requires a matched pool, rate normalization, kinetic shape, load and time scale. In particular 120 is not asserted to mean 120 minutes.

Remark 2.1 (Three meanings of reserve)**.** A *stock* reserve is a present amount of reduced carrier; a *rate* reserve is the ability to regenerate it relative to demand; a *functional* reserve is the ability to complete a stated task within a stated horizon. Large stock can carry a transient despite slow regeneration, and fast regeneration is useless without an available pool. In this paper g is a stock coordinate, V is a rate parameter, and (6) is the functional quantity. Carrier arrival at the target is not survival, viability, membrane service or low oxidative burden, none of which is modelled here.

Remark 2.2 (Two levels of uncertainty)**.** Section 3 constrains the coefficients of *one* cell; Section 4 constrains a *mixture* of cells at the level of rates. These do not compose naively. Reciprocal linearization is an individual-realization statement, and for a heterogeneous mixture $\bar{v} = \mathbb{E}[1/(\phi \cdot \beta)]$, whereas the reciprocal of the pooled rate is $1/\bar{v}$, which differs from $\phi \cdot \mathbb{E}[\beta]$ in general. Applying an individual reciprocal certificate to a pooled assay as though the pool were a single kinetic realization is therefore invalid; a bridge between the two levels is a substantive additional hypothesis, and we do not assert one.

3 From kinetic observations to individual recovery

3.1 Reciprocal coordinates

Fix $N, S > 0$ and put

$$\beta = \Big(\frac{1}{V}, \frac{K_g}{V}, \frac{K_n K_g}{V}, \frac{K_n K_g}{V K_h}, \frac{K_n K_g}{V K_a}, \frac{K_n K_g}{V K_b}\Big), \qquad \phi = \Big(1, \frac{1}{S}, \frac{1}{NS}, \frac{H}{NS}, \frac{A}{NS}, \frac{B}{NS}\Big). \tag{8}$$

Expanding (2) gives $1/v = \phi \cdot \beta$. The map (8) from positive parameters to positive coefficients is a bijection, with inverse

$$(V, K_g, K_n, K_h, K_a, K_b) = \Big(\frac{1}{\beta_0}, \frac{\beta_1}{\beta_0}, \frac{\beta_2}{\beta_1}, \frac{\beta_2}{\beta_3}, \frac{\beta_2}{\beta_4}, \frac{\beta_2}{\beta_5}\Big). \tag{9}$$

Symbol	Meaning	Role and normalization
V	catalytic capacity	the only quantity varying across the population; mean one by (5)
g	reduced-carrier coordinate	scalar state in $[0, 56]$; stock, not a whole-cell state
q	demand	constant load $3/10$ in model units
T	deadline	120 in model units unless stated; closed interval
p	recovery fraction	weighted mass of cells with an actual hit of 28 by T
s, f	readout sensitivity, false-positive rate	positive-readout fractions *within* the successful and failing classes for the task (6)
r	observed positive fraction	$r = sp + f(1-p)$, no independence assumed
R_h	cell-weight ratio	largest over smallest positive cell weight, used to convert weighted to number fractions
β, ϕ	reciprocal coefficients, features	$1/v = \phi \cdot \beta$ for one cell, Section 3
c_T	capacity cutoff at deadline T	connected support only; V_∞ is the deadline-free threshold
η	off-band mass	weight violating the designed support gap, Proposition 6.4

Table 1: Notation. Sensitivity and specificity here refer to the functional task (6), not to any clinical G6PD deficiency label.

Consequently a band $v \in [v_-, v_+]$ with $v_- > 0$ observed at inputs (N, S, H, A, B) is exactly the pair of linear inequalities $1/v_+ \le \phi \cdot \beta \le 1/v_-$, and an observation class described by finitely many such bands is a polyhedron in β. Working in β therefore replaces a nonlinear parameter-estimation problem by linear programming duality, without leaving the original kinetic family: every positive β is the coefficient vector of an actual parameter vector through (9).

With $N = P - g$, $H = g$ and S, A, B fixed, (8) collapses to

$$v(g) = \frac{P - g}{a(P - g) + bg + c}, \qquad a = \beta_0 + \frac{\beta_1}{S}, \quad b = \frac{\beta_3}{S}, \quad c = \frac{\beta_2 + \beta_4 A + \beta_5 B}{S}, \tag{10}$$

with $a, b, c > 0$, and the target-state reciprocal rate is $1/v(R) = \phi_R \cdot \beta$ for the feature vector ϕ_R of (8) evaluated at $N = P - R$, $H = R$.

Lemma 3.1 (Exact secant identity)**.** *For $a, b, c > 0$ and $0 \le x < y \le P$, with $D(g) = a(P - g) + bg + c > 0$,*

$$v(x) - v(y) = \frac{(bP + c)(y - x)}{D(x)D(y)} > 0. \tag{11}$$

Proof. $D > 0$ on $[0, P]$ because it is affine with positive values $aP + c$ and $bP + c$ at the endpoints. Clearing denominators, $(P - x)D(y) - (P - y)D(x) = b\,[y(P - x) - x(P - y)] + c(y - x) = (bP + c)(y - x)$. □

So v is strictly decreasing: the state both feeds and inhibits, and inhibition wins uniformly. Two consequences are used repeatedly. First, the drift $v - q$ is smallest at the target among all states below it, so a bound at the target bounds the whole approach. Second, arrival is decided by the sign of the target drift, with no need for a long numerical run.

3.2 A certificate that implies a uniform deadline

Theorem 3.2 (Certified individual recovery)**.** *Let $P > 0$, $S > 0$, $A, B \ge 0$, $q > 0$ and $0 \le x < R < P$. Let $F_i \in \mathbb{R}^6$ and $b_i \in \mathbb{R}$, i in a finite index set, describe an observation class*

$\{\beta > 0 : F_i \cdot \beta \le b_i \ \forall i\}$. *Suppose there are multipliers* $y_i \ge 0$ *and a number* $U > 0$ *with*

$$\sum_i y_i F_i = \phi_R, \qquad \sum_i y_i b_i \le U, \qquad q < \frac{1}{U}. \tag{12}$$

Then for every *positive* β *in the class the initial value problem* $g' = v(g) - q$, $g(0) = x$, *with* v *the literal rate* (2) *at the parameters* (9) *and inputs* $N = P - g$, $H = g$, *has a solution that remains in* $[0, P]$ *and satisfies*

$$g(t^\star) \ge R \quad \textit{for some } t^\star \le \frac{R - x}{1/U - q}. \tag{13}$$

Proof. Fix a compatible β and let a, b, c be as in (10); they are positive. Nonnegative combination of the observed inequalities gives $\phi_R \cdot \beta = \sum_i y_i F_i \cdot \beta \le \sum_i y_i b_i \le U$, that is $v(R) \ge 1/U$, hence the target drift is at least $m := 1/U - q > 0$. By Lemma 3.1, $v(g) \ge v(R)$ for all $g \le R$, so

$$v(g) - q \ge m > 0 \qquad (0 \le g \le R). \tag{14}$$

In particular $v(0) \ge q$, so the field points inward at $g = 0$; at $g = P$ it equals $-q < 0$. The field is locally Lipschitz on $[0, P]$ because D is bounded away from zero there. Local existence, uniqueness and continuation, together with forward invariance of $[0, P]$, give a solution on every finite horizon; equivalently, extend the field to a globally bounded Lipschitz field agreeing with it on $[0, P]$, solve, and observe that the solution never leaves $[0, P]$, so it solves the original problem. Suppose $g(t) < R$ for all $t \in [0, \bar{t}\,]$ with $\bar{t} = (R - x)/m$. Then (14) applies throughout, and integrating gives $g(\bar{t}\,) \ge x + m\bar{t} = R$, a contradiction. Hence (13) holds. Finally, the rate used in the argument is the literal (2) at the parameters (9), including at the pool boundary, by the reciprocal expansion of Section 3. □

Theorem 3.2 is the composition that matters for experiment design: the observations are constraints, the multipliers are a proof, and the output is a deadline. A numerical optimizer may propose y; exact arithmetic then checks (12), and no property of the optimizer is used afterwards. A feasible positive β certifies in addition that the class is nonempty, which is what makes the conclusion non-vacuous.

Example 3.3 (An observation design)**.** Take $P = 56$, $q = 3/10$, $x = 10$, $R = 28$, and a class in which $a = 2$, $c = 3$ are pinned and only the inhibition coefficient varies, $b = 3/K$ with $K > 0$ unknown. Then $1/v(28) = (59 + 84/K)/28$, so the exact decision boundary is

$$v(28) > q \iff K > \frac{252}{103}.$$

Suppose the available inhibitory observation is the designed functional $y(K) = K/(3K + 56)$, increasing in K, so that the decision boundary reads $y > 9/233$. An observed band $y \in [.24, .26]$ is the interval $K \in [48, 728/11]$ and yields the exact conclusions in Table 2: a worst-case target reciprocal rate $243/112$, a guaranteed target drift $391/2430$, and the uniform deadline $43740/391 \approx 111.8670$ for every compatible K in the band.

Two features of Example 3.3 deserve emphasis. First, without an observation that constrains the inhibition direction, the class contains β with arbitrarily large b, and the worst-case target rate tends to zero: no amount of extra precision in the *remaining* coordinates repairs this, because those coordinates do not appear in the missing direction. This is structural rather than statistical non-identifiability, and it is visible before any data are seen; profile-likelihood diagnostics address the

Quantity	Value
Decision boundary in the unknown coefficient	$K > 252/103$
Equivalent boundary in the observed functional	$y > 9/233$
Designed observation band	$y \in [.24, .26]$
Worst-case target reciprocal rate on the band	$243/112$
Guaranteed target rate	$112/243$
Guaranteed target drift	$391/2430$
Uniform arrival deadline	$43740/391 \approx 111.8670$
A compatible realization	$K = 56$, $y = 1/4$
First passage time of that realization	≈ 103.8058 (numerical)

Table 2: Exact consequences of the designed band in Example 3.3. The band is an illustrative design, not a same-preparation measurement. Only the last row is numerical.

practical variant of the same question for fitted models [12]. The useful experimental question is which observation sees the direction that the task depends on. Second, the compatible realization $K = 56$ is not an incidental choice: it gives $b = 3/56$, hence $v(g) = (56-g)/(2(56-g)+(3/56)g+3) = k(g)$, which is exactly the capacity-one cell of (4). The individual and population parts of this paper therefore concern the same scalar law, approached from two different observation models.

Remark 3.4 (Eventual recovery without a uniform deadline)**.** Let $K \downarrow 252/103$ within the class of Example 3.3. Every member has positive target drift, hence recovers in finite time by Theorem 3.2 applied to its own singleton class, but the drift margin tends to zero and the arrival times diverge. Pointwise eventual recovery of every member of a noncompact class therefore does not produce one finite deadline for the class. A uniform statement needs a positive uniform margin, as in (12), or a separate uniform passage-time argument.

4 The sharp population theorem

Theorem 4.1 (Exact bulk and repaired identified sets)**.** *Every population satisfying* (5) *admits solutions of* (4) *on* $[0, 120]$ *that remain in* $[0, 56]$*. For every such family of actual solutions:*

1. *its recovery fraction* (6) *satisfies* $20/41 \le p \le 1$*;*
2. *its entire pooled clamped response surface is* $\bar{a}(z) = \psi(z)$*;*
3. *if* $r = sp + f(1-p)$ *and*

$$\frac{9}{10} \le s \le 1, \qquad 0 \le f \le \frac{1}{50}, \qquad \frac{17}{25} \le r \le \frac{18}{25}, \tag{15}$$

then $33/49 \le p \le 4/5$*.*

Conversely, every value in $[20/41, 1]$ *is realized by an admitted population together with actual solutions, and every value in* $[33/49, 4/5]$ *is realized in the same class with readout parameters satisfying* (15)*. At most three positive-weight capacities are needed for either family.*

The readout quantities are conditional rates for the actual task (6): s is the positive-readout fraction among successful cells, f the positive-readout fraction among failing cells, and r the overall positive fraction. The mixture identity $r = sp + f(1-p)$ is a decomposition of a total mass, not an independence assumption; if a class has zero mass one works with joint masses and no conditional ratio. The calibration bounds (15) are designed inputs.

4.1 Existence and classification

The denominator of k is affine and positive on $[0, 56]$, with value 115 at $g = 0$ and 6 at $g = 56$. For every admitted capacity,

$$Vk(0) - q \geq \frac{5}{8} \cdot \frac{56}{115} - \frac{3}{10} = \frac{1}{230} > 0, \qquad Vk(56) - q = -\frac{3}{10} < 0,$$

so the field points inward at both ends of the pool, and the argument of Theorem 3.2 gives solutions on $[0, 120]$ staying in $[0, 56]$, unique by local Lipschitz continuity. Moreover

$$k'(g) = -\frac{6}{(115 - (109/56)g)^2} < 0, \qquad Vk(28) = \frac{56V}{121}, \tag{16}$$

which is Lemma 3.1 in the present normalization.

A low-band cell, $V \leq 9/14$, has target drift at most $\frac{9}{14} \cdot \frac{56}{121} - \frac{3}{10} = -\frac{3}{1210} < 0$. It cannot cross the target: at a first hit from below a differentiable trajectory would have nonnegative left derivative, contradicting the strictly negative prescribed derivative there. A high-band cell, $V \geq 1$, has drift at least $\frac{56}{121} - \frac{3}{10} = \frac{197}{1210}$ while below the target by (16), so if no hit had occurred by

$$T_0 = \frac{28 - 10}{197/1210} = \frac{21780}{197} \approx 110.5584 < 120, \tag{17}$$

integrating the drift bound would force $g(T_0) \geq 28$, a contradiction. Hence for every admitted capacity and every admitted trajectory

$$\exists t \in [0, 120] : g(t) \geq 28 \quad \Longleftrightarrow \quad V \geq 1. \tag{18}$$

Success is therefore membership in the upper band, for all solutions simultaneously; the designed support gap is what makes the classification this clean, and Section 6 quantifies its role.

4.2 Averaging, compensation and sharpness

Let $I = \mathbf{1}\{V \geq 1\}$. On the support $\mathcal{S}$ the pointwise bound

$$V \leq \frac{9}{14} + \frac{41}{56} I \tag{19}$$

holds, with equality exactly at the two capacities $9/14$ and $11/8$. Taking expectations and using $\mathbb{E}[V] = 1$ gives $1 \leq \frac{9}{14} + \frac{41}{56} p$, that is $p \geq 20/41$; normalization gives $p \leq 1$; and (7) gives the response-surface identity. The mechanism is a compensation budget: a population can hide failing cells only if its successful cells make up the capacity deficit, and the largest admitted successful capacity caps how much compensation a unit of successful mass can supply.

That reading dictates the extremal construction. To make p as small as possible, failing mass should be as heavy as failing mass may be, hence at the largest failing capacity $9/14$; and the compensating successful mass should be as efficient as possible, hence at the largest successful capacity $11/8$. Two atoms already give the endpoint, with weights $21/41$ at $9/14$ and $20/41$ at $11/8$. A third atom exactly at the mean, $V = 1$, is neutral for the mean constraint and successful for the task, so any amount of it can be substituted in to raise p continuously to one. Concretely, for $p \in [20/41, 1]$ put

$$\Big(1 - p,\ \frac{20(1-p)}{21},\ \frac{41p - 20}{21}\Big) \quad \text{at capacities} \quad \Big(\frac{9}{14}, \frac{11}{8}, 1\Big). \tag{20}$$

The weights are nonnegative on the whole interval, they sum to one, the mean capacity is one, and the high-band mass is p. By (18) and the existence statement these are dynamical witnesses, not distributions to which success labels have been assigned by fiat. A count realization at the lower endpoint is 42 cells at $9/14$ and 40 at $11/8$.

4.3 Measurement repair and joint feasibility

For $0 \le p \le 1$ the calibration bounds (15) give $\frac{9}{10}p \le r \le \frac{1}{50} + \frac{49}{50}p$, which with $17/25 \le r \le 18/25$ yields

$$\frac{33}{49} \le p \le \frac{4}{5}. \tag{21}$$

Necessary inequalities alone would not prove sharpness: kinetics, the mean constraint and the readout must be satisfiable *simultaneously*. Take the witness (20) and set

$$(s, f, r) = \begin{cases} (1,\ (17/25 - p)/(1-p),\ 17/25), & 33/49 \le p \le 17/25, \\ (1,\ 0,\ p), & 17/25 \le p \le 18/25, \\ ((18/25)/p,\ 0,\ 18/25), & 18/25 \le p \le 4/5. \end{cases} \tag{22}$$

In the first branch the readout is perfectly sensitive and the observed total is held at its lower bound by admitting exactly enough false positives; the required f is nonnegative and at most $1/50$ precisely on that subinterval. In the middle branch the readout is perfect in both directions, so $r = p$, which lies in the observed band there. In the third branch there are no false positives and sensitivity is reduced to hold r at its upper bound; $(18/25)/p \ge 9/10$ exactly when $p \le 4/5$. Each branch therefore satisfies every bound in (15) on its portion of (21), so the same mean-constrained kinetic population supplies the readout. For an explicit binary-labelled population, split each capacity atom into positive- and negative-readout atoms in the conditional proportions just computed; this leaves capacities and trajectories unchanged and uses at most three distinct capacities with at most six labelled atoms. This proves Theorem 4.1.

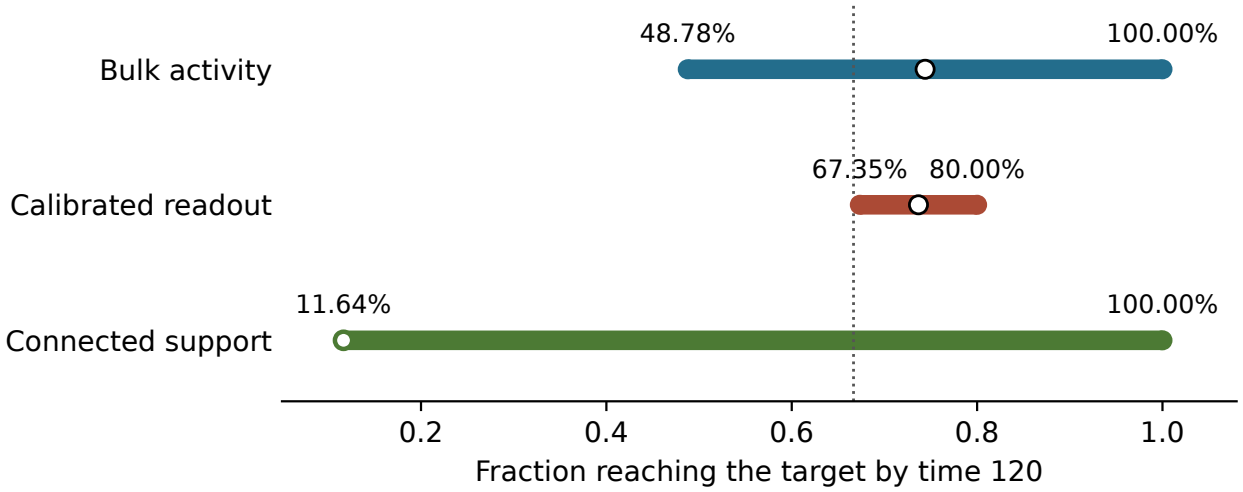

Figure 1: Exact sets of compatible recovery fractions. Filled endpoints are attained and white points mark minimax midpoints; the open endpoint of the connected-support set (Proposition 6.1) is an unattained infimum. The dotted line is the designed two-thirds requirement. These are identified sets for the stated assumptions, not statistical confidence intervals.

5 Decision consequences

5.1 Unavoidable estimation error

Corollary 5.1 (Absolute-error limits)**.** *Any estimate of p based only on the bulk information in Theorem 4.1 has worst-case absolute error at least* $21/82$*, and the midpoint* $61/82$ *attains that bound. Under the interval-valued readout information* (15) *the optimal midpoint and worst-case error are* $361/490$ *and* $31/490$*.*

Proof. For an identified interval $[L, U]$ and any estimate $\hat{p}$, the triangle inequality gives $U - L \leq |U - \hat{p}| + |\hat{p} - L|$, so one of the two endpoint errors is at least $(U - L)/2$; both endpoints are attained by Theorem 4.1. Conversely the midpoint has error at most $(U - L)/2$ everywhere on the interval. Substituting the two intervals gives the constants. □

The two radii are about 25.61 and 6.33 percentage points of recovery, a reduction of roughly 75.3% in identified-set width. They are not sampling standard errors, and they cannot be reduced by measuring the same pooled quantity more precisely: a deterministic procedure must return the same value for two populations whose observations are identical, and by Theorem 4.1 such pairs differ in recovery fraction by up to $21/41$. Reporting an exact readout value rather than the interval (15) would define a smaller class and a different, smaller radius.

5.2 A precision requirement for a performance target

If $s \leq s_+$, $f \leq f_+$ and $s_+ > f_+$, then $r \leq f_+ + (s_+ - f_+)p$, so a lower readout bound r_- certifies $p \geq p_*$ whenever

$$r_- \geq f_+ + (s_+ - f_+)p_*. \tag{23}$$

For the designed requirement $p_* = 2/3$ with $s_+ = 1$, $f_+ = 1/50$ the threshold is $101/150$, and the available bound $17/25$ leaves

$$\frac{17}{25} - \frac{101}{150} = \frac{1}{150}, \qquad \frac{33}{49} - \frac{2}{3} = \frac{1}{147}. \tag{24}$$

These two numbers are different quantities: the first is the additional downward readout error the certificate tolerates, the second is the slack in the certified recovery fraction itself. Replacing r_- by $17/25 - \varepsilon$ preserves the non-strict two-thirds guarantee exactly for $\varepsilon \leq 1/150$. Uncertainty already included in (15) must not be subtracted twice, and improving only the lower sensitivity bound moves the *upper* recovery bound, not this lower certificate. Two-thirds is a designed performance requirement used to exhibit decision sensitivity; it is not a clinical threshold.

As a transparent sampling illustration, suppose one positive-readout proportion is estimated from n independent representative Bernoulli draws. Hoeffding's inequality [6] gives one-sided error $\varepsilon_n = \sqrt{\log(1/\delta)/(2n)}$ at confidence $1 - \delta$; at $\delta = .05$, requiring $\varepsilon_n \leq 1/150$ gives the conservative sufficient count $n \geq 33{,}702$, and controlling both sides of that one proportion gives 41,500. These are model precision calculations under an assumed sampling model, not clinical sample-size recommendations: cell clustering, donor-level sampling, gating selection and systematic calibration error each require separate treatment.

5.3 Normalization can consume the margin

Proposition 5.2 (Change of population weights)**.** *Suppose a justified observation model supplies a weighted successful fraction r_h, and that positive cell weights have ratio at most $R_h \geq 1$. Then the*

number fraction p_n satisfies

$$\frac{r_h}{R_h(1-r_h)+r_h} \le p_n \le \frac{R_h r_h}{1-r_h+R_h r_h}, \tag{25}$$

and the weighted interval $[33/49, 4/5]$ *gives the number-fraction enclosure*

$$\Big[\frac{33}{33+16R_h},\ \frac{4R_h}{1+4R_h}\Big], \tag{26}$$

whose lower endpoint is at least $2/3$ *exactly when* $R_h \le 33/32$. *At* $R_h = 6/5$ *the enclosure is* $[55/87, 24/29] \approx [63.22\%, 82.76\%]$.

Proof. Divide all weights by their smallest positive value, so each lies in $[1, R_h]$. On a number-normalized population let A and B be the successful and failing weighted masses. Then $p_n \le A \le R_h p_n$ and $1-p_n \le B \le R_h(1-p_n)$, and $r_h = A/(A+B)$. Extremizing the ratio gives

$$\frac{p_n}{p_n + R_h(1-p_n)} \le r_h \le \frac{R_h p_n}{R_h p_n + 1 - p_n},$$

and inverting these Möbius maps, which are increasing in p_n, gives (25). Both bounds in (25) increase in r_h, since their derivatives have numerator $R_h > 0$; substituting the endpoints of $[33/49, 4/5]$ gives (26), and clearing positive denominators proves the remaining claims. □

At $R_h = 6/5$ the converted lower endpoint no longer certifies two-thirds. Proposition 5.2 converts a weighted *success fraction*; it does not convert a mean activity in U/g Hb into a recovery fraction, for which a separate justified map is needed. If capacities, weights and calibration are coupled by further constraints, (25) remains valid, but sharpness would have to be re-established with those constraints imposed jointly. No conversion is needed when the observation and the target are already number weighted.

5.4 One margin, two claimants

Proposition 5.3 (Exact error–normalization frontier)**.** *Keep* $s_+ = 1$, $f_+ = 1/50$, *replace the lower readout bound by* $17/25 - \varepsilon$ *with* $\varepsilon \ge 0$, *and let* $\alpha(\varepsilon) = (33 - 50\varepsilon)/49$ *be the resulting weighted lower bound, assumed to lie in* $[0, 1]$. *For a cell-weight ratio* $R_h \ge 1$, *the composed certificate* $p_n \ge 2/3$ *obtained by applying Proposition 5.2 to* $\alpha(\varepsilon)$ *holds if and only if*

$$\varepsilon \le \frac{33 - 32R_h}{50(1+2R_h)}. \tag{27}$$

In particular the budget is $1/150$ *at* $R_h = 1$, *only* $9/3800 \approx .0023684$ *at* $R_h = 51/50$, *and empty for* $R_h > 33/32$.

Proof. By (23) with $r_- = 17/25 - \varepsilon$ the weighted bound is $\alpha(\varepsilon)$. By (25) the converted lower endpoint is at least $2/3$ iff $3\alpha \ge 2R_h(1-\alpha) + 2\alpha$, i.e. iff $\alpha \ge 2R_h/(1+2R_h)$. Substituting $\alpha = (33 - 50\varepsilon)/49$ and clearing the positive denominators $49(1+2R_h)$ gives $33 - 32R_h \ge 50\varepsilon(1+2R_h)$, which is (27). □

Independent error allowances therefore cannot each consume the whole original margin: they are claims on one exactly computable budget (Figure 2). At $R_h = 6/5$ the certificate would need a weighted successful fraction of at least $12/17$, hence a lower positive-readout bound of $121/170 \approx .71176$, which the available $17/25 = .68$ does not supply. Proposition 5.3 is an exact algebraic frontier for this composed enclosure. It is not a claim that the two error sources are biologically independent, and a joint model coupling weights to capacity would need its own analysis.

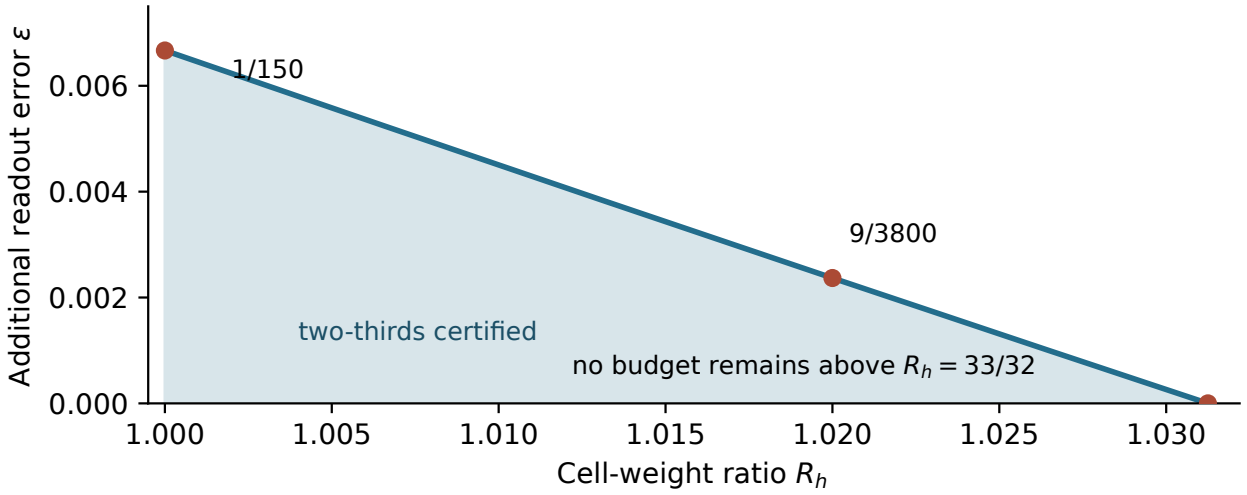

Figure 2: The frontier (27). Inside the shaded region the composed certificate still gives a number fraction of at least two-thirds. The region closes at $R_h = 33/32$, where weight normalization alone exhausts the margin.

6 What the support and deadline assumptions contribute

The results of this section are ordinary mathematics, with the passage-time brackets supported by exact rational arithmetic. They are not claimed as machine-verified theorems; Table 6 records the distinction.

6.1 Longer horizons on the separated support

For any $T \geq T_0$ from (17), construct solutions of (4) on the actual interval $[0, T]$ by the same compact-interval argument. Low-band cells cannot hit on that interval, by the same first-hit contradiction. Restricting a high-band solution to $[0, T_0]$ and applying the positive-drift bound shows it hits by $T_0 \leq T$. Hence the successful fraction is again the high-band mass, and the witnesses (20) prove the same exact interval $[20/41, 1]$ for every such horizon: on this support, waiting longer removes none of the bulk ambiguity. This is a statement about solutions constructed on $[0, T]$; it does not extrapolate a trajectory known only on $[0, 120]$.

6.2 Connected support and the inclusive cutoff

Now admit every capacity in $[a, d] = [5/8, 11/8]$, keeping mean one, the same load and the same task. The deadline-free threshold is

$$V_\infty = \frac{q}{k(28)} = \frac{363}{560} \approx .6482. \tag{28}$$

For $V > V_\infty$ the drift is positive up to the target, and the first passage time is

$$\tau(V) = \int_{10}^{28} \frac{dg}{Vk(g) - q}. \tag{29}$$

To justify (29) for an actual solution rather than as a separation-of-variables ansatz, note that the primitive $G(g) = \int_{10}^{g} (Vk - q)^{-1}$ satisfies $\frac{d}{dt} G(g(t)) = 1$ before the first hit; the positive target drift

supplies a hit by the argument of Theorem 3.2, and evaluating at the hitting time gives (29). For $V \le V_\infty$ no finite hit occurs: negative target drift excludes first arrival, and in the equality case the target is an equilibrium, which uniqueness forbids reaching from a different initial state.

The integrand is positive and strictly decreasing in V, and on a compact neighbourhood of any $V > V_\infty$ its denominator is bounded away from zero, so τ is continuous and strictly decreasing there. With $B_0 = 109/56$, $\alpha = 56V - 115q$ and $\beta = V - qB_0 > 0$, partial fractions give the closed form

$$\tau(V) = \frac{18B_0}{\beta} + \frac{6V}{\beta^2} \log \frac{\alpha - 10\beta}{\alpha - 28\beta}. \tag{30}$$

Since $\alpha - 28\beta = 28(V - V_\infty)$, the logarithm diverges as $V \downarrow V_\infty$ while its numerator stays positive, so $\tau(V) \to \infty$; and $\tau(V) \le 18/(Vk(28) - q) \to 0$ as $V \to \infty$. Hence for each $T > 0$ there is a unique cutoff $c_T > V_\infty$ with $\tau(c_T) = T$, and success by T is exactly $V \ge c_T$.

Proposition 6.1 (Connected-support identified set at the deadline 120)**.** *Let $c = c_{120}$. Then $19/20 < c < 951/1000$, and the exact set of compatible recovery fractions under mean one and support $[a, d]$ is*

$$\Big(\frac{1-c}{11/8-c},\ 1\Big], \tag{31}$$

whose lower endpoint is an infimum and is not *attained. The bracket alone gives $49/424 < \inf p < 2/17$, and numerically $c \approx .9506018463$, $\inf p \approx .1163957789$.*

Proof. Since k decreases, $h_V(g) = (Vk(g) - q)^{-1}$ increases on $[10, 28]$ for $V > V_\infty$, so for a uniform partition with $n = 128$ the left and right rectangle sums bracket (29) from below and above. For rational V these sums are rational and are evaluated exactly in Table 3; the rows $V = 19/20$ and $V = 951/1000$ give $\tau(19/20) > 120 > \tau(951/1000)$, and monotonicity of τ gives the bracket for c.

For validity, $I = \mathbf{1}\{V \ge c\}$ satisfies $V \le c + (d - c)I$ on $[a, d]$, because a failing capacity is below c and a successful one is at most d; averaging against mean one gives $p \ge (1 - c)/(d - c)$, and normalization gives $p \le 1$. For sharpness, let $x \in [a, c)$ and mix atoms at x and d with successful weight $p_x = (1 - x)/(d - x)$, which has mean one; p_x decreases to $(1 - c)/(d - c)$ as $x \uparrow c$. Mixing such a population with the uniform capacity-one population attains every larger value up to one. Finally the endpoint is not attained: equality in the averaged inequality requires every unit of failing mass to sit exactly at c, where it would in fact succeed, and if there is no failing mass equality requires all capacity at d, contradicting mean one. □

Capacity V	Lower passage bound	Upper passage bound	Hit by 120?
19/20	120.171857	120.285392	No
951/1000	119.792912	119.905845	Yes
7/8	157.558903	157.739481	No
9/8	77.356392	77.411824	Yes
11/8	51.267354	51.297023	Yes

Table 3: Outward decimal roundings of exact rational left and right rectangle sums with 128 subintervals. Their reading as passage-time bounds uses the ordinary argument for (29); no numerical integration certificate is involved, and the decimal value of c_{120} is not a premise anywhere.

The support gap does substantial inferential work: removing it moves the guaranteed fraction from $20/41 \approx .4878$ to an unattained infimum near .1164, a far larger change than any improvement

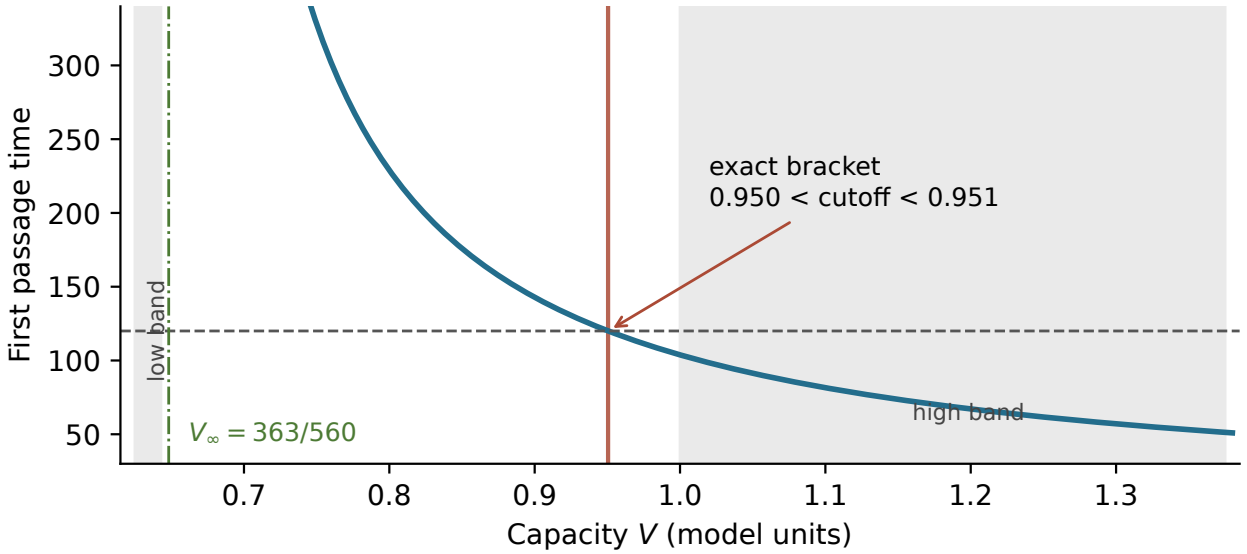

Figure 3: First passage time (29) against capacity. Shaded vertical strips are the two designed support bands of (5); the gap between them is what makes the classification (18) elementary. The plotted curve is a numerical evaluation; the marked cutoff has the separate exact rational bracket $(.950, .951)$, and V_∞ is the deadline-free threshold (28), at which τ diverges.

in the precision of an already exact bulk mean. A small explicit example: 15 cells at $V = 19/20$ and two at $V = 11/8$ have mean one, but only the two recover by 120, whereas all 17 cells of the uniform capacity-one population recover. The 15 slower cells do eventually recover; the deadline is essential to the example.

6.3 Mean and variance do not identify the outcome

The two connected-support populations

$$\mu_A = \tfrac{3}{4}\delta_{7/8} + \tfrac{1}{4}\delta_{11/8}, \qquad \mu_B = \tfrac{1}{4}\delta_{5/8} + \tfrac{3}{4}\delta_{9/8} \tag{32}$$

both have mean one and variance $3/64$, yet their recovery fractions are $1/4$ and $3/4$. Table 3 classifies the interior atoms $7/8$ and $9/8$, and $5/8 < V_\infty$ supplies the remaining failure. So adding a second moment to the bulk information does not identify the recovery fraction on this class; it does not follow that variance never tightens a bound, and Section 7 gives the criterion that decides such questions.

6.4 Deadline dependence and the deadline-free limit

Proposition 6.2 (Sharp limits at a general deadline)**.** *Fix mean one and support $[a, d]$. If $a < c_T < d$, the sharp lower and upper limits of the recovery fraction are*

$$L(T) = \max\left\{0, \frac{1 - c_T}{d - c_T}\right\}, \qquad U(T) = \min\left\{1, \frac{1 - a}{c_T - a}\right\}. \tag{33}$$

For $c_T > 1$ the value $L(T) = 0$ is attained by the uniform capacity-one population and $U(T)$ by a mixture at a and c_T; for $c_T \le 1 < d$ the lower endpoint is an unattained infimum, including the

value 0 *at* $c_T = 1$*, and* $U(T) = 1$ *is attained. If* $c_T > d$ *no cell succeeds; if* $c_T \le a$ *all cells succeed; and at* $c_T = d$ *the set of possible fractions is* $[0, (1-a)/(d-a)]$*. Finally, as* $T \to \infty$*,*

$$c_T \downarrow V_\infty, \qquad L(T) \longrightarrow \frac{197}{407} < \frac{1}{2}, \tag{34}$$

and the limiting value is attained, by weights 210/407 *and* 197/407 *at* V_∞ *and* 11/8*.*

Proof. Validity of (33) follows from the two affine majorants $a + (c_T - a)I \le V \le c_T + (d - c_T)I$ exactly as in Proposition 6.1, and the attainment statements from the two-atom mixtures named, whose means are one by construction. The case distinctions merely avoid dividing by a vanishing denominator at the support endpoints. For (34), eventual success is $V > V_\infty$ by the discussion after (30), and averaging $V \le V_\infty + (d - V_\infty)\mathbf{1}\{V > V_\infty\}$ gives the same expression with c_T replaced by V_∞; here the threshold atom itself *fails*, which is why the limiting endpoint is attained although every finite-deadline endpoint with $c_T \le 1$ is not. □

Deadline T	Cutoff c_T	Sharp lower limit	Sharp upper limit
60	1.26730622	0.00000000	0.58383367
120	0.95060185	0.11639578	1.00000000
240	0.79237408	0.35636231	1.00000000
1000	0.67309407	0.46574037	1.00000000

Table 4: Numerical evaluations of (33) for connected support. The lower limit is attained in the row $T = 60$, where $c_T > 1$, and is an unattained infimum in the other rows. Only the deadline 120 carries the exact rational bracket of Table 3.

Even with no deadline at all, the guarantee stays below one half: a population with mean activity one, on this support, can leave a majority of its cells unable ever to reach the target.

Proposition 6.3 (Rate and constant of the long-time approach)**.** *Write* $\delta = V - V_\infty \downarrow 0$*. Then*

$$\tau(V) = \frac{8470}{9}\log\frac{1}{\delta} + 545 + \frac{8470}{9}\log\frac{81}{1960} + O\big(\delta\log(1/\delta)\big), \tag{35}$$

where $545 + \frac{8470}{9}\log\frac{81}{1960} \approx -2453.6158$*. Equivalently*

$$c_T - V_\infty = \frac{81}{1960}\, e^{-9(T-545)/8470}\,(1 + o(1)), \qquad T \to \infty, \tag{36}$$

and consequently $L(\infty) - L(T) = \Theta\big(e^{-9T/8470}\big)$*, with first-order sensitivity* $dL/dc = -(d-1)/(d - V_\infty)^2 = -117600/165649$ *at* $c = V_\infty$*.*

Proof. At $V = V_\infty$ one has $\beta_\infty = 9/140$, $\alpha_\infty = 9/5$ and $\alpha_\infty - 10\beta_\infty = 81/70$, so $18B_0/\beta_\infty = 545$ and $6V_\infty/\beta_\infty^2 = 8470/9$. In (30) substitute $\alpha - 28\beta = 28\delta$ and split the logarithm:

$$\tau(V) = \frac{18B_0}{\beta} + \frac{6V}{\beta^2}\log\frac{1}{\delta} + \frac{6V}{\beta^2}\log\frac{\alpha - 10\beta}{28}.$$

The maps $V \mapsto 18B_0/\beta$, $V \mapsto 6V/\beta^2$ and $V \mapsto \log((\alpha - 10\beta)/28)$ are smooth at V_∞ with the values just computed, so each contributes its value plus $O(\delta)$; the only term that interacts with the divergence is $(6V/\beta^2 - 8470/9)\log(1/\delta)$, which is $O(\delta\log(1/\delta))$. This gives (35), and solving $\tau(c_T) = T$ for δ gives (36). Differentiating $c \mapsto (1-c)/(d-c)$ gives the stated sensitivity, and L is smooth and strictly decreasing near V_∞, so the difference $L(\infty) - L(T)$ has the same exponential order as $c_T - V_\infty$. □

The rate constant $9/8470 \approx 1.06 \cdot 10^{-3}$ per unit model time is the quantitative form of diminishing returns from waiting: to halve the gap between $L(T)$ and its limit one must roughly add $8470 \log 2/9 \approx 652$ time units. Two cautions belong with (36). It is an asymptotic statement in the selected units, so no physiological horizon should be read off it; and its remainder has a large coefficient, so the leading approximation is not yet numerically accurate at moderate deadlines. At $T = 1000$, for instance, the exact cutoff is .67309 while (36) gives .67379; agreement to eight digits requires horizons of order $3 \cdot 10^4$.

6.5 Sensitivity to the support gap itself

The designed gap is an assumption about a population, and it is the kind of assumption that data can violate partially rather than entirely. The following proposition interpolates, with the violating mass as an explicit, experimentally reportable parameter.

Proposition 6.4 (Off-band robustness)**.** *Put $b = 9/14$, $d = 11/8$, $c = c_{120}$, and fix $\eta \in [0, 1]$. Consider populations with capacities in $[5/8, d]$, mean one, and total weight at most η on capacities outside the two designed bands $[5/8, b] \cup [1, d]$. Then*

$$\inf p = L_\eta := \max\Big\{\frac{1-c}{d-c},\ \frac{1-b-(c-b)\eta}{d-b}\Big\}, \tag{37}$$

and the two expressions cross at $\eta^\star = (d-1)/(d-c) \approx .8836$. The infimum is attained only for $\eta = 0$, where $L_0 = 20/41$ recovers Theorem 4.1; for every $\eta > 0$ it is unattained. Numerically $L_{1/10} \approx .4458$.

Proof. Validity. Classify each cell as successful, off-band failing, or on-band failing. A successful capacity is at most d; an off-band failing capacity is below c; an on-band failing capacity is at most b, since the low designed band ends at b and the high band consists of successes. Hence pointwise

$$V \le b + (c-b)\mathbf{1}\{\text{off-band failing}\} + (d-b)\mathbf{1}\{\text{successful}\},$$

and averaging against mean one gives $1 \le b + (c-b)\xi + (d-b)p$, where $\xi \le \eta$ is the off-band failing mass; this is the second expression in (37). The first expression is the connected-support bound of Proposition 6.1, which holds under these hypotheses as well.

Sharpness for $\eta \le \eta^\star$. Take three atoms: mass η at some off-band $x < c$, the successful mass w_d at d, and the remainder at b. Mean one forces $w_d = (1 - b - \eta(x-b))/(d-b)$, which tends to the second expression in (37) as $x \uparrow c$. The weights are admissible iff $w_d \le 1 - \eta$, and at $x = c$ this reads $\eta(d-c) \le d - 1$, i.e. $\eta \le \eta^\star$; by continuity it holds for x close enough to c whenever $\eta < \eta^\star$, and for $\eta = \eta^\star$ the two expressions in (37) coincide.

Sharpness for $\eta \ge \eta^\star$. Use the two-atom connected-support family of Proposition 6.1 at $x \uparrow c$ and d; its off-band mass is $1 - p_x \uparrow 1 - (1-c)/(d-c) = \eta^\star \le \eta$, so it is admissible and approaches the first expression.

Nonattainment for $\eta > 0$. If a population has no off-band failing mass, then $p \ge (1-b)/(d-b) = 20/41 > L_\eta$ because $L_\eta < 20/41$ for $\eta > 0$. Otherwise the averaging above is strict, since off-band failing capacities are strictly below c. At $\eta = 0$ the second expression is $20/41$, attained by the two-atom witness of (20). □

Proposition 6.4 turns a yes-or-no modelling assumption into a graded one: reporting an upper bound on off-band mass, which a cell-resolved assay can do directly, yields a guarantee that degrades smoothly from $20/41$ at $\eta = 0$ to the connected-support infimum once η reaches $\eta^\star$. No smooth distributional family has to be fitted to obtain it.

7 When can a further observation help?

Sections 3–6 exhibit one useful additional observation, the calibrated readout, and two uninformative ones: added precision in a coefficient direction the task does not see (Section 3), and a second moment of the capacity distribution (Section 6). The distinction is not a matter of assay quality, and it can be decided before any data are collected.

Proposition 7.1 (Row-span criterion)**.** *Fix finitely many distinct capacities $v_1, \dots, v_m$ and let the rows of $A \in \mathbb{R}^{k\times m}$ be observed population features, including the constant row $\mathbf{1} = (1, \dots, 1)$. Let $h \in \mathbb{R}^m$ be the target row, $h_j = 1$ if v_j succeeds and 0 otherwise. Then the target $h \cdot w$ is determined by the observations Aw for* every *weight vector w in the simplex if and only if h lies in the row span of A.*

Proof. If $h = \lambda^\top A$ then $h \cdot w = \lambda^\top(Aw)$ is a function of the observations. Conversely suppose $h \notin \operatorname{rowspan} A$. Then there is $z \in \mathbb{R}^m$ with $Az = 0$ and $h \cdot z \neq 0$. Decompose $z = z^+ - z^-$ into positive and negative parts. The constant row gives $\mathbf{1} \cdot z = 0$, so z^+ and z^- have the same total mass m_0, which is positive because $z \neq 0$. Then $w = z^+/m_0$ and $w' = z^-/m_0$ lie in the simplex, satisfy $Aw = Aw'$, and $h \cdot w - h \cdot w' = (h \cdot z)/m_0 \neq 0$. □

Two remarks make Proposition 7.1 usable. It is a statement about identification for all weights; at a *particular* observed value the nonnegativity constraints may restrict the feasible face enough to pin the target down even when the row-span condition fails, which is exactly the situation of Theorem 4.1, where the identified set is a proper subinterval of $[0, 1]$ rather than a point or everything. And the criterion is constructive in both directions: a certificate is a vector of multipliers, and a refutation is a null direction, which also localizes *where* in capacity space the ambiguity lives.

Example 7.2 (A correctly measured second assay with zero value)**.** Reparameterize the family of Example 3.3 by the inhibition coordinate $x = 3K/(3K + 56) \in (0, 1)$, so that the target rate is $56x/(109x + 9)$ and larger x means weaker inhibition. Take the designed support $\{1/20, 1/10\} \cup \{1/5, 4/5\}$, mean $\mathbb{E}[x] = 9/80$, and deadline 300; the two upper atoms then recover and the two lower atoms have negative target drift and never recover. Let the observations be $A_1(x) = x$ and the nonlinear $A_2(x) = 2x/(1 + x)$, both measured exactly. From the mean alone, the affine majorants $x \le 1/10 + (7/10)\mathbf{1}$ and $x \ge 1/20 + (3/20)\mathbf{1}$ give the sharp interval $p \in [1/56, 5/12]$, attained by

$$\nu_A = \tfrac{55}{56}\delta_{1/10} + \tfrac{1}{56}\delta_{4/5}, \qquad \nu_B = \tfrac{7}{12}\delta_{1/20} + \tfrac{5}{12}\delta_{1/5}.$$

Both endpoint populations satisfy $\mathbb{E}[A_2] = 7/36$, as do all their mixtures, so adding the second assay leaves the entire interval unchanged. The diagnostic behind this coincidence is the exact null vector $(-98, 165, -70, 3)$ on the atoms $(1/20, 1/10, 1/5, 4/5)$: it annihilates the rows $\mathbf{1}$, A_1 and A_2 while pairing to -67 against the success row, so by Proposition 7.1 no function of those three observations can determine the target. A denser search for a better estimator based on the same two assays is therefore pointless, whereas the calibrated classification readout of Theorem 4.1 works precisely because its row is not annihilated.

8 From a scalar threshold to multidimensional dynamics

The population argument used only one structural fact about the dynamics: the set of successful capacities is an up-set of the form $\{V \ge c\}$. The following proposition isolates hypotheses under which that survives in a multidimensional model, and the subsequent counterexample shows which plausible weakening does not suffice. Neither statement is verified for any particular erythrocyte network.

Proposition 8.1 (Monotone transfer of the threshold)**.** *Let $D \subset \mathbb{R}^n$ be compact, order convex for the componentwise order, and forward invariant for $x' = F(x, V, t)$, where F is continuous and locally Lipschitz in x, solutions from a common $x_0 \in D$ are unique on $[0, T]$, and V ranges over a compact interval $\mathcal{V}$. Assume*

1. *F is quasimonotone in the state: $\partial F_i/\partial x_j \geq 0$ for $i \neq j$;*
2. *$V \mapsto F(x, V, t)$ is componentwise nondecreasing;*
3. *the success set $K \subseteq D$ is closed and upward closed.*

Then $\{V \in \mathcal{V} : x(t; V) \in K \text{ for some } t \in [0, T]\}$ is a closed up-set, hence empty, all of $\mathcal{V}$, or $[c_T, \max \mathcal{V}]$ for some cutoff c_T.

Proof. Hypotheses (i) and (ii) make $V_1 \leq V_2$ a pair of ordered vector fields on an order-convex invariant domain, so the comparison theorem for quasimonotone systems [17, 15, 2] gives $x(t; V_1) \leq x(t; V_2)$ for all $t \in [0, T]$. If the smaller-capacity trajectory enters the upward closed K at some time, the larger one is in K at that time, so the successful set is an up-set. For closedness, let $V_j \to V$ with hits at $t_j \in [0, T]$; pass to a convergent subsequence $t_j \to t$, and use continuous dependence on parameters and closedness of K to conclude $x(t; V) \in K$. □

Adding cumulative service as an extra state variable is permitted, provided its order and every accompanying constraint are checked; a different order may be needed when increased capacity is favourable for one coordinate and adverse for another. Substrate competition, coupled consumption and changing initial states can all break hypothesis (i) or (ii).

Example 8.2 (Cooperativity in the state is not enough)**.** Consider the mathematical resource system

$$y' = 1 - (1 + V^2)y, \qquad x' = Vy - x, \qquad y(0) = 1, \quad x(0) = 0, \quad V \geq 0. \tag{38}$$

The nonnegative quadrant is invariant and, for each fixed V, the system is cooperative: $\partial x'/\partial y = V \geq 0$ and $\partial y'/\partial x = 0$. But the parameter acts in two directions, increasing production of x while depleting its support y faster, so hypothesis (ii) fails. Explicitly

$$x(t; V) = \frac{V}{1 + V^2}\Big(1 - e^{-(1+V^2)t}\Big), \tag{39}$$

so with success defined as $x \geq 1/4$ at some time in $[0, 1]$, capacity $V = 1$ succeeds, since $x(1; 1) = (1 - e^{-2})/2 > 1/4$, while capacity $V = 4$ fails at *every* time, since $x(t; 4) < 4/17 < 1/4$ for all t. The successful set is therefore not an up-set. This is an illustrative positive system, not a proposed G6PD mechanism; its role is to show that positivity, or cooperativity in the state alone, cannot justify a monotone capacity threshold.

When a clean threshold is unavailable, scalar envelopes still give something. If r is a reserve coordinate of an actual multidimensional trajectory and validated reachable-state bounds give $\underline{f}(r, V, t) \leq r' \leq \overline{f}(r, V, t)$, then solving the two scalar comparison equations from compatible initial data yields $r_- \leq r \leq r_+$ [17]. If r_- reaches the target by T, the actual reserve does; if r_+ stays strictly below the target on $[0, T]$, the actual reserve cannot succeed. This partitions capacities into certified success, certified failure and an unresolved class, and the population bounds of Section 4 apply to any such certified classification. The envelopes must follow from the joint model or from explicitly stated assumptions about its reachable states; independent extrema over incompatible hidden states may be conservative, and a trajectory that saturates an envelope need not be realizable, so envelope validity does not confer sharpness.

9 Using the results as a measurement requirement

Consider the number-weighted example. A population of 82 cells at capacity one all recover by time 120. A population of 42 cells at 9/14 and 40 at 11/8 has the same mean, the same *entire* clamped response surface (7), and only 40 recovering cells. Bulk activity cannot separate them, and by Corollary 5.1 no estimator based on that information can do better than a worst-case error of 21/82 on this class.

A readout satisfying (15) restricts the class and certifies $p \geq 33/49 > 2/3$. Note what this does to the example: the 42/40 population has $p = 40/82 < 33/49$, so it is *excluded* by the readout information rather than reconciled with it; the populations compatible with both the mean and the readout are those of (20)–(22) with $p \in [33/49, 4/5]$. The certificate's margin is thin in two independent senses: the tolerable additional readout error is 1/150, and interpreting a weighted fraction as a number fraction consumes the same budget, which Proposition 5.3 shows closes entirely at a weight ratio of 33/32.

A different route is available when a cell-resolved clamped rate $A_i = V_i\psi$ is measured with known $\psi > 0$ and error $|\widehat{A}_i - A_i| \leq \varepsilon_A$. Setting $Z_i = \widehat{A}_i/\psi$ and $e = \varepsilon_A/\psi$ gives $|Z_i - V_i| \leq e$, so for any cutoff enclosure $c \in [c_-, c_+]$,

$$\sum_i w_i \mathbf{1}\{Z_i \geq c_+ + e\} \leq p \leq \sum_i w_i \mathbf{1}\{Z_i \geq c_- - e\}. \tag{40}$$

The left event forces $V_i \geq c$, and actual success forces the right event; the gap between the two sides is exactly the observed mass in $[c_- - e, c_+ + e)$. With $e = .01$ and the bracket of Proposition 6.1 that ambiguity band is $[.940, .961)$. The design requirement this expresses is specific and testable: resolve the population mass *near the functional threshold*. Uncertainty can be small without high precision everywhere if little mass sits in that band, and conversely high nominal resolution is worthless if the calibration is biased or if cells near the threshold are selected out. No joint sharpness under retained moment constraints is claimed for (40).

Fluorescence intensity is not automatically the variable Z: it requires a justified measurement map, a calibration factor and an error budget. Existing cell-resolved G6PD methods [13, 7] motivate this task but do not supply this model's functional calibration, and more cells reduce sampling error without correcting selection bias or an unsupported kinetic map. What an analysis must record in order to support any of the conclusions above is listed in Table 5; a software implementation of these bounds should return a compatible interval, the assumption responsible for each endpoint, and the measurement whose uncertainty limits the declared decision, and should return "not certified from these inputs" rather than a falsely precise activity-to-outcome conversion.

10 Scope, formal evidence and reproduction

The results separate three questions that a single enzyme name tends to merge: a kinetic task decides which capacities succeed; pooled observations constrain population moments; and a calibrated additional observation restricts the mass near the functional threshold. The same reaction-rate structure can therefore coexist with markedly different finite-time operating capability while pooled observations fail to distinguish the cases. This is a task-specific kinetic example with exact identified sets, not a general classification theorem for biochemical networks, and Theorem 4.1 identifies the failure of *one* observation class for *one* functional task: bulk assays remain informative for other tasks, or for this task under additional justified assumptions.

What remains open is specific. The scalar state is not a full NADPH balance, and nothing here establishes simultaneous energy service, peroxide control and whole-cell restart, which would require

Record field	Why it changes the inference
Target event and deadline	define which capacities count as successful, (6)
Initial state and relevant pools	set the distance and the resources needed, (4)
Assay substrate, inhibitor, temperature, preparation	decide whether one common clamped rate map is justified, (7)
Activity, fraction and target normalization	separate number fractions from mass-weighted quantities, Proposition 5.2
Observation map and calibration bounds	connect signal to task or capacity, (15) and (40)
Sampling and gating rules	determine representativeness and selection bias
Support evidence, including off-band mass	fix which robustness statement applies, Proposition 6.4
Joint covariates and their uncertainty	prevent combining separately plausible but jointly infeasible cases
Functional observations from the same preparation	test the declared kinetic-to-task relationship at all

Table 5: Minimum analysis record for applying the bounds of this paper. This is a computational interpretation specification, not a clinical protocol.

matched source dynamics, observations and endpoint definitions in a single preparation. Neither the support, the load, the deadline nor the readout calibration has been fitted to donor measurements, and Remark 2.2 records why the individual and pooled uncertainty levels of Sections 3 and 4 do not compose without an additional hypothesis. Machine verification certifies stated implications under their hypotheses; it does not validate the hypotheses experimentally.

Claim	Evidence
Literal solutions, the two-band classification (18), and both exact identified sets of Theorem 4.1	machine-verified root plus the ordinary proof of Section 4
Minimax constants, decision rule (23), readout error budget, weight conversion (25)	machine-verified consequences plus the ordinary proofs of Section 5
Individual dual certificate, uniform deadline (13), coefficient bijection (9)	machine-verified; Example 3.3 is exact rational arithmetic
Rate-error normalization and the individual threshold implications behind (40)	machine-verified lemmas; summing the indicators is ordinary
Frontier (27), arbitrary horizons, connected-support cutoff and nonattainment, deadline dependence, eventual limit, asymptotics (35), off-band robustness (37), row-span criterion, monotone transfer and its counterexample	ordinary analysis, Sections 5–8
Passage-time brackets of Table 3 and the finite-population identities	exact rational arithmetic, read through the ordinary passage argument
Plotted curves, the cutoff and horizon decimals, sampling counts	numerical evaluation, or calculation from the cited concentration bound
Matched biological calibration of task, support or readout	not established

Table 6: Evidence scopes. Verification of the primary root does not silently extend to every analytic or statistical consequence, and an ordinary proof is mathematical evidence whether or not it has been formalized.

Formal development. The proofs are developed in Lean 4 [4] with mathlib [10]; the release pins Lean 4.30.0 and its mathlib dependencies. All declarations live in the namespace `G6PDReserve`, the population ones in `G6PDReserve.Population`, and all modules in one proof directory; Table 7 maps the statements of this paper to them. The primary declaration composes the whole population result: it constructs literal solutions of (4) for every admitted finite population, quantifies over all such trajectory families, and proves both realizability equivalences, with a companion covering arbitrary finite index types. Compilation treats warnings as errors, the source closure uses only the standard logical axioms `Classical.choice`, `Quot.sound` and `propext`, and no numerical computation is a premise of any verified statement. Two caveats on scope: the verified identification lemmas are the coefficient-space statements, so Proposition 7.1 as stated for finite-support populations rests on its ordinary proof above; and the nonarrival lemma is proved for general decreasing rates with nonpositive target drift, which is slightly more general than the strict inequality used in Section 4.

Statement in this paper	Declaration
PopulationRoot.lean	
Theorem 4.1, composed	`population_root`
arbitrary finite index type	`population_guarantee`
both realizability parts	`sharp_bulk`, `sharp_repaired`
WeightedPopulation.lean	
Bound (19), surface (7), witness (20)	`fraction_bounds`, `bulk_surface`, `witness_fraction`
ReadoutBounds.lean	
Repair (21) and its joint witness (22)	`readout_bounds`, `readout_feasible`
Normalization and threshold steps of (40)	`normalized_rate_error`, `capacity_error_implications`
PopulationConsequences.lean	
Corollary 5.1	`bulk_minimax`, `repaired_minimax`
Budget (24), conversion (25), boundary 33/32	`extra_error_budget`, `weight_conversion`, `normalization_boundary`
CapacityDynamics.lean, *SourceExistence.lean*, *RecoveryClassification.lean*	
Classification (18), existence	`hits_iff`, `solution_exists`
Invariant existence and uniqueness	`scalar_solution_exists`, `compact_solution_unique`
Drift arrival and nonarrival, secant bound	`positive_drift_reaches`, `decreasing_rate_no_finite_arrival`, `scalar_secant_bound`
CoefficientRecovery.lean, *ObservationBounds.lean*, *SourceRate.lean*, *FunctionalIdentification.lean*	
Theorem 3.2	`certified_observations_recover`
Bijection (9), certificate (12)	`coefficients_parameters`, `finite_upper_certificate`
Reciprocal expansion $1/v = \phi \cdot \beta$	`reciprocal_source_rate`
Identification in coefficient space	`row_combination_identifies`, `null_direction_nonidentification`

Table 7: Machine-verified declarations, grouped by module. Namespaces are `G6PDReserve` and, for the population results, `G6PDReserve.Population`. Statements absent from this table, including everything in Sections 6–8, carry the ordinary proofs given in the text.

Reproduction. The source package contains `main.tex`, the bibliography, the three vector figures and `check_paper.py`. The script recomputes, in exact rational arithmetic, every fraction printed in this paper, including the witness weights, the calibration branches, the minimax and budget constants, the conversion endpoints, the frontier (27), the rectangle sums of Table 3 and their outward decimal roundings, the null-vector identities of Example 7.2 and the exact quantities in Table 2; it verifies the asymptotic constant of Proposition 6.3 in high precision, labels every numerical value as such in `results.json`, and regenerates Figures 1–3. Proof sources, verification receipts and the toolchain lock accompany the formal development.

References

[1] Steven J. Altschuler and Lani F. Wu. Cellular heterogeneity: do differences make a difference? *Cell*, 141(4):559–563, 2010. https://doi.org/10.1016/j.cell.2010.04.033.

[2] David Angeli and Eduardo D. Sontag. Monotone control systems. *IEEE Transactions on Automatic Control*, 48(10):1684–1698, 2003. https://doi.org/10.1109/TAC.2003.817920.

[3] Dimitris Bertsimas and Ioana Popescu. Optimal inequalities in probability theory: A convex optimization approach. *SIAM Journal on Optimization*, 15(3):780–804, 2005. https://doi.org/10.1137/S1052623401399903.

[4] 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. https://doi.org/10.1007/978-3-030-79876-5_37.

[5] Zachary B. Haiman, Alicia Key, Angelo D'Alessandro, and Bernhard O. Palsson. RBC-GEM: A genome-scale metabolic model for systems biology of the human red blood cell. *PLOS Computational Biology*, 21(3):e1012109, 2025. https://doi.org/10.1371/journal.pcbi.1012109.

[6] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. https://doi.org/10.1080/01621459.1963.10500830.

[7] Michael Kalnoky, Germana Bancone, Maria Kahn, Cindy S. Chu, Nongnud Chowwiwat, Pornpimon Wilaisrisak, Sampa Pal, Nicole LaRue, Brandon Leader, Francois Nosten, and Gonzalo J. Domingo. Cytochemical flow analysis of intracellular G6PD and aggregate analysis of mosaic G6PD expression. *European Journal of Haematology*, 100(3):294–303, 2018. https://doi.org/10.1111/ejh.13013.

[8] Lucio Luzzatto and Paolo Arese. Favism and glucose-6-phosphate dehydrogenase deficiency. *New England Journal of Medicine*, 378(1):60–71, 2018. https://doi.org/10.1056/NEJMra1708111.

[9] Charles F. Manski. *Partial Identification of Probability Distributions*. Springer Series in Statistics. Springer, New York, 2003. https://doi.org/10.1007/b97478.

[10] 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. https://doi.org/10.1145/3372885.3373824.

[11] Taiko Nishino, Ayako Yachie-Kinoshita, Akiyoshi Hirayama, Tomoyoshi Soga, Makoto Suematsu, and Masaru Tomita. Dynamic simulation and metabolome analysis of long-term erythrocyte storage in adenine–guanosine solution. *PLOS ONE*, 8(8):e71060, 2013. https://doi.org/10.1371/journal.pone.0071060.

[12] Andreas Raue, Clemens Kreutz, Thomas Maiwald, Julie Bachmann, Marcel Schilling, Ursula Klingmüller, and Jens Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, 2009. https://doi.org/10.1093/bioinformatics/btp358.

[13] Shivang S. Shah, Seidina A. S. Diakite, Karim Traore, Mahamadou Diakite, Dominic P. Kwiatkowski, Kirk A. Rockett, Thomas E. Wellems, and Rick M. Fairhurst. A novel cytofluorometric assay for the detection and quantification of glucose-6-phosphate dehydrogenase deficiency. *Scientific Reports*, 2:299, 2012. https://doi.org/10.1038/srep00299.

[14] Hiroyuki Shimo, Taiko Nishino, and Masaru Tomita. Predicting the kinetic properties associated with redox imbalance after oxidative crisis in G6PD-deficient erythrocytes: A simulation study. *Advances in Hematology*, 2011:398945, 2011. https://doi.org/10.1155/2011/398945.

[15] Hal L. Smith. *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems*, volume 41 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1995.

[16] Elie Tamer. Partial identification in econometrics. *Annual Review of Economics*, 2:167–195, 2010. https://doi.org/10.1146/annurev.economics.050708.143401.

[17] Wolfgang Walter. *Differential and Integral Inequalities*, volume 55 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer, Berlin, 1970. https://doi.org/10.1007/978-3-642-86405-6.