Efficient Parameter Estimation for Oscillatory Biochemical Reaction Networks via a Genetic Algorithm with Adaptive Simulation Termination

Sekiguchi, Tatsuya
Department of Life Engineering, Faculty of Engineering, Maebashi Institute of Technology

Hamada, Hiroyuki
Department of Bioscience and Biotechnology, Faculty of Agriculture, Kyushu University

Okamoto, Masahiro
School of Interdisciplinary Science and Innovation, Kyushu University

出版情報：AppliedMath. 6 (3), pp.47-, 2026-03-16. Multidisciplinary Digital Publishing Institute : MDPI
バージョン：

AppliedMath

MDPI

Article

Efficient Parameter Estimation for Oscillatory Biochemical Reaction Networks via a Genetic Algorithm with Adaptive Simulation Termination

Tatsuya Sekiguchi [1,*], Hiroyuki Hamada [2] and Masahiro Okamoto [3]

[1] Department of Life Engineering, Faculty of Engineering, Maebashi Institute of Technology, 460-1, Kamisatori-cho, Maebashi 371-0816, Japan
[2] Department of Bioscience and Biotechnology, Faculty of Agriculture, Kyushu University, 744, Motooka, Nishi-ku, Fukuoka 819-0395, Japan; hamada@brs.kyushu-u.ac.jp
[3] School of Interdisciplinary Science and Innovation, Kyushu University, 744, Motooka, Nishi-ku, Fukuoka 819-0395, Japan; okahon@kyoso.kyushu-u.ac.jp
* Correspondence: sekiguchi@maebashi-it.ac.jp

Abstract

Parameter estimation for biochemical reaction networks is computationally demanding, especially for systems with oscillatory nonlinear dynamics, where standard iterative optimization strategies, including genetic algorithms, often struggle with prohibitive computational costs. We introduce an efficient parameter estimation framework that combines a real-coded genetic algorithm with a novel adaptive simulation termination strategy. This strategy defines a time-dependent termination boundary based on population quantiles, which is permissive during early transients and becomes progressively stricter as simulations advance, explicitly accounting for the temporal structure of oscillatory behavior. Crucially, this mechanism facilitates the efficient identification and early simulation termination of poor parameter candidates, thus avoiding the computational expense of full-horizon simulations. The framework further integrates global exploration with the modified Powell method for rapid local refinement. Numerical experiments on two benchmark oscillatory models—the Lotka–Volterra and Goodwin oscillators—demonstrate that the framework reduces computational cost by approximately 30–50% compared to a baseline GA without this strategy. For the parameter-sensitive Goodwin model, the framework efficiently identifies candidates evolving toward damped oscillations caused by subtle parameter variations. Sensitivity analysis also confirms robustness across diverse hyperparameter settings, indicating that adaptive simulation termination provides a practical acceleration mechanism for inverse problems in systems biology where iterative objective function evaluation dominates runtime.

Keywords: systems biology; inverse problems; parameter estimation; genetic algorithms

Academic Editor: Chih-Wen Chang

Received: 4 February 2026
Revised: 7 March 2026
Accepted: 12 March 2026
Published: 16 March 2026

1. Introduction

Parameter estimation for biochemical reaction networks is a fundamental challenge in systems biology [1–5]. Ordinary differential equation (ODE) models, including those based on Michaelis–Menten kinetics and their extensions, are commonly used to describe nonlinear dynamical systems such as enzymatic reactions, metabolic pathways, and cellular signaling processes. The predictive accuracy of these models critically depends on accurate kinetic parameters—such as reaction rate constants and affinity coefficients—inferred from

experimentally observed time course data. Because many kinetic parameters cannot be measured directly, computational inference is essential. However, solving this inverse problem remains difficult owing to the strong nonlinearity, complex interactions among parameters, and limited observability of internal system states.

Evolutionary computation methods, particularly real-coded genetic algorithms (GAs), have been widely applied to parameter estimation in biochemical reaction and gene regulatory networks owing to their robustness in nonconvex, multimodal search landscapes. Real-coded GAs are especially advantageous when gradient information is unavailable, a common scenario in large-scale biochemical reaction networks. Beyond systems biology, GAs have also demonstrated significant utility in solving diverse inverse problems across other engineering and physical domains, such as semiconductor device modeling [6] and thermodynamics [7]. To efficiently solve these inverse problems, hybrid strategies have been proposed in which a real-coded GA first identifies promising regions of the search space, followed by rapid local refinement using the modified Powell method [8]. Furthermore, advanced real-coded GAs employing the Just Generation Gap (JGG) scheme [9] and Adaptive Real-coded Ensemble Crossover (AREX) [10] have demonstrated high performance in exploring complex parameter spaces [11].

Although standard optimization strategies based on iterative solution methods including GAs can improve practical convergence, they remain computationally demanding. A major bottleneck arises because they require many fitness evaluations, each involving numerical integration of nonlinear ODE systems across all experimental conditions. This burden is particularly pronounced when time course data span long horizons or when multiple initial conditions must be considered. In highly multimodal landscapes, substantial computational effort is often wasted evaluating clearly inferior candidates, especially during the early stages of the search. The Goodwin oscillator illustrates typical nonlinear oscillatory dynamics in biochemical reaction networks: while it can exhibit stable sustained oscillations (limit cycles) over long periods, slight parameter variations may produce weakly damped oscillations with gradual amplitude decay (asymptotic stability) [12–16]. Accurately distinguishing such damped behavior from truly sustained oscillations generally requires long-horizon simulations and the detailed inspection of the resulting time course data, further increasing the computational cost. These challenges highlight the need for strategies that can efficiently eliminate parameter candidates prone to gradually decaying oscillatory behavior.

To address this limitation, we exploit the sequential structure of time course data to accelerate optimization. During numerical simulation, poor parameter candidates often diverge considerably from experimental observations at early time points, long before the full simulation horizon is reached. This suggests that full-horizon simulations are not always necessary to eliminate clearly inferior candidates. However, oscillatory biochemical reaction networks present a specific challenge: early simulation phases often reflect transient responses that differ from the steady oscillations observed experimentally. As a result, candidates that poorly match early data may still reproduce the correct long-term oscillatory behavior, while others that initially align well may fail to sustain oscillations. Therefore, naive early termination based only on initial errors risks discarding potentially valid solutions (false negatives), whereas standard full-horizon evaluations expend substantial resources on candidates that gradually decay.

In this study, we propose an efficient parameter estimation framework for oscillatory biochemical reaction networks that incorporates a novel adaptive simulation termination strategy. The framework combines three key components: (i) a real-coded GA for the global exploration of parameter space, (ii) a local optimization scheme based on the modified Powell method for rapid, gradient-free convergence, and (iii) an adaptive termination

mechanism that evaluates ODE simulations sequentially, discarding candidates exhibiting persistent discrepancies before completing full-horizon simulations. The rejection criterion is transient tolerant but progressively stricter, making it particularly well suited for estimating parameters that govern nonlinear oscillatory dynamics and for distinguishing sustained oscillations from weakly damped ones.

The proposed framework naturally accommodates multiple experimental conditions, such as varying initial concentrations or external perturbations, by defining the objective function as an aggregated, normalized error across all conditions and observables. Optimization is performed in a log-transformed parameter space to enforce positivity and stabilize the search across heterogeneous parameter scales. In the GA stage, a one-generation lag and a reference retention rule decouple boundary updates from candidate evaluation and prevent overly aggressive tightening when no valid offspring is produced. By reducing the average cost of objective function evaluation through adaptive termination, this approach provides a scalable and practical iterative solution for parameter estimation in computationally demanding oscillatory biochemical reaction networks.

In summary, to address the computational bottleneck of ODE-constrained inverse problems while preserving robustness in oscillatory dynamics, we make the following contribution. Our main novelty lies in a time-adaptive, quantile-based simulation termination rule with stabilization mechanisms (warm-up evaluation, a one-generation boundary lag, and reference retention), integrated into a real-coded GA. By explicitly tolerating early transient mismatches in oscillatory systems while terminating persistently poor candidates, the proposed method reduces the end-to-end computational burden (including overhead) without sacrificing practical estimation performance, and it generalizes to other ODE-constrained inverse problems where objective function evaluation dominates runtime.

2. Proposed Framework

2.1. Problem Formulation

We consider a biochemical reaction network consisting of n chemical species governed by p unknown kinetic parameters. Let $\boldsymbol{x}(t) \in \mathbb{R}^n$ denote the vector of species concentrations at time t, and let $\boldsymbol{\theta} \in \mathbb{R}^p$ denote the vector of kinetic parameters to be estimated. These parameters satisfy a component-wise positivity constraint:

$$\theta_i > 0 \quad (i = 1, \ldots, p). \tag{1}$$

Experiments are performed under multiple conditions indexed by $c = 1, \ldots, C$ (e.g., different initial concentrations or external perturbations). For each condition c, the system dynamics are given by

$$\begin{aligned} \frac{d\boldsymbol{x}(t)}{dt} &= f\big(\boldsymbol{x}(t); \boldsymbol{\theta}, \boldsymbol{u}_c\big), \\ \boldsymbol{x}(0) &= \boldsymbol{x}_0(\boldsymbol{u}_c). \end{aligned} \tag{2}$$

Here, $f(\cdot)$ represents the reaction kinetics, $\boldsymbol{u}_c$ denotes condition-specific inputs, and $\boldsymbol{x}_0(\boldsymbol{u}_c)$ is the corresponding initial state.

Let $y_{c,k}^{\mathrm{exp}}(t_{c,j})$ denote the experimental value of the k-th observable at time $t_{c,j}$ under condition c, and let $y_{c,k}^{\mathrm{sim}}(t_{c,j}; \boldsymbol{\theta})$ denote the corresponding simulated value obtained by numerically solving Equation (2). We define the objective function as the aggregated, normalized absolute relative error over all conditions, observables, and time points:

$$J(\boldsymbol{\theta}) = \frac{1}{C} \sum_{c=1}^{C} \left(\frac{1}{n^{\mathrm{exp}}} \sum_{k=1}^{n^{\mathrm{exp}}} \left(\frac{1}{T_c} \sum_{j=1}^{T_c} \left| \frac{y_{c,k}^{\mathrm{sim}}(t_{c,j}; \boldsymbol{\theta}) - y_{c,k}^{\mathrm{exp}}(t_{c,j})}{\max\big(\varepsilon, |y_{c,k}^{\mathrm{exp}}(t_{c,j})|\big)} \right| \right) \right). \tag{3}$$

Here, n^{exp} is the number of observables and T_c is the number of observation time points for condition c. Equation (3) is defined as a nested average: first over the sampled time points $j = 1, \ldots, T_c$ within each condition–observable pair (c, k), then over observables k, and finally over conditions c. Thus, the factor $1/T_c$ is explicitly tied to the inner summation over j for condition c, while $1/n^{\mathrm{exp}}$ and $1/C$ normalize the subsequent averages over observables and conditions, respectively. To avoid numerical instability when experimental observations are close to zero, we introduce a stabilization constant $\varepsilon > 0$ in the denominator via $\max\left(\varepsilon, |y^{\mathrm{exp}}_{c,k}(t_{c,j})|\right)$, thereby preventing excessively large contributions when $|y^{\mathrm{exp}}_{c,k}(t_{c,j})| \approx 0$. In this near-zero regime, the normalized error behaves similarly to an absolute error scaled by ε, providing a smooth and practical way to handle zero or near-zero measurements without excluding them. In our experiments, we set

$$\varepsilon = \eta \max_{c,k,j} |y^{\mathrm{exp}}_{c,k}(t_{c,j})|, \quad \eta = 10^{-3}. \tag{4}$$

This stabilizes the objective function while leaving the relative error behavior unchanged for typical observation magnitudes. In practice, the results are generally insensitive to the exact value of η as long as it is sufficiently small, because ε mainly affects the low-signal regime $|y^{\mathrm{exp}}_{c,k}(t_{c,j})| \lesssim \varepsilon$.

The parameter estimation problem is formulated as

$$\min_{\boldsymbol{\theta} \in \mathbb{R}^p} J(\boldsymbol{\theta}) \quad \text{subject to} \quad \theta_i > 0 \ (i = 1, \ldots, p). \tag{5}$$

Remark 1. *$J(\boldsymbol{\theta})$ is highly nonlinear, nonconvex, and typically multimodal owing to nonlinear reaction kinetics and strong parameter interactions. Moreover, each evaluation of $J(\boldsymbol{\theta})$ requires numerically solving Equation (2) for all experimental conditions, which can be computationally expensive.*

2.2. Log-Transformed Parameterization

To handle positivity, optimization is performed in log space. Introduce $\boldsymbol{z} \in \mathbb{R}^p$ via

$$z_i = \log \theta_i \quad (i = 1, \ldots, p). \tag{6}$$

This transformation enforces $\theta_i > 0$ by construction, thereby eliminating the need for explicit inequality constraints. The logarithmic parameterization (i) improves scaling when parameters span multiple orders of magnitude, (ii) enhances numerical stability during ODE simulations, and (iii) aligns naturally with relative error objectives such as Equation (3), where proportional deviations are meaningful. All evolutionary operations and local refinements are performed in $\boldsymbol{z}$-space, while model simulations use $\boldsymbol{\theta} = \exp(\boldsymbol{z})$.

2.3. Global Exploration Using a GA

Global exploration is performed using a real-coded GA in log space. Individuals represent $\boldsymbol{z}$, and offspring are generated by Adaptive Real-coded Ensemble Crossover (AREX). We employ the Just Generation Gap (JGG) scheme with a population $P_g = \{\boldsymbol{z}_m\}_{m=1}^{N}$ of size N at generation g [9–11].

Offspring generation (AREX): In each generation, λ offspring $\Omega_g = \{\boldsymbol{z}_s^{\mathrm{off}}\}_{s=1}^{\lambda}$ are generated in $\boldsymbol{z}$-space using AREX with μ parents selected from P_g. Each offspring is assigned a fitness value, evaluated as $J(\exp(\boldsymbol{z}))$ using either full-horizon simulation (baseline) or adaptive termination (proposed; Section 2.5).

Generational replacement (JGG with partial replacement): In JGG, only a subset of parents is replaced in each generation. The μ parents selected for AREX are treated as

replacement targets. After evaluating all offspring, we collect the subset of valid offspring $\Omega_g^{\text{valid}} \subseteq \Omega_g$ for which the simulation is completed over the full horizon, and define

$$r = \min\left(\mu, |\Omega_g^{\text{valid}}|\right). \tag{7}$$

We then perform partial replacement by: (i) sorting the target parents by fitness (worst first); (ii) sorting Ω_g^{valid} by fitness (best first); and (iii) replacing the worst r target parents with the best r valid offspring to form P_{g+1}.

2.4. Local Refinement Using the Modified Powell Method

After the GA stage terminates, local refinement is performed using the modified Powell method [17], a derivative-free optimizer that often achieves fast local convergence. To control the computational cost, Powell refinement is launched from only the top-ranked K GA individuals, balancing refinement quality and runtime. The modified Powell method is well-suited for this task because it has a proven track record in parameter estimation for biochemical reaction networks.

2.5. Partial Objective for Adaptive Simulation Termination

To reduce the computational cost, the objective function in Equation (3) is evaluated incrementally over observation indices. At index j, a partial objective $J^{(j)}(\boldsymbol{\theta})$ is computed using only observations up to $t_{c,j}$. For simplicity, we assume aligned observation indices within each model, i.e., $T_c = T$, and the index j is shared across conditions:

$$J^{(j)}(\boldsymbol{\theta}) = \frac{1}{C}\sum_{c=1}^{C}\left(\frac{1}{n^{\text{exp}}}\sum_{k=1}^{n^{\text{exp}}}\left(\frac{1}{j}\sum_{l=1}^{j}\left|\frac{y_{c,k}^{\text{sim}}(t_{c,l};\boldsymbol{\theta}) - y_{c,k}^{\text{exp}}(t_{c,l})}{\max\left(\varepsilon, |y_{c,k}^{\text{exp}}(t_{c,l})|\right)}\right|\right)\right). \tag{8}$$

Equation (8) is also defined as a nested average, but restricted to the first j sampled time points. In particular, the factor $1/j$ normalizes the inner summation over $l = 1, \ldots, j$, while $1/n^{\text{exp}}$ and $1/C$ average over observables and conditions, respectively.

The full objective function corresponds to $J(\boldsymbol{\theta}) = J^{(T)}(\boldsymbol{\theta})$. For clarity, the optimization target remains the full-horizon objective $J(\boldsymbol{\theta}) = J^{(T)}(\boldsymbol{\theta})$; the partial objectives $J^{(j)}(\boldsymbol{\theta})$ are used only as an evaluation shortcut for adaptive termination and do not redefine the problem being solved. As with any early termination rule, the adaptive simulation termination does not guarantee global optimality for $\min_{\boldsymbol{\theta}} J(\boldsymbol{\theta})$ and may, in rare cases, reject candidates whose fit improves only at later time points. The design in Section 2.6 mitigates this risk by using a transient-tolerant, time-adaptive boundary and a one-generation lag. During the numerical integration of Equation (2), the solver advances sequentially in time, and $J^{(j)}(\boldsymbol{\theta})$ is updated at each observation time $t_{c,j}$. If the partial objective exceeds the termination boundary at some index j^*, numerical integration is terminated, and the candidate is rejected. For candidates rejected at j^*, partial objectives at later indices $j > j^*$ are treated as undefined (missing) rather than being assigned artificially large values, preventing distortion of population-level statistics used for thresholding.

2.6. Quantile-Based Reference and Updating Strategy

2.6.1. Trade-Offs and Strategies for Oscillatory Systems

The update schedule for the termination reference creates a trade-off between computational savings and robustness. Frequent updates can tighten thresholds quickly and improve efficiency, whereas delayed updates reduce false negatives and help maintain population diversity (Table 1). This balance is especially important for oscillatory biochemical reaction networks, where early phases are dominated by transients. Even if candidates show large mismatches initially, they may still develop into sustained oscillations over time.

Table 1. Trade-offs in reference update timing for simulation termination.

Aspect	Early Update	Delayed Update
Computational efficiency	High	Moderate
Risk of false negatives	High	Low
Preservation of solution diversity	Low	High

Accordingly, we use a time-indexed, quantile-based reference designed for oscillatory dynamics. We define reference values for each observation index j and update the vector $\{Q_g^{(j)}\}_{j=1}^{T}$ (and thus $\{B_g^{(j)}\}_{j=1}^{T}$) once per generation from the post-replacement population. This allows for the termination boundary to adapt to the temporal evolution of the system.

- **Early Phases (Transient):** The population distribution is broad, i.e., partial fitness values have a wide range, yielding permissive thresholds that tolerate transient deviations.
- **Late Phases (Stabilized):** The population distribution becomes more concentrated, leading to stricter thresholds.

The thresholds are defined relative to the population distribution rather than based on a single best individual. Using quantile-based references preserves candidates that exhibit large transient deviations but eventually converge to sustained oscillations, while efficiently rejecting candidates that deviate persistently. Because the termination boundary acts as a screening mechanism, it introduces an efficiency–robustness trade-off. Overly aggressive screening could reduce exploration and may reject late-improving candidates. For this reason, the boundary is defined from population quantiles (rather than a single candidate) and is intentionally permissive at early indices, tightening only as τ_j increases. The warm-up evaluation and one-generation boundary lag further stabilize the screening pressure and reduce the likelihood of premature over-tightening in early generations.

Note on informativeness of the partial objective: we do not assume that very early errors reliably predict full-horizon performance, which is important to parameter estimation for oscillatory systems due to transient relaxation and phase alignment. Instead, $J^{(j)}(\boldsymbol{\theta})$ provides progressively more evidence as j increases, and persistent mismatch accumulates while short-lived transient deviations are diluted. This is why the boundary is intentionally permissive at small τ_j and tightens only toward later indices, so that termination decisions are driven primarily by sustained discrepancies rather than by early transients. In practical terms, this design is most effective when the observation horizon includes post-transient behavior (so late indices reflect the target oscillatory regime) and when clearly suboptimal candidates exhibit a mismatch that persists over a non-negligible portion of the horizon.

2.6.2. Population-Based Update with Generation Lag

We first perform a warm-up evaluation on the initial population P_0 without adaptive termination so that all individuals have full-horizon fitness values, ensuring that the initial quantiles are well defined. In each subsequent generation, we calculate quantile references for all indices $j = 1, \ldots, T$ from the post-replacement population:

$$Q_g^{(j)} = \text{Quantile}_q\left(\{J^{(j)}(\boldsymbol{\theta}) \mid \boldsymbol{\theta} \in P_g\}\right), \quad q \in (0, 1]. \tag{9}$$

The quantile level q controls screening pressure. Larger q yields a more permissive reference (higher $Q_g^{(j)}$ and thus a higher boundary), whereas smaller q yields a stricter reference, complementing the roles of α (global margin) and β (time-dependent tightening rate).

Note on quantile estimation under early termination: $Q_g^{(j)}$ in Equation (9) is defined as the q-quantile of partial objective values over the *post-replacement population* P_g, rather than as an unbiased estimate of the quantile of the offspring-generating distribution. In our

implementation, individuals in P_g have well-defined partial objectives $\{J^{(j)}\}_{j=1}^{T}$ because (i) the initial population is fully evaluated in the warm-up step ($g = 0$) and (ii) only offspring that survive the early termination test and complete full-horizon evaluation are inserted into the population. Early-rejected offspring are treated as *missing* for indices $j > j^*$ and are excluded from the computation of $Q_g^{(j)}$, to avoid any distortion of the quantile reference by censored evaluations. To prevent circular dependence, offspring in generation g are evaluated using the previous generation boundary (one-generation lag).

The termination boundary is

$$B_g^{(j)} = \alpha\, Q_g^{(j)} \exp(-\beta\tau_j), \quad \tau_j = t_{c,j}/t_{c,T} \in (0, 1], \tag{10}$$

where $\alpha > 1$ is a tolerance factor and $\beta > 0$ controls the rate at which the tolerance decreases.

Note on the termination boundary: $J^{(j)}(\boldsymbol{\theta})$ in Equation (8) is a running average of per-time-point normalized discrepancies (averaged over conditions and observables). As more observation points are accumulated, such running averages typically become less variable and more informative of long-horizon agreement, whereas very early indices can be dominated by transient relaxation and phase alignment in oscillatory systems. This motivates a boundary that is permissive early and progressively tighter later in the horizon. We implement this principle with a simple smooth monotone schedule $\exp(-\beta\tau_j)$, where β controls how quickly the tolerance decreases. The quantile scale $Q_g^{(j)}$ provides a population-relative, data-driven normalization at each index j, and $\alpha > 1$ acts as a global safety margin controlling the robustness–efficiency trade-off.

To avoid circular dependence within a generation, a one-generation lag is used: offspring in generation g are evaluated using $B_{g-1}^{(j)}$, and a candidate $\boldsymbol{\theta}$ is early rejected at index j if

$$J^{(j)}(\boldsymbol{\theta}) > B_{g-1}^{(j)}. \tag{11}$$

If no valid offspring is obtained (i.e., $|\Omega_g^{\text{valid}}| = 0$; hence, $r = 0$), no replacement occurs, and the references are retained:

$$Q_g^{(j)} = Q_{g-1}^{(j)} \quad \text{and} \quad B_g^{(j)} = B_{g-1}^{(j)}. \tag{12}$$

After offspring evaluation and replacement, the updated references $Q_g^{(j)}$ and $B_g^{(j)}$ are stored and used in generation $g + 1$. JGG partial replacement ensures that the population size remains N. Together with the warm-up evaluation at $g = 0$, the quantiles in Equation (9) remain well defined. For GA ranking and selection, rejected candidates are assigned a fixed penalized fitness M_{reject} and treated as invalid offspring, while quantile references are calculated from the post-replacement population (Equation (9)) and remain unaffected by the penalized values.

To provide intuition, Figure 1 visualizes partial fitness trajectories during evaluation. The green line represents the boundary $B_g^{(j)}$, which is designed to be permissive early and progressively stricter over time. The red trajectories correspond to candidates that deviate substantially and are terminated early, avoiding wasteful full-horizon simulations. The blue trajectories correspond to promising candidates that remain below the boundary. Importantly, the relaxed early threshold (small τ_j) acts as a buffer that tolerates transient errors inherent to oscillatory dynamics, reducing the risk of early termination for valid solutions (false negatives).

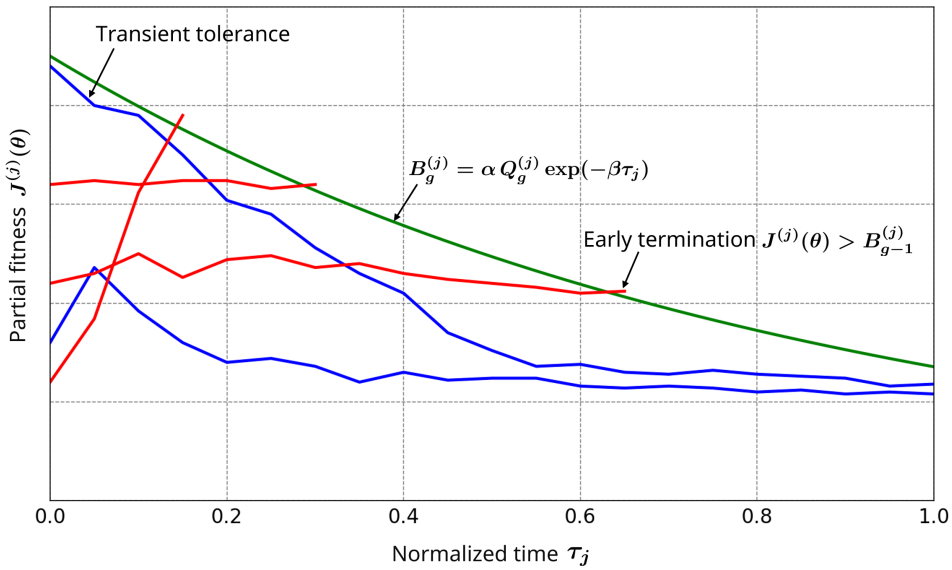

Figure 1. Schematic illustration of the adaptive simulation termination strategy. Valid candidates (blue lines) remain below the time-dependent termination boundary (green line), while inferior candidates (red lines) are terminated early.

3. Numerical Experiments

3.1. Test Models and Synthetic Data

We evaluate the proposed framework on two benchmark models: (a) an oscillatory mass action system (Lotka–Volterra predator–prey model) [18,19], and (b) an oscillatory biochemical feedback model (Goodwin oscillator) [12–16].

The Lotka–Volterra model serves as a canonical, fully observed oscillatory benchmark in theoretical population dynamics, whereas the Goodwin oscillator is a prototypical negative-feedback biochemical oscillator, widely used as a minimal model for gene-regulatory rhythms.

(a) Lotka–Volterra (fully observed)

Consider the reaction scheme $X \xrightarrow{k_1} 2X,\ X + Y \xrightarrow{k_2} 2Y,\ Y \xrightarrow{k_3} \varnothing$, yielding

$$\begin{aligned} \frac{dx_1}{dt} &= k_1 x_1 - k_2 x_1 x_2, \\ \frac{dx_2}{dt} &= k_2 x_1 x_2 - k_3 x_2. \end{aligned} \tag{13}$$

Here, $x_1(t)$ and $x_2(t)$ denote the prey and predator concentrations. The parameter vector is $\boldsymbol{\theta} = (k_1, k_2, k_3)$, with ground truth values $k_1^* = 1.0$, $k_2^* = 0.5$, $k_3^* = 1.0$. Two experimental conditions ($C = 2$) are used with initial states $(2.0, 1.0)$ and $(1.0, 2.0)$. All state variables are observed: $\boldsymbol{y}(t) = (x_1(t), x_2(t))$.

(b) Goodwin oscillator (partially observed)

We use the Goodwin model, as follows:

$$\begin{aligned} \frac{dx_1}{dt} &= k_1 \frac{1}{1 + x_3^n} - k_2 x_1, \\ \frac{dx_2}{dt} &= k_3 x_1 - k_4 x_2, \\ \frac{dx_3}{dt} &= k_5 x_2 - k_6 x_3, \end{aligned} \tag{14}$$

with the Hill coefficient fixed at $n = 10$ to ensure sustained oscillations [14,15]. The parameter vector is $\boldsymbol{\theta} = (k_1, k_2, k_3, k_4, k_5, k_6)$, with ground truth values $k_1^* = k_3^* = k_5^* = 1.0$

and $k_2^* = k_4^* = k_6^* = 0.5$. Two experimental conditions ($C = 2$) are used with initial states $(0.2, 0.5, 1.2)$ and $(0.3, 0.5, 1.0)$. Only the third state variable is observed: $y(t) = x_3(t)$. However, although only $x_3(t)$ is observed, we show that the framework remains capable of parameter estimation under this partial observability.

Synthetic data are generated by numerically solving the ODEs using the ground truth parameters and specified initial conditions. Noise-free observations are used to ensure a fair evaluation under identical settings. For both the Lotka–Volterra and Goodwin models, observations are sampled uniformly at 100 post-initial time points over $t \in (0, 50]$ (excluding the initial state at $t = 0$), resulting in a total of 200 observations across the two initial conditions.

3.2. Compared Frameworks and Settings

We compare two frameworks under identical parameter bounds, optimization budgets, and solver settings, with the only difference being whether adaptive termination is applied during evaluation.

- **Baseline (GA + Powell; full-horizon simulation):** A real-coded GA explores the log-transformed parameter space, followed by local refinement using the modified Powell method. Each candidate is evaluated with the full-horizon objective $J(\boldsymbol{\theta})$ (Equation (3)).
- **Proposed (GA + Powell; adaptive termination):** The same GA and Powell method settings are used, but evaluation follows the adaptive simulation termination strategy: the partial objective $J^{(j)}(\boldsymbol{\theta})$ (Equation (8)) is computed sequentially, and each simulation is terminated once Equation (11) is satisfied.

GA settings: The population size is $N = 50$; $\lambda = 30$ offspring are generated per generation using AREX with $\mu = 3$ parents. The GA runs for $G = 100$ generations with JGG partial replacement (Equation (7)). In both models, each kinetic parameter is constrained to lie between 0.1 and 10 times its ground truth value $\boldsymbol{\theta}^*$. All GA operations are performed in the log-transformed parameter space. The penalized fitness for rejected candidates is set to $M_{\text{reject}} = 10^6$.

Powell method settings: After the GA terminates, the modified Powell method is applied, starting from the top $K = 10$ GA individuals ranked by the full objective $J(\boldsymbol{\theta})$. Each Powell run stops when improvements fall below a specified tolerance or when a maximum number of objective evaluations is reached; these stopping criteria are the same for both frameworks. During Powell refinement, the adaptive termination boundary is held fixed at the final GA boundary $B_G^{(j)}$.

Adaptive termination settings: Unless otherwise stated, we use quantile level $q = 0.75$, tolerance factor $\alpha = 1.2$, and decay rate $\beta = 0.7$ in Equations (9)–(12) (Sections 3.6.5 and 4.2). The initial population P_0 is evaluated without adaptive termination to obtain $Q_0^{(j)}$ and $B_0^{(j)}$ (Equations (9) and (10)). Offspring in generation g are tested against $B_{g-1}^{(j)}$ with a one-generation lag (Equation (11)); if no valid offspring are obtained, the references are retained (Equation (12)).

3.3. Evaluation Metrics

ODE simulations use the same solver and tolerance settings across all frameworks. The computational cost is measured by the total number of right-hand-side (RHS) evaluations performed by the solver. For each model, we performed 100 independent trials with different random seeds. In each trial, we record the best fitness achieved after Powell refinement. To quantify the computational cost, we report the total number of RHS evaluations aggregated over the entire optimization process (GA + Powell). We also report the

relative reduction in the computational cost of the proposed framework compared with the baseline:

$$\Delta_{\text{cost}} = 1 - \frac{\text{Cost}_{\text{proposed}}}{\text{Cost}_{\text{baseline}}}, \tag{15}$$

where Cost denotes the total number of RHS evaluations.

To characterize the behavior of adaptive termination, we report the mean normalized termination index. For each candidate, let $j^* \in \{1, \ldots, T\}$ denote the last evaluated time index (with $j^* = T$ indicating that adaptive termination was not triggered). We define

$$\rho = \frac{j^*}{T} \in (0, 1], \tag{16}$$

and report its mean value $\bar{\rho}$, aggregated over all evaluations (GA + Powell) in each trial. Across the 100 trials, the results are summarized using the median and interquartile range (IQR) for both fitness and computational cost.

3.4. Implementation

All simulations and optimizations are implemented in Python (version 3.14.2), using NumPy (version 2.4.2) and SciPy (version 1.17.0). ODE systems are solved using the explicit Runge–Kutta integrator `RK45` in SciPy with fixed tolerance settings (rtol $= 10^{-6}$, atol $= 10^{-9}$). For each candidate and each condition c, the simulated observations $y^{\text{sim}}_{c,k}(t_{c,j}; \boldsymbol{\theta})$ are evaluated at the prescribed observation times $\{t_{c,j}\}_{j=1}^{T}$.

3.4.1. Time-Sequential Evaluation Without Restarting from $t = 0$

To implement adaptive termination, the solver advances sequentially between observation times, reusing the terminal state as the initial condition for the next interval. For each condition c, let $t_{c,0} = 0$ and $\boldsymbol{x}_c(t_{c,0}) = \boldsymbol{x}_0(\boldsymbol{u}_c)$. For $j = 1, \ldots, T$, we integrate Equation (2) over the interval $[t_{c,j-1}, t_{c,j}]$ and update the partial objective $J^{(j)}(\boldsymbol{\theta})$ (Equation (8)). Evaluation terminates as soon as $J^{(j)}(\boldsymbol{\theta}) > B^{(j)}_{g-1}$ (Equation (11)), and the index j^* is recorded as the termination index. For the baseline, the same sequential procedure is used, but integration always proceeds to $j = T$, i.e., adaptive termination is disabled.

3.4.2. Computational Cost Accounting

The computational cost is measured by the total number of RHS evaluations. For each solver call, the number of RHS evaluations is recorded, and these counts are aggregated over all candidate evaluations and experimental conditions. To prevent method-dependent step size growth over long simulation horizons, we enforce a common maximum step size equal to the uniform observation interval:

$$\Delta t_c = \frac{t_{c,T} - t_{c,0}}{T}, \tag{17}$$

ensuring that numerical effort is comparable across candidates.

3.5. Summary of Experimental Settings

In this section, we summarize the model configurations, data generation protocol, optimization hyperparameters, evaluation metrics, and solver settings used in the numerical experiments. Table 2 lists the experimental settings used in the numerical experiments. To make the provenance of the listed symbol values explicit, Table 2 is a consolidated summary of values defined in Sections 3.1–3.4: model-specific ground truth parameters, initial conditions, and observables are taken from Section 3.1; optimization and adaptive simulation termination hyperparameters are taken from Section 3.2; evaluation metrics are defined

in Section 3.3; and solver/implementation settings are taken from Section 3.4. Unless otherwise stated, these values were fixed a priori and shared across trials. Symbols are consistent with the main text and equation numbering (Equations (4), (6), (7), (9)–(12) and (17)). Notably, the Lotka–Volterra model is fully observed, whereas the Goodwin oscillator is partially observed (only x_3). The sampling grid consists of 100 post-initial time points over $t \in (0, 50]$ for each initial condition (excluding $t = 0$), yielding a total of 200 observations across the two initial conditions. The adaptive termination boundary is updated once per generation based on population quantiles, incorporating a warm-up, a one-generation lag, and a reference retention rule. During the Powell refinement stage, the boundary is held fixed at the value obtained at the final GA generation.

Table 2. Experimental settings used in the numerical experiments (GA/Powell/adaptive simulation termination/solver/data). The listed symbol values are summarized from Sections 3.1–3.4 (model definitions and observations, optimization/termination boundary settings, evaluation definitions, and solver/implementation settings) and, unless otherwise stated, were fixed a priori across trials.

Category	Name	Symbol/Value	Notes
Genetic algorithm	Population size	$N = 50$	Real-coded GA in log space (Equation (6))
	Offspring	$\lambda = 30$	Generated via AREX
	Parents	$\mu = 3$	JGG partial replacement (Equation (7))
	Generations	$G = 100$	Maximum generation
	Parameter bounds	$[0.1, 10] \times \boldsymbol{\theta}^*$	Kinetic parameter bounded relative to ground truth
Powell method	Function	SciPy minimize	Maximum number of iterations is set to 200
	Starts	$K = 10$	Initialized from top-ranked GA solutions
	Termination boundary	$B_G^{(j)}$	Fixed (no updates during Powell)
Adaptive termination	Quantile level	$q = 0.75$	For $Q_g^{(j)}$ (Equation (9))
	Tolerance factor	$\alpha = 1.2$	Controls the tolerance boundary (Equation (10))
	Decay rate	$\beta = 0.7$	Controls time-dependent strictness (Equation (10))
	Warm-up	—	P_0 evaluated without adaptive termination
	Boundary lag	1 generation	Evaluate at g using $B_{g-1}^{(j)}$ (Equation (11))
	Reference retention rule	—	If no valid offspring, retain $Q_g^{(j)}$, $B_g^{(j)}$ (Equation (12))
	Rejection penalty	$M_{\text{reject}} = 10^6$	For ranking; excluded from quantile computation
Objective	Stabilization constant	$\varepsilon = \eta \max_{c,k,j} \lvert y_{c,k}^{\exp}(t_{c,j}) \rvert$	Treatment of zero/near-zero measurements (Equation (4))
Solver	Integrator	SciPy RK45	Explicit Runge–Kutta; shared across frameworks
	Tolerances	rtol $= 10^{-6}$, atol $= 10^{-9}$	Fixed across runs
	Maximum step	$\Delta t_c = (t_{c,T} - t_{c,0})/T$	Equal to the sampling interval (Equation (17))
Data	Sampling grid	100 points over $t \in (0, 50]$	Per condition; excludes $t = 0$
	Observables (Lotka–Volterra)	(x_1, x_2)	Fully observed
	Observables (Goodwin)	x_3	Partially observed (single variable)

3.6. Results

Unless otherwise noted, all box plots summarize 100 independent trials per model. Boxes represent the IQR, center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean where shown.

3.6.1. Estimation Accuracy

Figure 2 shows the best fitness (objective value; lower is better) over 100 independent trials for (a) the Lotka–Volterra model and (b) the Goodwin model. Notably, the proposed framework achieves fitness values that are comparable to, and, in the case of

the Lotka–Volterra model, slightly more consistent than the baseline. The IQRs for the proposed framework are similarly tighter or narrower compared to the baseline, indicating that the adaptive termination strategy effectively filters out inferior candidates without discarding valid solutions that are critical for convergence. Although some outliers are observed in both frameworks—a natural consequence of the stochastic nature of GAs across a larger number of trials—the medians remain of the same order. This result confirms that the computational savings (Sections 3.6.2 and 3.6.3) are not achieved at the expense of estimation accuracy. Instead, the adaptive boundary successfully focuses computational resources on promising regions of the search space.

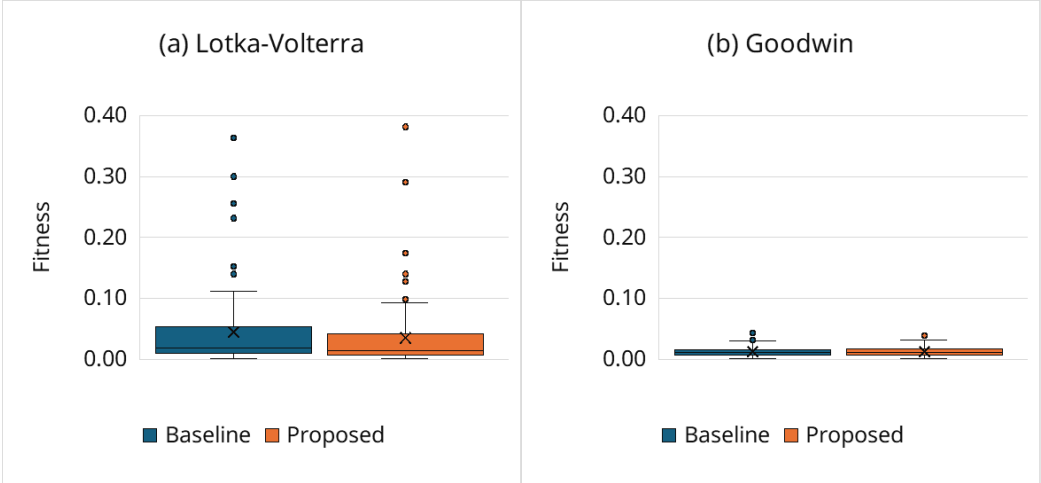

Figure 2. Best fitness (objective value; lower is better) over 100 independent trials for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model. Each panel compares the baseline framework (full-horizon evaluation) and the proposed framework (adaptive simulation termination). Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. The distributions indicate that the proposed framework preserves estimation accuracy while reducing the computational cost (see Figures 3–6).

Note on the Goodwin model: In Figure 2, the Goodwin model generally exhibits lower relative error than the Lotka–Volterra model; this is partly due to partial observability (only $x_3(t)$ is measured), which renders the objective effectively less constrained than when the full state vector is observed.

3.6.2. Computational Cost Reduction

Figure 3 shows the distribution of the computational cost, measured as RHS evaluations (Section 3.3), aggregated over the entire optimization process (GA + Powell). The proposed framework substantially reduces RHS evaluations compared with the baseline for both models. Quantitatively, for the Lotka–Volterra model, the median decreases from 17.8 million under the baseline to 11.2 million with the proposed framework. For the Goodwin model, the median decreases from 18.9 million to 10.2 million.

To complement the solver-reported computational cost in terms of RHS evaluations (Figure 3), we also report the elapsed wall clock time per trial, measured on a desktop PC (Intel Core i7-13700KF (Intel Corporation, Santa Clara, CA, USA), Windows 11). This timing includes the full algorithmic overhead, including GA operators, boundary evaluation/termination checks, and Powell refinement, in addition to ODE integration. As shown in Figure 4, the proposed framework consistently reduces wall clock time for both models, and the reduction closely tracks the decrease in RHS evaluations. This indicates that the overhead introduced by adaptive simulation termination is modest relative to the cost of numerical integration and does not offset the efficiency gains.

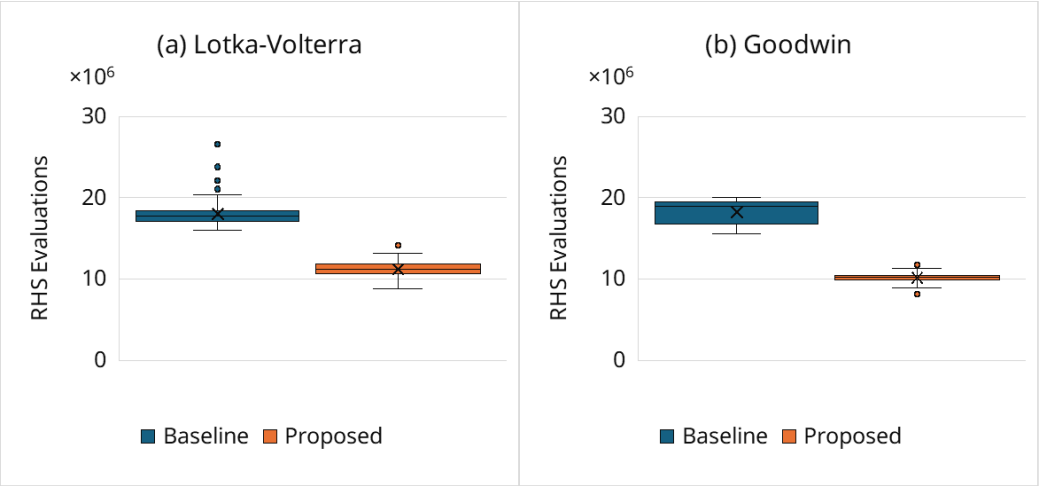

Figure 3. Total computational cost measured by the number of ODE right-hand-side (RHS) evaluations over the entire optimization process (GA + Powell) for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model, summarized over 100 independent trials. Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. The baseline and proposed frameworks use the same optimization budgets and solver settings; thus, differences reflect the effect of adaptive simulation termination. Lower values indicate lower computational cost. The proposed framework consistently reduces the total RHS evaluations in both models.

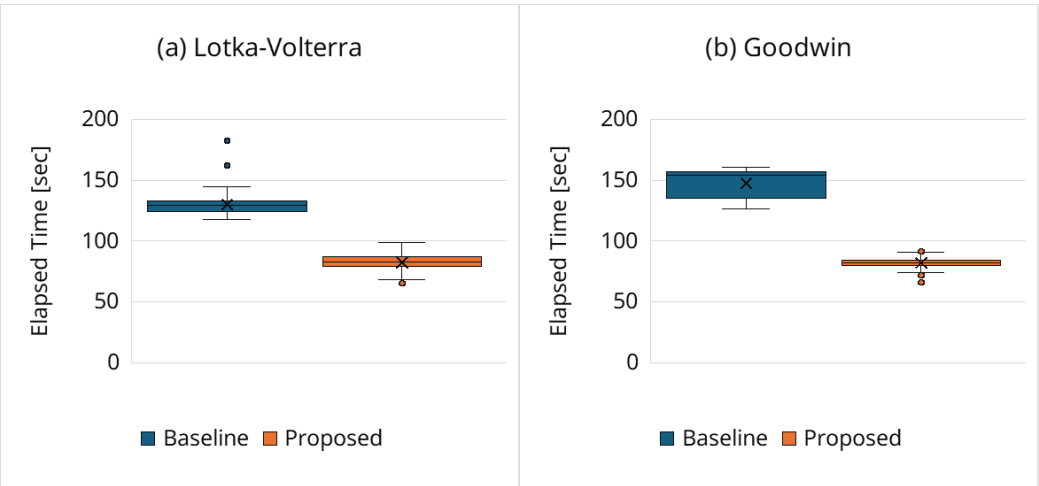

Figure 4. Elapsed wall clock time per trial for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model, summarized over 100 independent trials. Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. Timing was measured on a desktop PC (Intel Core i7-13700KF (Intel Corporation, Santa Clara, CA, USA), Windows 11) and includes all algorithmic overhead (GA operators, adaptive termination boundary evaluation/termination checks, and Powell refinement) in addition to ODE integration.

Figure 5 summarizes the relative cost reduction Δ_{cost} defined in Equation (15). The proposed framework achieves substantial savings, with reductions falling within the 30–50% range, reflecting markedly fewer full-horizon simulations of clearly suboptimal candidates.

The proposed framework consistently lowers the computational cost relative to the baseline without changing solver or optimization settings. This demonstrates that the early termination of clearly inferior candidates decreases the average evaluation burden. Notably, cost savings persist not only during the GA stage but also throughout Powell refinement because the same time-sequential evaluation routine is applied with a fixed boundary $B_G^{(j)}$ during Powell (Section 3.2).

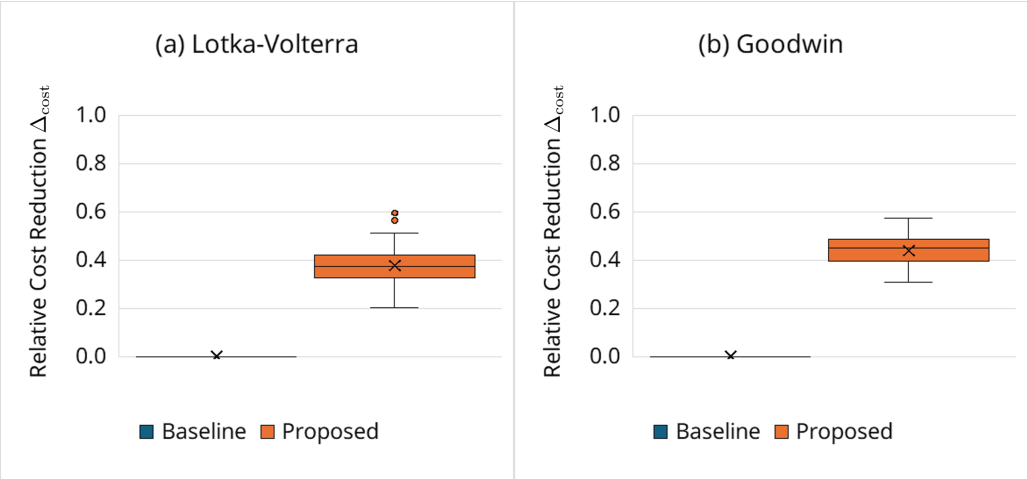

Figure 5. Trial-wise relative cost reduction Δ_{cost} for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model, summarized over 100 independent trials. Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. The baseline is shown as a horizontal reference line at $\Delta_{\text{cost}} = 0$ by definition, while the proposed framework is shown as a distribution of the achieved cost reductions relative to the corresponding baseline runs. Values in the proposed distributions falling in the range 0.3–0.5 indicate substantial computational savings due to adaptive simulation termination.

3.6.3. Behavior of Adaptive Termination via the Normalized Index

To directly assess how adaptive termination contributes to cost reduction, Figure 6 shows the mean normalized termination index $\bar{\rho}$ defined in Equation (16). Values substantially below one indicate that, on average, numerical integrations are stopped before reaching the end of the observation horizon. Across trials, $\bar{\rho}$ is consistently smaller for the proposed framework, confirming that the cost reduction mainly comes from truncated integrations rather than from incidental implementation overhead.

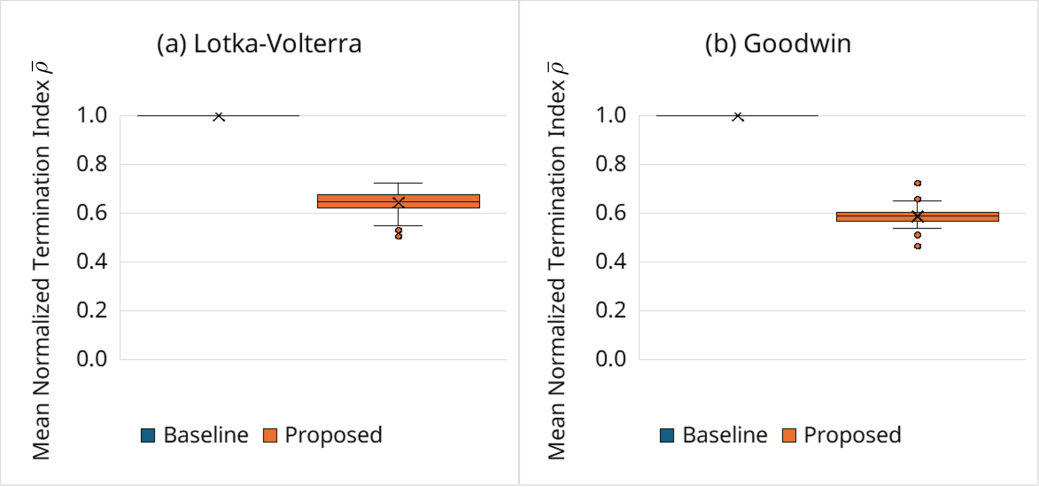

Figure 6. Mean normalized termination index $\bar{\rho}$ for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model, summarized over 100 independent trials. Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. $\bar{\rho} = 1$ (baseline reference) corresponds to full-horizon evaluation without adaptive termination, whereas values below one indicate that candidate evaluations are terminated before the final observation index on average. The proposed framework yields $\bar{\rho} < 1$ consistently, confirming that computational savings arise from the early truncation of inferior candidates.

Importantly, the proposed framework is designed to be permissive during early transient phases (Section 2.6). Therefore, early termination typically occurs only after persistent

mismatches have accumulated, rather than being triggered by short-lived transient deviations. As a result, $\bar{\rho}$ typically falls within the 0.5–0.7 range, which is particularly suitable for oscillatory biochemical reaction networks, where transient-dominated early phases may provide limited information about long-horizon fitness.

3.6.4. Evolution of Quantile References and Termination Boundaries

Figure 7 shows one result of the generation-wise evolution of the quantile references $Q_g^{(j)}$ and the corresponding termination boundaries $B_g^{(j)}$ as functions of the observation index $j = 1, \ldots, T$. Each curve represents a single GA generation, and the overlaid curves illustrate how the boundaries evolve throughout the optimization run.

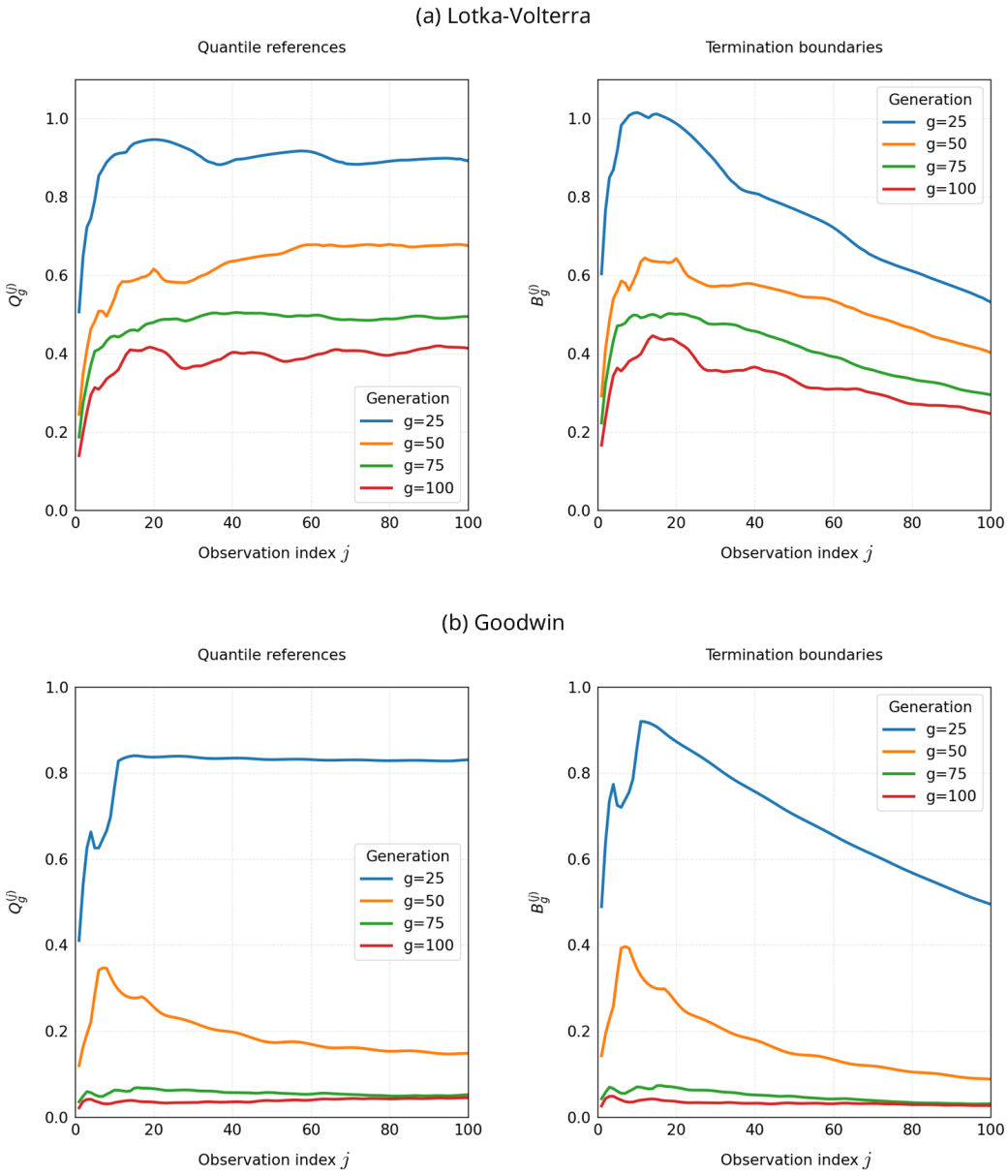

Figure 7. Generation-wise evolution of $Q_g^{(j)}$ (**left**) and the time-adaptive termination boundaries $B_g^{(j)}$ (**right**) as functions of the observation index j, for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model. Early indices are intentionally permissive, while later indices are stricter (Equation (10)). Across generations, the curves shift downward and stabilize, reflecting population-level improvement and threshold convergence. At small j, trajectories remain near the initial states, and relative errors are small, resulting in higher Q and B values that decrease as j increases. Curves overlay generations $g = 25$, 50, 75, and 100.

Two systematic patterns are observed. First, in each generation, the termination boundary is intentionally more permissive at early observation indices and becomes increasingly strict toward $j = T$, consistent with the explicit time dependence in Equation (10). This design buffers transient mismatches that are typical in oscillatory systems while enforcing tighter agreement near the end of the observation horizon, where sustained oscillations are expected to be established. Second, as generations progress, the overall level of $Q_g^{(j)}$ decreases, and the boundary curves shift downward accordingly, reflecting population-level improvement in partial objectives. In later generations, the curves tend to stabilize, suggesting that both the population distribution and the resulting thresholds have converged to a relatively steady regime.

This visualization also illustrates the stability mechanisms introduced in Section 2.6: the warm-up evaluation at $g = 0$ provides well-defined initial references; the one-generation lag prevents circular dependence between threshold estimation and offspring evaluation (Equation (11)); and the reference retention rule ensures well-posed behavior even in generations with no valid offspring (Equation (12)). Overall, the observed boundary dynamics offer an intuitive explanation for why adaptive termination achieves substantial cost reduction while maintaining competitive final objective values.

Finally, for small indices j, the simulated trajectories remain close to their initial states, so relative errors are typically small (i.e., many candidates appear spuriously fit). This early effect naturally results in higher $Q_g^{(j)}$ and $B_g^{(j)}$ values at small j; both quantities then decline as trajectories move away from the initial conditions and the time-adaptive boundary tightens. Importantly, this permissive early-time regime also helps maintain population diversity by avoiding the premature elimination of candidates that initially appear as similar but later diverge in performance, thereby preserving exploratory breadth before the boundary becomes stricter.

3.6.5. Sensitivity Analysis of Hyperparameters

We performed a grid sensitivity analysis over (α and β) (six combinations) with 100 independent trials per setting for each model, using the default $q = 0.75$. We assess the sensitivity of the adaptive termination behavior to the tolerance factor $\alpha \in \{1.2, 1.5\}$ and the decay rate $\beta \in \{0.5, 0.7, 0.9\}$. Figure 8 summarizes the distributions of the mean normalized termination index $\bar{\rho}$ (Equation (16)) over 100 trials per setting, serving as a proxy for the average integration length and, hence, the computational cost.

Overall, the qualitative trends are consistent across both models: increasing β (faster tightening) modestly reduces $\bar{\rho}$ (earlier termination), whereas increasing α (more permissive initial scaling) slightly increases $\bar{\rho}$ (later termination). Crucially, all settings lie within a narrow operational band (approximately $\bar{\rho} \approx 0.55$–0.73 across models), with tight IQRs, indicating stable behavior without fine-tuning. These trends are consistent with the cost reductions reported in Section 3.6.2 and confirm that the efficiency gain is an intrinsic property of the time-adaptive boundary rather than a consequence of sensitive hyperparameter tuning.

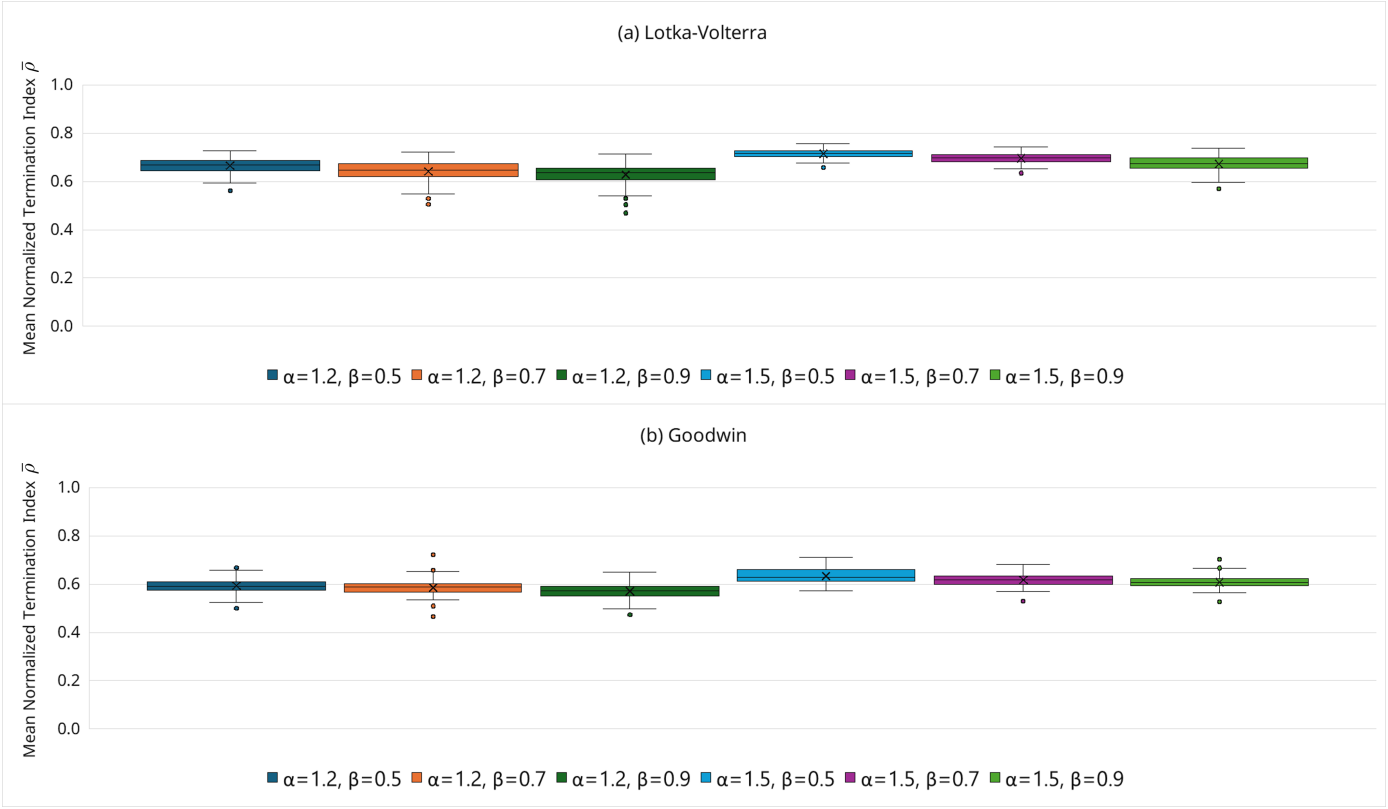

Figure 8. Sensitivity of the proposed framework to the adaptive termination hyperparameters (α and β), shown via the distribution of the mean normalized termination index $\bar{\rho}$ over 100 independent trials for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model. Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. The six box plots in each panel correspond to $\alpha \in \{1.2, 1.5\}$ and $\beta \in \{0.5, 0.7, 0.9\}$ (see legend). Lower $\bar{\rho}$ indicates earlier average termination (stronger screening), whereas higher $\bar{\rho}$ indicates later termination (more permissive screening). The narrow spread across settings indicates stable behavior without strong sensitivity to hyperparameter tuning.

3.6.6. Robustness Under Noisy Observations

To assess the robustness to observation noise, we additionally performed experiments using noisy synthetic observations for both the Lotka–Volterra and Goodwin models. In each trial, a new noisy dataset was generated by adding Gaussian perturbations to the clean synthetic observations (Section 3.1), and the same noisy dataset was shared by the baseline and proposed frameworks to ensure paired comparison. We considered three noise levels, $r = 0$ (noise-free), $r = 0.05$, and $r = 0.10$. The observable-wise noise standard deviation was defined as

$$\sigma_k = r \cdot \max_{c,j} |y_{c,k}^{\exp}(t_{c,j})|, \tag{18}$$

and the noisy observations (synthetic dataset) were generated as

$$\begin{aligned} y_{c,k}^{\text{noisy}}(t_{c,j}) &= y_{c,k}^{\exp}(t_{c,j}) + \xi_{c,k,j}, \\ \xi_{c,k,j} &\sim \mathcal{N}(0, \sigma_k^2). \end{aligned} \tag{19}$$

We performed 100 independent trials for each noise level. Moreover, we used the quantile level $q = 0.75$, tolerance factor $\alpha = 1.2$, and decay rate $\beta = 0.7$ in Equations (9) and (10) (Sections 3.2 and 4.2).

Figure 9 summarizes the effect of noise on the best fitness (objective value; lower is better). As expected, increasing noise worsened the fitness values for both methods

because the optimization criterion was evaluated directly on noise-contaminated observations. For the Lotka–Volterra model, the median fitness (baseline/proposed) changed from 0.0187/0.0144 (at $r = 0$) to 0.0923/0.0900 at $r = 0.05$, and to 0.229/0.219 at $r = 0.10$. For the Goodwin model, the corresponding medians changed from 0.0119/0.0116 to 0.0489/0.0493, and to 0.0967/0.102, respectively. The Goodwin model shows a smaller absolute fitness scale than the Lotka–Volterra model, which is partly attributable to partial observability (only x_3 is observed); accordingly, the y-axis ranges in Figure 9a,b are shown separately for visibility.

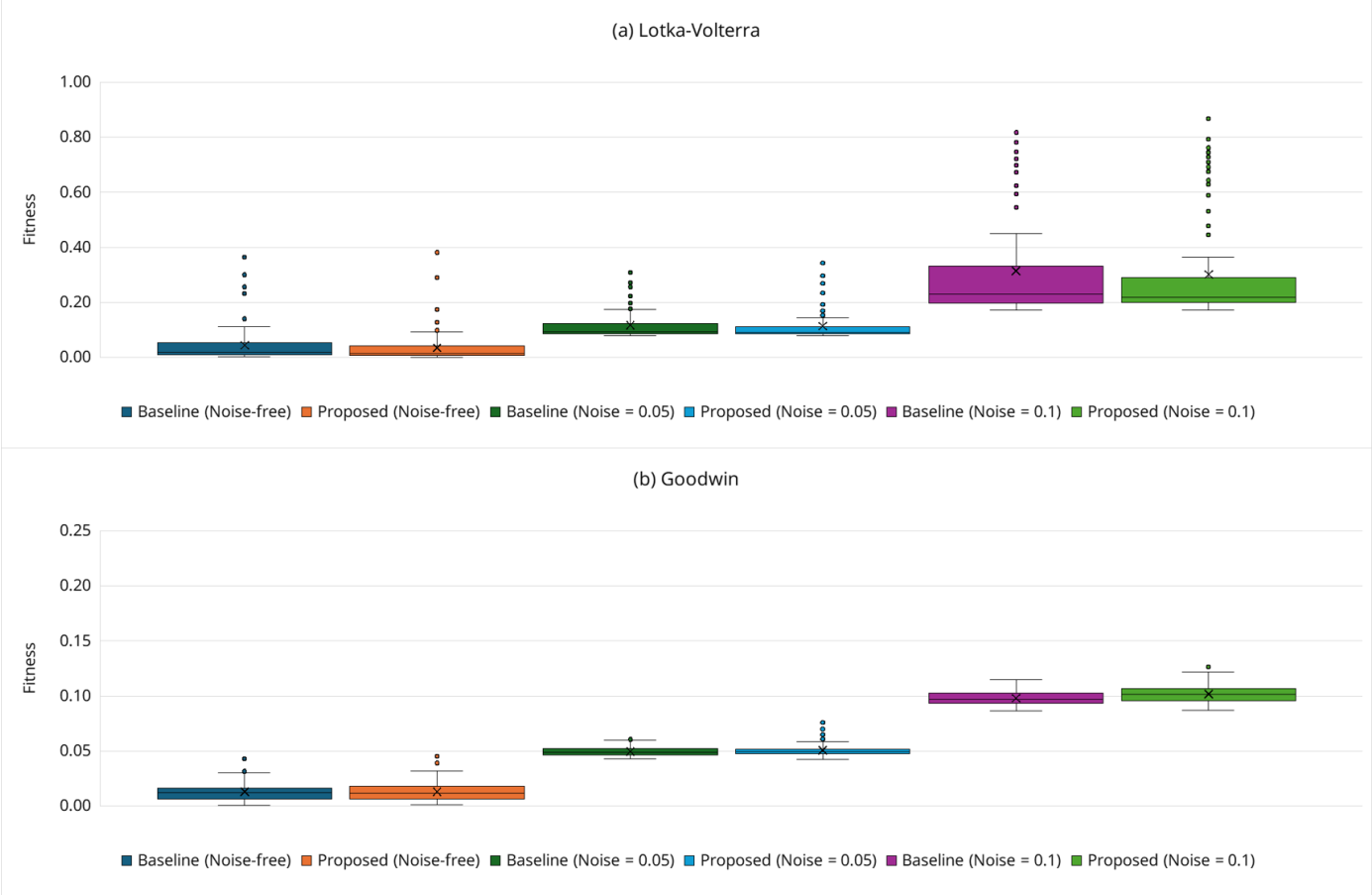

Figure 9. Best fitness (objective value; lower is better) under noisy observations for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model (100 trials per noise level). Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. Box plots compare the baseline and proposed frameworks under $r = 0$ (noise-free), $r = 0.05$, and $r = 0.10$ Gaussian noise. The y-axis ranges differ between panels because the absolute objective value scales differ between the two models.

Figures 10 and 11 show the computational cost and the relative cost reduction behavior under noise. The proposed framework consistently reduced the total number of RHS evaluations (Section 3.3) across all tested noise levels. For the Lotka–Volterra model, the median RHS evaluations decreased from 17.8 million (baseline) to 11.2 million (proposed) at $r = 0$, from 17.6 million to 11.2 million at $r = 0.05$, and from 18.0 million to 10.8 million at $r = 0.10$. For the Goodwin model, the corresponding median reductions were 18.9 million to 10.2 million, 19.1 million to 9.80 million, and 18.9 million to 9.70 million, respectively. The trial-wise median relative cost reduction Δ_{cost} (Equation (15)) also remained substantial. For the Lotka–Volterra model, 0.374, 0.372, and 0.405 at $r = 0$, 0.05, and 0.10, respectively; for the Goodwin model, 0.451, 0.475, and 0.464, respectively. These results indicate that

the acceleration effect of the adaptive termination mechanism is retained under moderate observation noise.

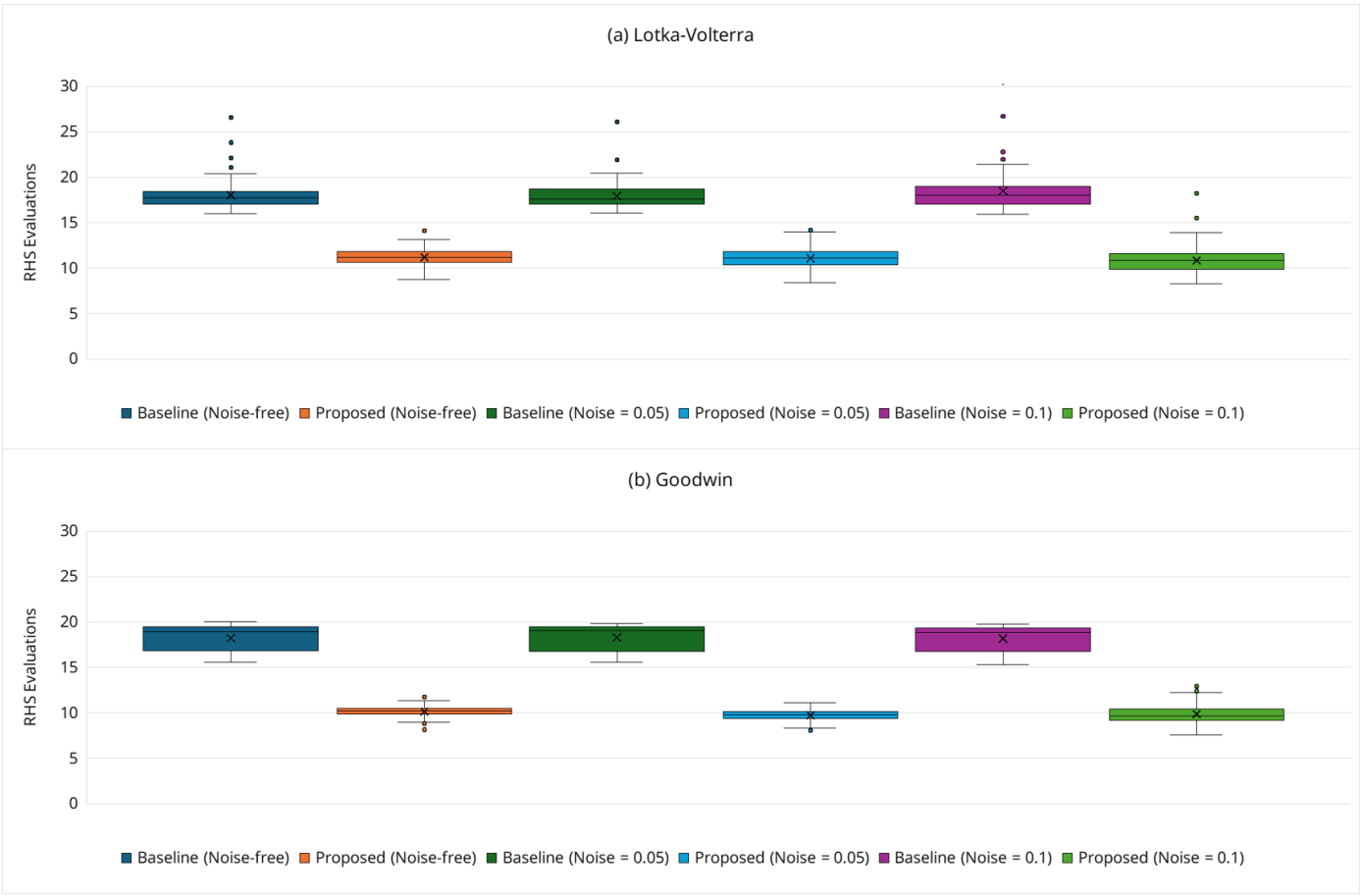

Figure 10. Total computational cost measured by ODE right-hand-side (RHS) evaluations under noisy observations for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model (100 trials per noise level). Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. The baseline and proposed frameworks are compared under $r = 0$, $r = 0.05$, and $r = 0.10$ using paired noisy datasets within each trial. Lower values indicate lower computational cost.

Figure 12 shows the mean normalized termination index $\bar{\rho}$ under noisy observations. For the proposed framework, the median $\bar{\rho}$ values for the Lotka–Volterra model were 0.646, 0.644, and 0.611 at $r = 0$, 0.05, and 0.10, respectively. For the Goodwin model, the corresponding medians were 0.587, 0.563, and 0.559 at $r = 0$, 0.05, and 0.10, respectively. Thus, even under noisy observations, the adaptive termination mechanism continued to terminate many evaluations of poor parameter candidates before reaching the end of the observation horizon, while preserving a competitive estimation performance.

Overall, these results support the practical robustness of the proposed framework as a computational cost reduction mechanism under noisy observations. Under the tested noise levels ($r \leq 0.10$), the observed cost reductions remained within a similar range to the noise-free setting (roughly 30–50%) without retuning the default hyperparameters, suggesting that substantial efficiency gains can be retained without widening the termination envelope in these cases.

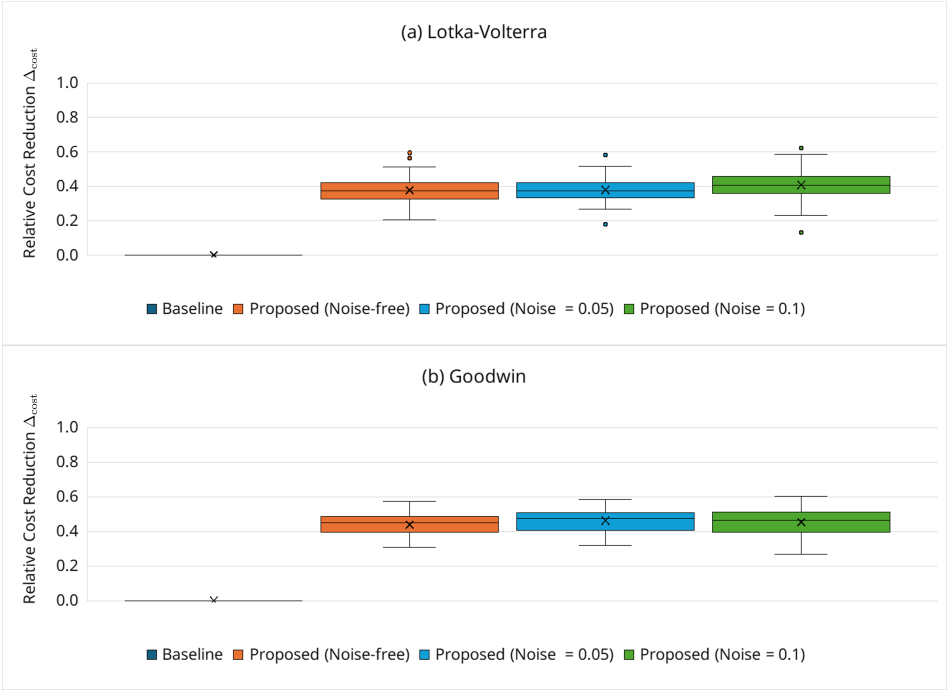

Figure 11. Trial-wise relative cost reduction Δ_{cost} for (**a**) Lotka–Volterra and (**b**) Goodwin models (100 trials per noise level). Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "$\times$" marks the mean. The baseline is shown as a horizontal reference line at $\Delta_{\text{cost}} = 0$ by definition, while the distributions summarize the trial-wise cost reduction under $r = 0$, 0.05, and 0.10. Substantial cost savings are retained under moderate noise.

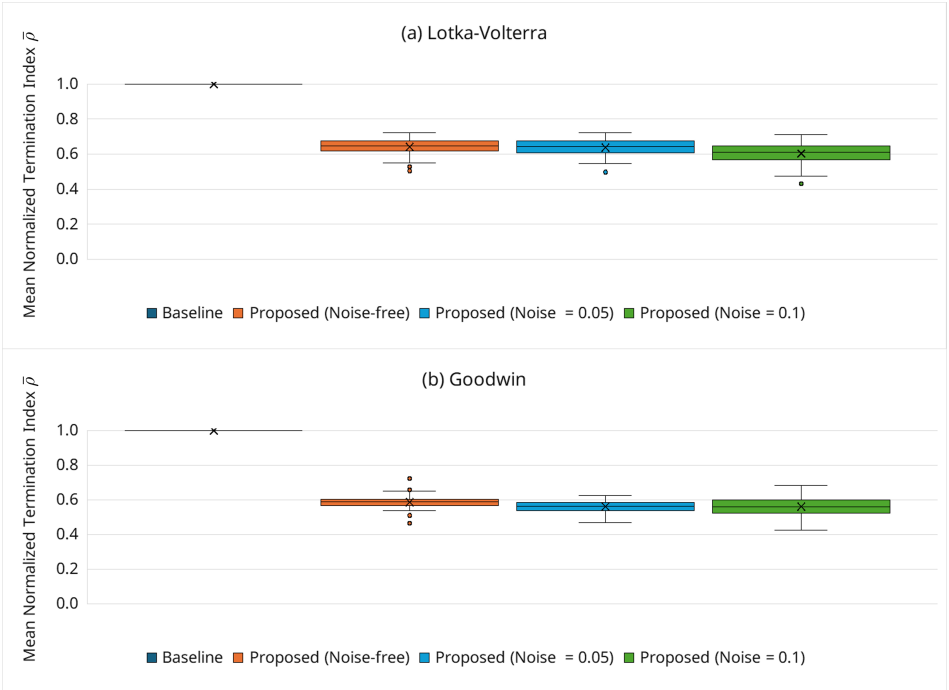

Figure 12. Mean normalized termination index $\bar{\rho}$ under noisy observations for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model (100 trials per noise level). Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "$\times$" marks the mean. The baseline corresponds to full-horizon evaluation and appears as a horizontal reference line at $\bar{\rho} = 1$. Values below one indicate that candidate evaluations are terminated before the full observation horizon on average, indicating that adaptive termination remains active under $r = 0$, 0.05, and 0.10.

3.6.7. Effect of Broader Search Ranges

To examine the behavior under less informative parameter bounds, we additionally performed experiments with a broader search range in which each kinetic parameter was bounded to $[10^{-2}, 10^{2}] \times \theta^{*}$. We performed 100 independent trials for each model using the same settings as in the main experiments (Section 3.2). In this wide-range setting, we set the ODE solver maximum step size to a conservative value (`max_step` $= 0.1$) to ensure stable integration across candidate trajectories.

Figure 13 summarizes the best fitness (objective value; lower is better) under broader search ranges ($[10^{-2}, 10^{2}] \times \theta^{*}$). While the proposed framework remains competitive in terms of its final objective values, the wider bounds increase variability because the search explores regions that can induce qualitatively different dynamics.

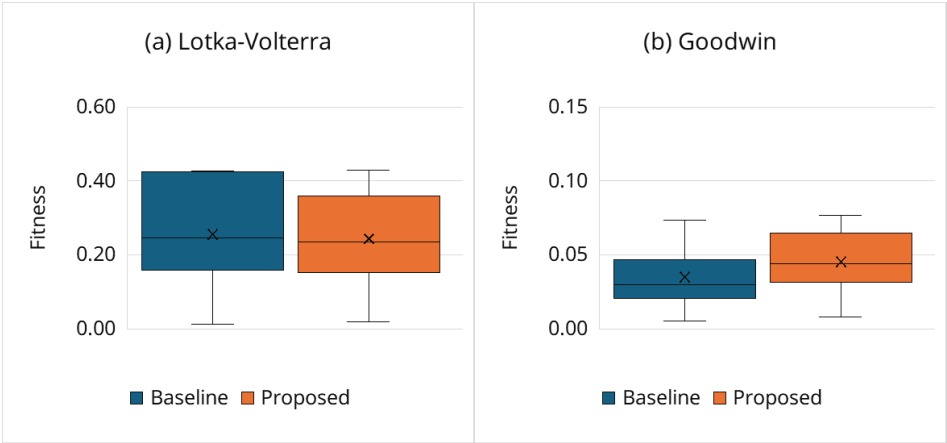

Figure 13. Best fitness (objective value; lower is better) under broader search ranges ($[10^{-2}, 10^{2}] \times \theta^{*}$) for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model (100 trials). Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. The y-axis ranges differ between panels because the absolute objective value scales differ between the two models.

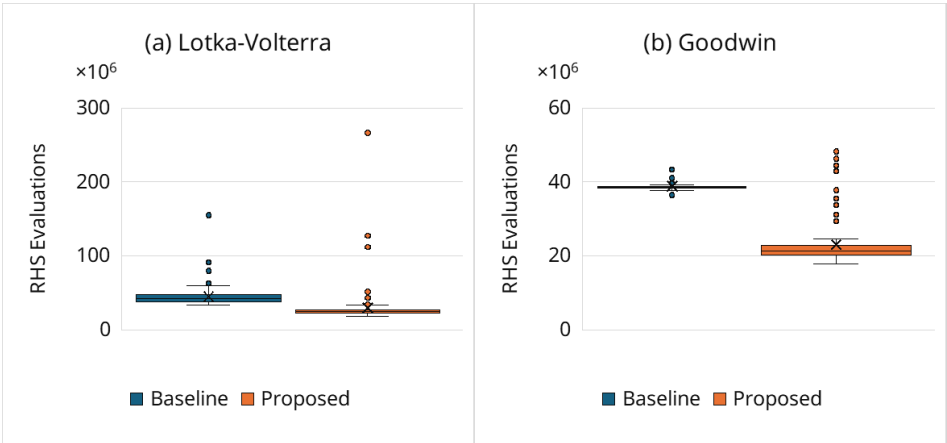

Figure 14. Total computational cost measured by ODE right-hand-side (RHS) evaluations under broader search ranges ($[10^{-2}, 10^{2}] \times \theta^{*}$) for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model (100 trials). Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. Wider bounds can induce numerically challenging trajectories, leading to heavier-tailed cost distributions. Lower values indicate a lower computational cost.

Figures 14–16 show the corresponding computational behavior. The proposed framework typically reduces the computational cost by terminating persistently poor candidates

early; however, when the search range is substantially widened, a small fraction of candidates can drive the ODE solver into numerically difficult regimes (e.g., stiff or near-stiff behavior). In such cases, the RHS evaluation count can increase markedly, producing heavy-tailed cost distributions and occasionally negative values of Δ_{cost} (Figure 15), meaning that the proposed run is more expensive than the baseline for those trials. Overall, these results suggest that broad parameter ranges are feasible but may require more conservative solver settings and/or additional safeguards for numerically challenging candidates, especially when stiff dynamics can arise (Section 4.2).

We note that negative values of Δ_{cost} in Figure 15 do not indicate a failure of the termination mechanism itself; rather, they occur when the widened search domain includes parameter regions that trigger numerically difficult trajectories (e.g., stiff or near-stiff behavior), for which the solver requires substantially more steps in the proposed run than in the corresponding baseline trial.

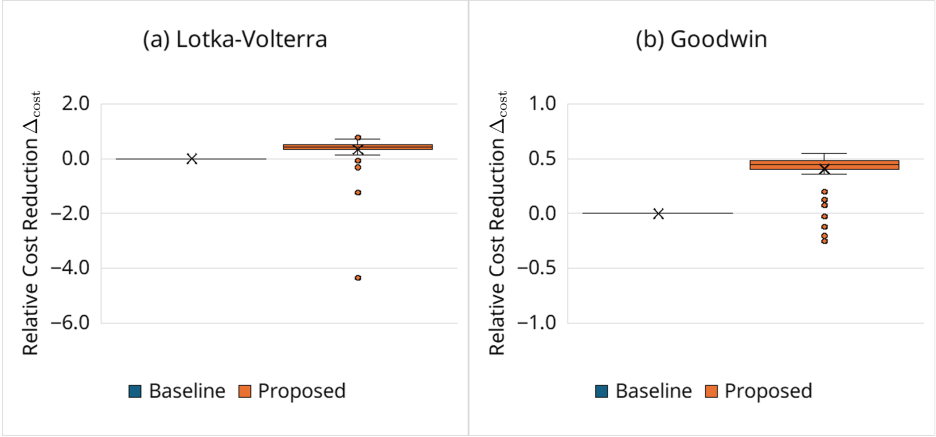

Figure 15. Relative cost reduction Δ_{cost} (Equation (15)) under broader search ranges ($[10^{-2}, 10^{2}] \times \theta^{*}$) for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model (100 trials). Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. The baseline corresponds to $\Delta_{\text{cost}} = 0$ by definition and is shown as a horizontal reference line. Negative values indicate trials in which the proposed framework required more RHS evaluations than the baseline (e.g., due to numerically difficult candidate trajectories).

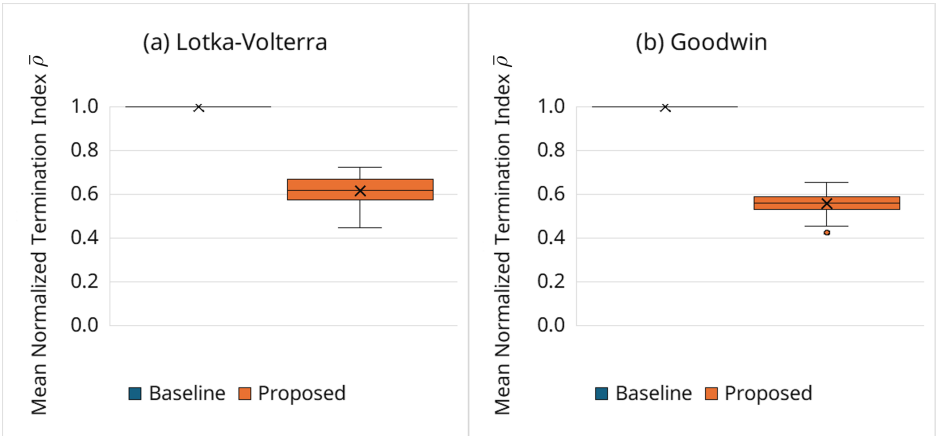

Figure 16. Mean normalized termination index $\bar{\rho}$ (Equation (16)) under broader search ranges ($[10^{-2}, 10^{2}] \times \theta^{*}$) for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model (100 trials). Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. The baseline corresponds to full-horizon evaluation and appears as a horizontal reference line at $\bar{\rho} = 1$. Values below one indicate that candidate evaluations are terminated before the full observation horizon on average.

3.6.8. Effect of a Grace Period

To explicitly investigate the effect of a "grace period" during early simulation phases—a period where early termination is disabled to allow trajectories to relax onto their attractors—we performed additional experiments. In this experiment, early termination was disabled for the initial portion of the observation horizon, and the termination criterion was applied only after a prescribed normalized observation index threshold. We compared three settings: no grace period, grace period $= 0.1$ (no termination for the first 10% of observation indices), and grace period $= 0.2$ (no termination for the first 20% of observation indices). We performed 100 independent trials for each model using the same settings as in the main experiments (Section 3.2).

Figures 17–20 summarize the effects of the grace period on fitness, computational cost, and termination behavior. Overall, introducing a grace period did not improve the cost–accuracy trade-off under the tested settings. For the Lotka–Volterra model, the median total RHS evaluations increased from 11.2 million (no grace period) to 13.5 million (grace period $= 0.1$) and 15.0 million (grace period $= 0.2$); for the Goodwin model, the corresponding medians increased from 10.2 million to 11.8 million and 12.8 million, respectively. Consistently, Δ_{cost} decreased (Lotka–Volterra: $0.37 \rightarrow 0.25 \rightarrow 0.15$; Goodwin: $0.45 \rightarrow 0.37 \rightarrow 0.32$), while the median-normalized termination index $\bar{\rho}$ increased (Lotka–Volterra: $0.65 \rightarrow 0.70 \rightarrow 0.72$; Goodwin: $0.58 \rightarrow 0.64 \rightarrow 0.68$), indicating that a longer grace period shifts termination to later indices and lengthens candidate evaluations.

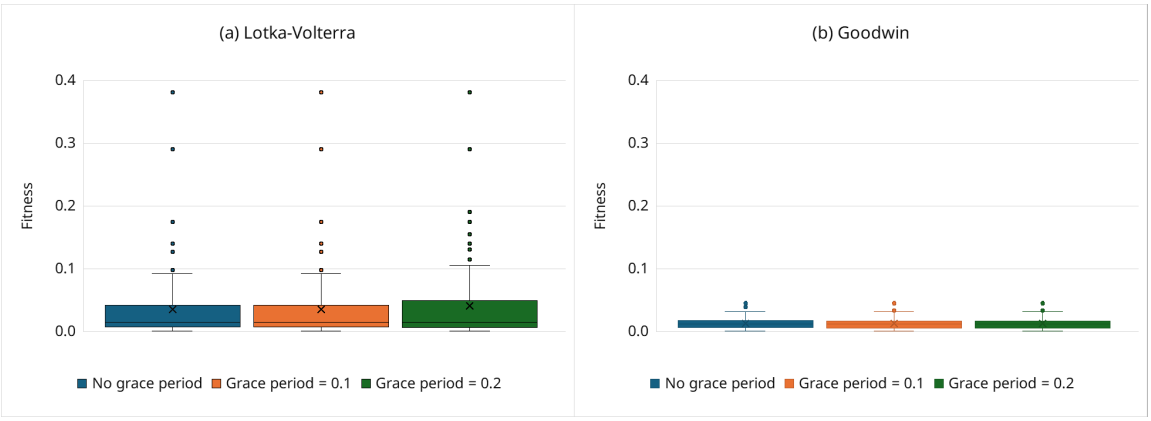

Figure 17. Best fitness (objective value; lower is better) under the grace period for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model over 100 independent trials. Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. Box plots compare three settings: no grace period, grace period $= 0.1$, and grace period $= 0.2$. Introducing a grace period does not significantly change final estimation performance under the tested settings.

The grace period shifts termination to later indices and lengthens candidate evaluations. This cost increase occurs because clearly inferior candidates that would otherwise be correctly rejected very early are forced to simulate until the grace period expires. These findings confirm that the standard proposed framework already accommodates early transients effectively. The exponential decay term in the termination boundary (Equation (10)) naturally provides a highly permissive, continuous envelope during the early phases, functioning as an effective built-in transient tolerance mechanism without forcing unnecessary computations for inherently poor candidates. Nevertheless, a grace period may be useful in problem settings with more severe transient mismatch or stronger observation noise, and could be considered as an optional variant.

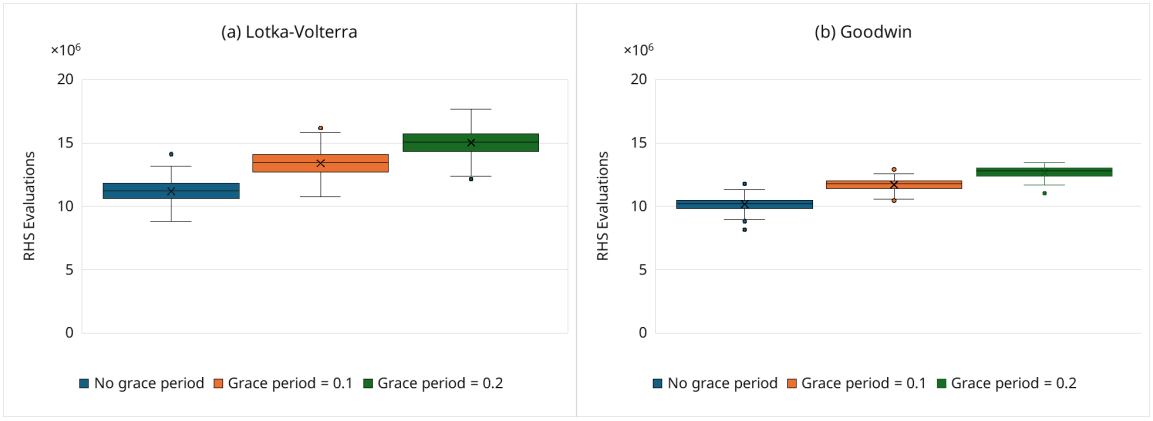

Figure 18. Total computational cost measured by ODE right-hand-side (RHS) evaluations under the grace period, for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model over 100 independent trials. Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. Introducing a grace period delays early termination and generally increases the average evaluation cost.

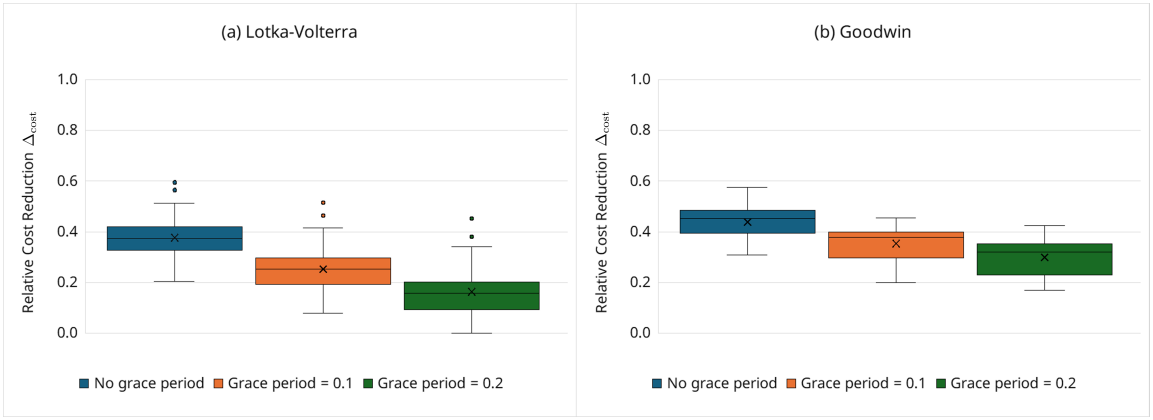

Figure 19. Trial-wise relative cost reduction Δ_{cost} under the grace period, for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model over 100 independent trials. Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. Larger grace periods delay early termination and tend to reduce the efficiency gain of adaptive termination by increasing the average evaluation length.

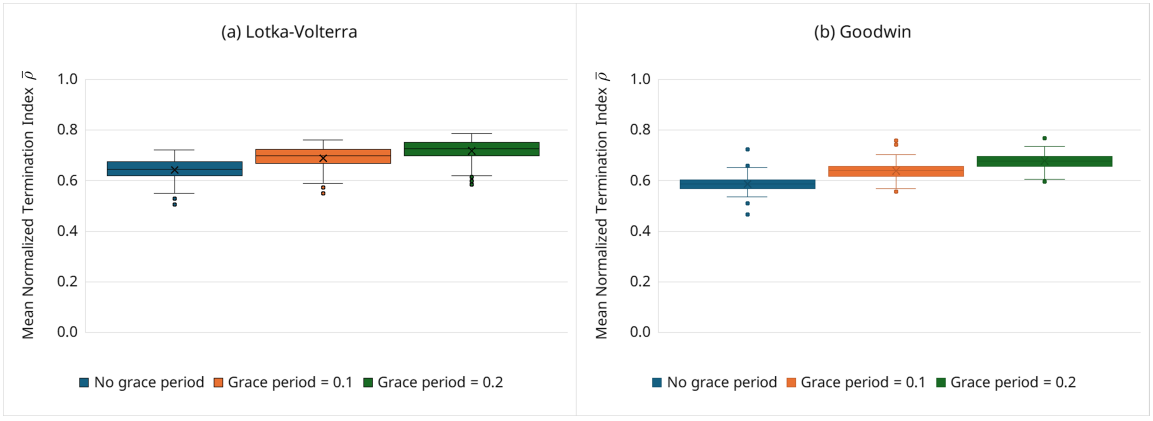

Figure 20. Mean normalized termination index $\bar{\rho}$ under the grace period, for (**a**) the Lotka–Volterra model and (**b**) the Goodwin model over 100 independent trials. Boxes represent the interquartile range (IQR), center lines indicate the median, whiskers show the full minimum–maximum range, and "×" marks the mean. Larger $\bar{\rho}$ indicates later termination and longer average candidate evaluations, consistent with the intended effect of grace period.

3.6.9. Ability to Distinguish Damped Oscillations

The Goodwin oscillator is a canonical example of oscillatory dynamics in biochemical reaction networks. Although it can exhibit stable, sustained oscillations over long horizons, small changes in parameters may cause trajectories to decay gradually toward a steady state, producing weakly damped oscillations. Accurately distinguishing such damped behavior from truly sustained oscillations typically requires long-horizon simulations and the careful inspection of the resulting time course data, making naive evaluation computationally expensive. This underscores the need for strategies that efficiently eliminate candidates trending toward damped behavior. The Goodwin model is defined in Equation (14); sustained oscillations arise for Hill coefficients $n > 8$ [14,15]. Figure 21 shows trajectories up to $T = 200$ with kinetic parameters $k_1^* = k_3^* = k_5^* = 1.0$ and $k_2^* = k_4^* = k_6^* = 0.1$, initial state $(x_1, x_2, x_3) = (0.01, 0.25, 2.5)$, and $n \in \{1, 5, 8, 10\}$. The case $n = 10$ exhibits sustained oscillations, whereas cases $n = 1, 5$, and 8 produce damped oscillations; notably, $n = 8$ shows gradual amplitude decay that is difficult to detect without long-horizon evaluation. To demonstrate the framework's ability to discriminate damped oscillations, we performed parameter estimation with the Hill coefficient treated as an integer decision variable, $n \in \{1, \ldots, 12\}$. Synthetic observations were generated from the ground truth setting $n = 10$ (Figure 21d) with a horizon $T = 200$ and 400 time points, and all three state variables were observed. Because the range of possible values for n is large, the population size is $N = 100$ and $\lambda = 30$ offspring are generated per generation using AREX with $\mu = 5$ parents. Across 100 independent trials, we recorded (i) whether a candidate was early terminated, (ii) the normalized termination index $\rho = j^*/T$ (Equation (17)) among terminated candidates, and (iii) the Hill coefficient of the best final solution in each trial.

Early termination rate by *n*: Figure 22a shows the early termination rate, which is 56.1% total. This demonstrates the proposed framework's ability to distinguish damped oscillations. Not all such candidates are eliminated immediately; instead, they are gradually replaced across GA generations through the JGG partial replacement scheme, which reinjects fitter offspring. A non-zero termination rate is also observed for $n > 8$, indicating that even when n satisfies the oscillatory condition, mismatched kinetic parameters can still produce decay or large discrepancies that are filtered out by the adaptive boundary.

When termination occurs: Figure 22b reports the mean ρ among the terminated candidates for each n. For $n \leq 8$, termination occurs earlier in the horizon (small ρ); for $n = 7$ and 8, termination tends to occur later (larger ρ), consistent with gradually decaying trajectories that initially appear plausible but only reveal decay in the second half of the observation window. This behavior aligns with the time-adaptive boundary, which is permissive early on and becomes stricter later (Equation (10)), exactly where it is designed to act.

Best solution *n* at the end of estimation: Table 3 lists the Hill coefficient of the best final solution over the 100 trials for both the baseline and proposed frameworks. In both cases, most of the best solutions satisfy $n > 8$ (91% for the baseline; 96% for the proposed), indicating that both frameworks reliably identify oscillatory regimes. However, the proposed framework achieves substantially lower computational cost (e.g., the median is reduced from 25.5 million to 16.6 million RHS evaluations; $\Delta_{\text{cost}} = 0.35$; Section 3.6.2). The few best solutions with $n \leq 8$ suggest that some kinetic parameter combinations can transiently mimic sustained oscillations within $T = 200$; conclusively resolving such borderline cases would require longer horizons. Even in these cases, the adaptive simulation termination strategy can reject many damped candidates before a full-length simulation, reducing the computational cost while maintaining accuracy.

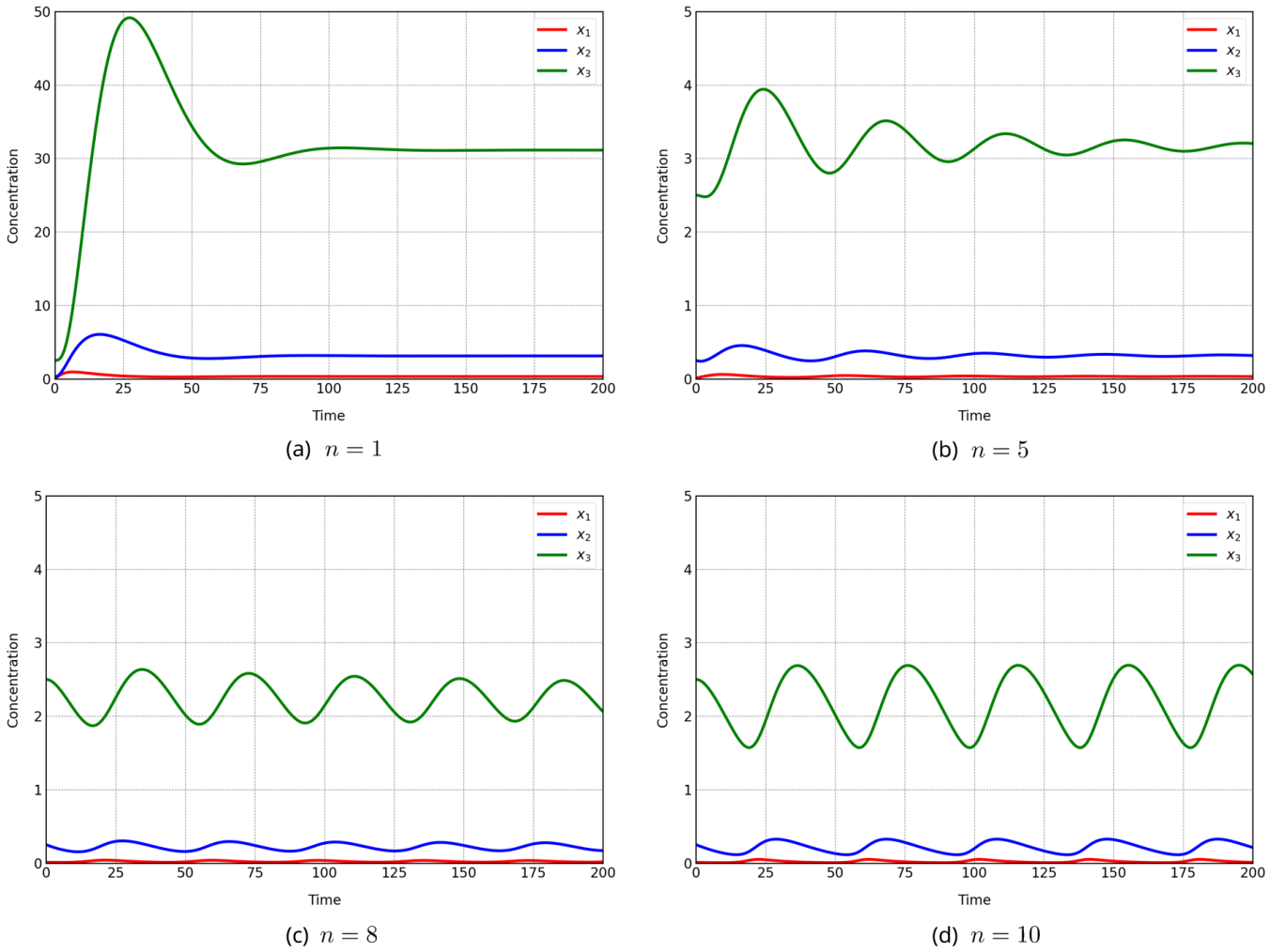

Figure 21. Goodwin trajectories for $n = \{1, 5, 8, 10\}$ at $T = 200$; $n = 10$ sustained and $n = 1, 5, 8$ damped (gradual amplitude decay for $n = 8$).

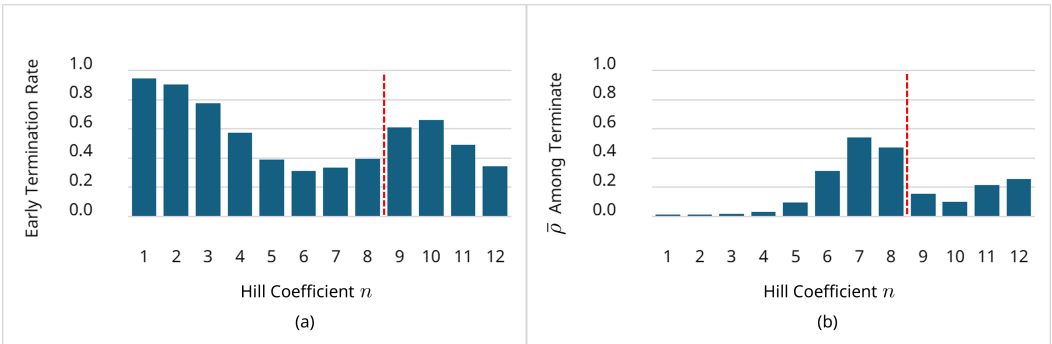

Figure 22. (**a**) Early termination rate vs. Hill coefficient n. (**b**) Mean ρ among terminated candidates vs. n: early termination for $n \leq 6$; later termination for $n = 7$ and 8 (weakly damped trajectories). The red dashed line indicates the nominal Hill coefficient used in the Goodwin experiments ($n = 10$).

Table 3. Distribution of the estimated Hill coefficient n among the best final solutions over 100 independent trials.

Hill Coefficient *n*	**1–6**	**7**	**8**	**9**	**10**	**11**	**12**
Baseline	0	1	8	42	44	5	0
Proposed	0	0	4	42	45	9	0

4. Discussion

4.1. Effectiveness of the Adaptive Simulation Termination Strategy

The primary contribution of this work is an adaptive simulation termination strategy designed specifically for oscillatory biochemical reaction networks. This strategy significantly enhances the efficiency of iterative optimization algorithms applied to nonlinear inverse problems. Across the tested benchmarks, the strategy consistently reduces computational cost—measured by RHS evaluations—by approximately 30–50% relative to the baseline, without compromising practical estimation accuracy (Sections 3.6.1 and 3.6.2). These savings arise because the method identifies and rejects clearly inferior candidates early, avoiding wasteful full-horizon simulations.

Crucially, the approach addresses the transient nature of oscillatory dynamics. In biochemical oscillators, early time windows often exhibit phase misalignment or amplitude adjustments that differ substantially from the steady limit cycle observed later. Any naive early termination rule with a static or overly strict threshold risks discarding otherwise valid solutions during this transient phase, resulting in false negatives. In contrast, our time-dependent boundary (Equation (10)) is permissive at early observation indices and becomes increasingly strict as the simulation proceeds, explicitly accounting for this temporal structure. The generation-wise evolution of the quantile references and boundaries (Figure 7) confirms the formation of a buffer zone at early indices and progressive tightening at later stages, achieving an effective balance between exploration and candidate screening.

A notable benefit of the time-decreasing boundary is its ability to discriminate sustained from damped oscillations (Section 3.6.9). Candidates that initially mimic the data but gradually decay toward a steady state, producing weakly damped oscillations, inevitably accumulate a persistent amplitude error relative to sustained trajectories. Even in the extreme case where the simulated trajectory decays to an approximately constant level (e.g., near the mean of an oscillatory observation), the mean absolute deviation from a sustained oscillation remains on the order of the oscillation amplitude—for a pure sine wave, it is $2/\pi \approx 63.7\%$ of the amplitude—and therefore does not vanish. As the boundary tightens over time, these gradually decaying candidates cross the boundary and are terminated at mid-to-late indices, whereas genuinely sustained oscillators, even those with phase shifts, remain valid candidates.

Another distinctive feature of the approach is its reliance on population-level statistics, specifically quantiles, rather than a single incumbent solution, to define rejection thresholds. In high-dimensional, multimodal landscapes, a current best solution may be local and unrepresentative; using it as a rigid threshold risks the premature loss of diversity and stagnation. By adapting to a population quantile (e.g., $Q_{0.75}$), the method adjusts the rejection pressure to the convergence state: thresholds are high in early, diverse generations and tighten naturally as the population improves. Additionally, the one-generation lag and warm-up evaluation (Section 2.6) decouple the threshold updates from the evaluation, preventing circularity and curbing over-aggressive termination.

As shown in Figure 2, the final fitness distributions overlap substantially between the proposed and baseline frameworks. Although early termination introduces a potential trade-off, in which rare candidates that improve later might be rejected, the empirical impact is negligible. This risk is effectively mitigated by the substantial cost savings and by design choices, including the transient-tolerant boundary with a one-generation lag, which reduces false negatives.

While surrogate model acceleration is a popular approach to reduce computational costs, constructing accurate surrogates for oscillatory systems is notoriously challenging. Small perturbations in parameters can induce bifurcations or qualitative regime changes, such as switching from weakly damped to sustained oscillations. Our framework avoids

this issue by performing the numerical integration of exact ODEs and achieving computational savings through time-adaptive truncation rather than approximation. This positions our framework as a robust iterative optimization method that maintains the mathematical rigor of ODE systems while addressing the computational bottlenecks of complex nonlinear simulations.

4.2. Limitations and Practical Guidance

Observation design and horizon length: For partially observed systems (e.g., the Goodwin model with $x_3(t)$ only, as in Section 3.1 (b)), the objective function may be less constrained, and borderline cases can appear plausible within limited simulation horizons. When borderline regimes (e.g., Hill coefficients $n = 7$ and 8 in the Goodwin model, as shown in Section 3.6.9) are of particular interest, using a longer simulation horizon or multiple complementary initial conditions can improve discrimination. When transients dominate most of the observation horizon (e.g., short horizons or strongly underdamped regimes), the partial objective $J^{(j)}$ may remain weakly informative until late times; in such cases, using a longer horizon and/or multiple initial conditions improves robustness of the termination decision.

Hyperparameter selection (α, β): The sensitivity study (Figure 8) demonstrates stable behavior for $\alpha \in \{1.2, 1.5\}$ and $\beta \in \{0.5, 0.7, 0.9\}$. As a default, we recommend $\alpha = 1.2$ (moderately permissive early) and $\beta = 0.7$ (moderate tightening). If false negatives are suspected (over pruning), α can be increased slightly or tightening slowed (decrease β); if cost reduction is the priority and accuracy is stable, α can be decreased or tightening accelerated (increase β). In addition, our noisy observation experiments (Section 3.6.6) indicate that, for the tested moderate noise levels ($r = 0.05$ and 0.10), substantial cost reduction was retained without substantially widening the boundary (i.e., using the same default q, α, and β settings). For stronger noise levels or different noise models, however, retuning the boundary parameters may be necessary, and the cost–accuracy trade-off may change.

Stiff or large-scale ODE systems: For very stiff systems or large networks, step size control and event handling may interact with the time-sequential evaluation. In such cases, it is advisable to set the maximum step size conservatively to ensure the stability of the numerical integration (Section 3.4). This interaction becomes more apparent when broader search domains are used; as shown in Section 3.6.7 (experiments for broader search range), some trials enter numerically difficult regions where the solver workload increases substantially, highlighting the importance of conservative step size control and, if needed, additional solver/termination tuning.

4.3. Future Work

While we focused on parameter estimation for known ODE structures, the proposed adaptive termination strategy has broader implications for data-driven structure identification. In many systems biology applications, the model structure itself is unknown. Genetic programming (GP) and symbolic regression are promising approaches for discovering state equations; however, the search space is combinatorial: determining and evaluating each candidate's structure requires even more kinds of solving ODE systems, making the computational burden even greater than in parameter estimation [20–23].

In our previous study, we proposed a GP approach that utilized divided experimental data to estimate state equations for oscillatory biochemical reaction networks [24]. A characteristic feature of that method was the imposition of perturbations through the addition of experimental data at various intervals, where adaptation to these changes can be driven the evolution of state equations with higher fitness. However, this approach suffered from high

computational costs because accurately detecting gradually decaying oscillatory behavior required lengthy numerical integration for every candidate.

Our approach is particularly well-suited to this challenge because invalid or ill-posed structures generated by GP often exhibit non-physical behavior or rapid divergence, making them ideal targets for early termination. Integrating the adaptive simulation termination strategy into genetic programming-based structure inference is a promising direction for future research, potentially enabling the efficient discovery of complex topologies and state equations for oscillatory biochemical reaction networks.

5. Conclusions

In this study, we proposed an efficient parameter estimation framework for oscillatory biochemical reaction networks that incorporates a novel adaptive simulation termination strategy. By evaluating the objective sequentially over observation indices and terminating ODE simulations when the partial objective exceeds a quantile-based, time-adaptive boundary, the framework aims to reduce wasted computation on persistently poor candidates while remaining tolerant to early transients. Under the tested settings and models (the Lotka–Volterra and the Goodwin model), the proposed method achieved substantial computational savings while maintaining competitive estimation performance; the efficiency gain was supported both by solver-reported effort (RHS evaluations) and by wall clock time measured with full overhead.

Additional experiments with moderate measurement noise suggest that the cost reduction can be retained without substantial retuning of the default parameters. Furthermore, evaluation of a grace period indicates that disabling adaptive termination in the early portion of the horizon increases cost without improving the cost–accuracy trade-off under the tested settings. Overall, the proposed approach provides a practical acceleration mechanism for ODE-constrained inverse problems in which objective function evaluation dominates runtime, particularly for oscillatory dynamics where transient behavior complicates early screening.

Author Contributions: Conceptualization, T.S. and M.O.; methodology, T.S.; software, T.S.; validation, T.S.; formal analysis, T.S.; investigation, T.S.; resources, T.S.; data curation, T.S.; writing—original draft preparation, T.S.; writing—review and editing, T.S., H.H. and M.O.; visualization, T.S.; supervision, T.S.; project administration, and T.S.; funding acquisition, T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by JSPS KAKENHI Grant Number 25K15344.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: During the preparation of this manuscript/study, the authors used ChatGPT 5.2 for the purposes of debugging the Python (version 3.14.2) scripts for numerical experiments. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kitano, H. Systems biology: A brief overview. *Science* **2002**, *295*, 1662–1664. [CrossRef] [PubMed]
2. Kikuchi, S.; Tominaga, D.; Arita, M.; Takahashi, K.; Tomita, M. Dynamic modeling of genetic networks using genetic algorithm and S-system. *Bioinformatics* **2003**, *19*, 643–650. [CrossRef]
3. Crampin, E.J.; Schnell, S.; McSharry, P.E. Mathematical and computational techniques to deduce complex biochemical reaction mechanisms. *Prog. Biophys. Mol. Biol.* **2004**, *86*, 77–112. [CrossRef]

4. Fernandez, M.R.; Mendes, P.; Banga, J.R. A hybrid approach for efficient and robust parameter estimation in biochemical pathways. *BioSystems* **2006**, *83*, 248–265. [CrossRef]
5. Tangherloni, A.; Spolaor, S.; Cazzaniga, P.; Besozzi, D.; Rundo, L.; Mauri, G.; Nobile, M.S. Biochemical parameter estimation vs. benchmark functions: A comparative study of optimization performance and representation design. *Appl. Soft Comput.* **2019**, *81*, 105494. [CrossRef]
6. El Yazidi, Y.; Ellabib, A. Reconstruction of the Depletion Layer in MOSFET by Genetic Algorithms. *Math. Model. Comput.* **2020**, *7*, 96–103. [CrossRef]
7. Zaki, G.; Akbarfam, A.J.; Irandoust-Pakchin, S. On the Inverse Problem of Time Fractional Heat Equation Using Crank-Nicholson Type Method and Genetic Algorithm. *Filomat* **2024**, *38*, 6829–6849. [CrossRef]
8. Okamoto, M.; Nonaka, T.; Ochiai, S.; Tominaga, D. Nonlinear numerical optimization with use of a hybrid Genetic Algorithm incorporating the Modified Powell method. *Appl. Math. Comput.* **1998**, *91*, 63–72. [CrossRef]
9. Akimoto, Y.; Nagata, Y.; Sakuma, J.; Ono, I.; Kobayashi, S. Analysis of The Behavior of MGG and JGG As A Selection Model for Real-coded Genetic Algorithms. *Trans. Jpn. Soc. Artif. Intell.* **2010**, *25*, 281–289. [CrossRef]
10. Akimoto, Y.; Nagata, Y.; Sakuma, J.; Ono, I.; Kobayashi, S. Proposal and Evaluation of Adaptive Real-coded Crossover AREX. *Trans. Jpn. Soc. Artif. Intell.* **2009**, *24*, 446–458. [CrossRef]
11. Komori, A.; Maki, Y.; Nakatsui, M.; Ono, I.; Okamoto, M. Efficient Numerical Optimization Algorithm Based on New Real-Coded Genetic Algorithm, AREX + JGG, and Application to the Inverse Problem in Systems Biology. *Appl. Math.* **2012**, *3*, 1463–1470. [CrossRef]
12. Goodwin, B.C. Oscillatory behavior in enzymatic control processes. *Adv. Enzyme Regul.* **1965**, *3*, 425–438. [CrossRef] [PubMed]
13. Higgins, J. The Theory of Oscillating Reactions. *Ind. Eng. Chem.* **1967**, *59*, 18–62. [CrossRef]
14. Griffith, J.S. Mathematics of cellular control processes I. Negative feedback to one gene. *J. Theor. Biol.* **1968**, *20*, 202–208. [CrossRef] [PubMed]
15. Gonze, D.; Abou-Jaoudé, W. The Goodwin Model: Behind the Hill Function. *PLoS ONE* **2013**, *8*, e69573. [CrossRef]
16. Gonze, D.; Ruoff, P. The Goodwin Oscillator and its Legacy. *Acta Biotheor.* **2021**, *69*, 857–874. [CrossRef]
17. Powell, M.J.D. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Comput. J.* **1964**, *7*, 155–162. [CrossRef]
18. Lotka, A.J. Analytical Note on Certain Rhythmic Relations in Organic Systems. *Proc. Natl. Acad. Sci. USA* **1920**, *6*, 410–415. [CrossRef]
19. Volterra, V. Fluctuations in the Abundance of a Species considered Mathematically. *Nature* **1926**, *118*, 558–560. [CrossRef]
20. Koza, J.R.; Mydlowec, W.; Lanza, G.; Yu, J.; Keane, A.M. Reverse engineering of metabolic pathways from observed data using genetic programming. *Pac. Symp. Biocomput.* **2001**, 434–445. [CrossRef]
21. Sugimoto, N.; Sakamoto, E.; Iba, H. Inference of Differential Equations by Using Genetic Programming. *J. Jpn. Soc. Artif. Intell.* **2004**, *19*, 450–459. [CrossRef]
22. Sugimoto, M.; Kikuchi, S.; Tomita, M. Reverse engineering of biochemical equations from time-course data by means of genetic programming. *BioSystems* **2005**, *80*, 155–164. [CrossRef] [PubMed]
23. Iba, H. Inference of differential equation models by genetic programming. *Inf. Sci.* **2008**, *178*, 4453–4468. [CrossRef]
24. Sekiguchi, T.; Hamada, H.; Okamoto, M. Inference of General Mass Action-Based State Equations for Oscillatory Biochemical Reaction Systems Using *k*-Step Genetic Programming. *Appl. Math.* **2019**, *10*, 627–645. [CrossRef]