Measuring regulatory network inheritance in dividing yeast cells using ordinary differential equations

Wenbin Wu
Department of Statistics, University of Washington
and
Taylor Kennedy
Department of Plant and Microbial Biology, North Carolina State University
and
Orlando Arguello-Miranda
Department of Plant and Microbial Biology, North Carolina State University
and
Kevin Z. Lin
Department of Biostatistics, University of Washington

November 23, 2024

Abstract

Quantifying the inheritance of regulatory networks among proteins during asymmetric cell division remains a challenge due to the complexity of these systems and the lack of robust mathematical definitions for inheritance. We propose a novel statistical framework called ODEinherit to measure how much a mother cell's regulatory network explains its daughter's trajectories, addressing this gap. Using time-lapse microscopy, we tracked the expression dynamics of six proteins across 85 dividing *S. cerevisiae* cells, observed over eight hours at 12-minute intervals. Our framework employs a two-step approach. First, we estimate an ordinary differential equation (ODE) system for each cell to characterize protein interactions, introducing novel adjustments for non-oscillatory time series and leveraging multi-cell data. Second, we assess inheritance by clustering cells based on cycling markers and quantifying how well a mother's regulatory network predicts her daughter's. Preliminary findings suggest stage-dependent differences in inheritance rates, paving the way for applications in cellular stress response and cell-fate prediction studies across generations.

Keywords: Asymmetric cell division, Cellular dynamics, Live-cell imaging, Protein interaction networks

1 Introduction

Understanding how cells divide and pass on regulatory information is fundamental to biology, with implications for processes such as cancer progression, immune responses, and tissue regeneration. Studying cell cycling, which is the process of how cells divide to yield new cells, unravels the potential for understanding tumor progression and cancer therapies (Otto and Sicinski, 2017; Ma and Gurkan-Cavusoglu, 2024). During cell cycling, the inheritance of molecular components can vary, especially in asymmetric divisions, where distinct daughter cell phenotypes emerge (Higuchi-Sanabria et al., 2014; Herrero et al., 2020). The *saccharomyces cerevisiae* (budding yeast) is a commonly used model organism to study cell cycling due to its fast division rate. In budding yeast, mother cells can influence the fate of their daughters. Notably, studies such as those by Argüello-Miranda et al. (2018) have shown that a daughter's fate can be predicted based on its mother's properties, even before it is physically formed. These observations highlight the need to investigate how regulatory networks shape cell fate across generations.

Despite these advances, quantifying the inheritance of protein regulatory networks from the mother to the daughter cell remains a challenge due to the dynamic and complex nature of these systems. Current approaches often fail to capture the network-level inheritance of cellular regulatory machinery. To address this gap, we propose a novel statistical framework that uses live-cell imaging data from yeast to rigorously measure the inheritance of protein regulatory networks. By bridging mathematical modeling with biological insights, our method seeks to determine whether cell fate is driven by the transmission of these networks, providing a new perspective on the interplay between cellular dynamics and regulatory inheritance.

2 Background

We motivate our statistical method by first explaining a preliminary analysis of the dataset we collected of budding yeast. This preliminary analysis attempts to investigate the amount of inheritance between a mother and daughter cell, but as we will see, existing methods are not quite apt for measuring this quantity. The lackluster results in the preliminary analysis will motivate our new statistical method in the later sections.

To investigate cellular inheritance, we collect time-lapse microscopy imaging data which will be the main dataset of interest through this entire paper. This data is generated by tracking the expressions of 6 proteins simultaneously using fluorescent reporters in dividing yeast cells. Each dataset of different combinations of six proteins contains 200+ cells over eight hours with a twelve-minute sampling rate, where cells divide up to four times. The proteins are markers of cell fate and cell cycle activities. There are 48 time points in 25 mother cells and their 60 daughter cells. The time series of mother cells are observed throughout the course, while those of daughter cells are incomplete and only observed after birth. See Appendix S1 and Ramakanth et al. (2024) for more details on how this data was collected and processed. Using this dataset, we hope to investigate whether daughter cells inherit certain protein dynamics from their mother cells. Because the yeast cells are at different cellular states at the start of the experiment, we do not necessarily expect that all mother-daughter relations are necessarily the same across all cell pairs. Therefore, we seek to perform a preliminary analysis that can investigate the amount of inheritance heterogeneity among the 60 mother-daughter pairs.

2.1 Preliminary analysis

In our preliminary analysis, we demonstrate that existing statistical frameworks, such as Granger causality, are insufficient to model cellular inheritance. To start, we first provide qualitative evidence that there are differences in the amount of inheritance between

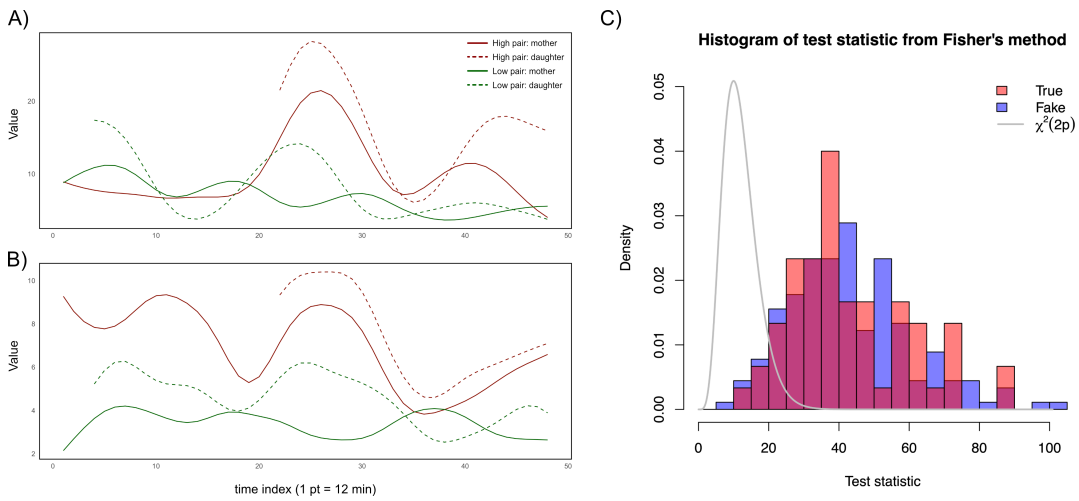

Figure 1: **A)** and **B)**: Time series of cell pairs with high (red) and low (green) degrees of inheritance, qualitatively speaking, with mother and daughter cells shown in solid lines and dashed lines, respectively. Each plot shows one (out of the six) protein as a time series. **C)**: Test statistics from preliminary Granger causality analysis, with true pairs in red and fake pairs in blue, compared to the $\chi^2(2p)$ reference distribution in gray.

mother-daughter pairs in our data. Figures 1A and B present time series examples for two mother-daughter pairs, with each figure representing a different protein variable. We see that in both proteins, in the red pair, the daughter cell after birth visually follows the cyclical expression patterns of the mother cell — this suggests there is a higher amount of inheritance in this mother-daughter pair. In contrast, in the green pair, the daughter cell after birth seems much more uncorrelated with the mother cell, both in terms of amplitude and frequency in its cyclical expression pattern — this suggests there is a lower amount of inheritance in this mother-daughter pair. These two example mother-daughter pairs, among the 60 mother-daughter pairs in our dataset, indicate a possible variability among how much the daughter cell inherits from its mother in terms of protein dynamics.

To quantify this phenomenon rigorously, one might consider using Granger causality. Granger causality is commonly used to evaluate how much one time series is predictive of another (Shojaie and Fox, 2022), making it a natural choice for assessing inheritance in our setting. However, as demonstrated in the following preliminary analysis, this strategy is insufficient to quantify how much regulation is passed from a mother to its daughters.

For a pair of univariate time series x_t and y_t, the series x_t is said to be Granger causal of y_t if the inclusion of x_t adds explanatory power to the autoregression of y_t. Specifically, it performs an F-test in the following model on the hypothesis $H_0 : b_1 = \cdots = b_q = 0$, namely, the joint significance of the lag effects in the mother series,

$$y_t = a_0 + a_1 y_{t-1} + \cdots + a_p y_{t-q} + b_1 x_{t-1} + \cdots + b_q x_{t-q} + \epsilon_t,$$

where the lag q is specified in advance.

Now, we show that this Granger causality analysis is inadequate for measuring inheritance, as it yields close similarity measures for true mother-daughter pairs and unrelated pairs. For each mother-daughter pair, we perform the Granger causality test on their time series using a maximum lag of 2. A p-value is obtained for each variable, denoted by pval_j for $j = 1, \ldots, p$ where $p = 6$. We aggregate the p-values across the p proteins for each mother-daughter pair by Fisher's method, calculating the test statistic $X^2 = -2\sum_{j=1}^{p} \log(\text{pval}_j)$, which follows a $\chi^2(2p)$ distribution under the null hypothesis that the mother series is not predictive of the daughter series for any variable. A larger value of this test statistic indicates that the mother cell's protein expression is predictive of the daughter cell's protein expression, which could suggest a statistically significant amount of inheritance in this mother-daughter pair. As a comparison, we also pair each daughter cell with three non-mother cells from the mother generation and compute the same test statistic for these "fake pairs" as a null distribution. Since the mother cells are treated as independent biological replicates, we expect the test statistic to be larger (indicating more inheritance) in true mother-daughter pairs and smaller (indicating less inheritance) in fake pairs. In other words, if protein regulatory networks are indeed inherited by daughter cells, we hypothesize that the degree of inheritance would be higher in true pairs than in fake pairs.

Figure 1C presents the histograms of the test statistic for true pairs and fake pairs, with the reference $\chi^2(2p)$ distribution shown in gray. We observe considerable overlap between

the distributions for true pairs and fake pairs, both of which show significance under the null hypothesis. This phenomenon suggests that any two cells exhibit a high degree of inheritance, irrespective of an actual mother-daughter relationship. Based on additional analyses beyond the scope of this paper, we assess that the problematic quality is more foundational than simply recalibrating the Granger causality's null distribution. This result provides strong indication that the Granger causality test is inadequate for distinguishing true inheritance between genuine mother-daughter pairs and unrelated pairs.

While multivariate and non-parametric extensions of Granger causality could fix certain statistical shortcomings, we note one substantive shortcoming that is fundamental to any Granger causality framework. Granger causality assumes a stationary lag where the mother can predict the daughter's protein expression. This statistical premise is potentially unrealistic, since after the daughter cell separates from the mother, there is little biological explanation on why the mother and daughter cells' protein expression will remain in-sync. Instead, we hypothesize that the daughter inherits the *regulatory machinery* from the mother. This means the daughter cell can inherit post-translational modifications and other intermediary proteins from its mother, all of which influence how the proteins interact in the daughter cell (Infant et al., 2021; Hamey and Wilkins, 2023). This means even if the mother cell's expression is not predictive of the daughter's expression with a stationary lag, the protein-protein interactions across time in both cells could be nearly identical. We hypothesize that this latter framework is a more accurate depiction of the underlying biology. Therefore, we seek to take a different statistical approach — can we first model a mother cell's regulatory network and then assess how much its regulatory network explains the variability in the daughter cell's time series?

2.2 Modeling as an ordinary differential equation system

Protein-protein interaction networks offer a framework for understanding how proteins dynamically interact and influence cellular processes. Each node represents a protein, while edges indicate interactions such as activation, inhibition, or stabilization. In yeast, such networks can differ between cells due to factors like post-translational modifications, which alter protein function by chemical additions such as phosphorylation or ubiquitination (Infant et al., 2021; Hamey and Wilkins, 2023). Furthermore, cellular asymmetry during division can result in daughter cells inheriting distinct initial protein abundances (Higuchi-Sanabria et al., 2014; Herrero et al., 2020). This variation, coupled with dynamic regulatory feedback, implies that while regulatory dynamics may be similar, observed protein levels across time could seem uncorrelated between mother and daughter cells. Statistically capturing these subtleties can help uncover fundamental principles of cellular regulation and inheritance.

Mathematical modeling of these regulatory dynamics as a network in cell cycles has been a long-standing area of research. Graphical models, while effective for certain applications, are not well-suited for capturing the time-dependent regulatory dynamics in our time-series data. In contrast, Ordinary Differential Equation (ODE) systems excel at modeling feedback circuits within the oscillatory regulations that drive cell cycles (Pomerening et al., 2003; Tyson and Novák, 2015). They have been proven to be an effective framework for describing protein interaction networks (Sible and Tyson, 2007). In previous work, ODE models have been extensively applied to study the budding yeast cell cycle, effectively characterizing the dynamics between various cell cycle activities (Pomerening et al., 2003; Chen et al., 2004; Radde and Kaderali, 2009; Boczko et al., 2010). In one of our recent studies, we employed ODE systems to model the regulatory network governing the entry into meiosis, demonstrating their utility in addressing the specific biological problems we investigate (Kociemba et al., 2024). However, estimating the parameters of an ODE system

based on data is hard, and instead, much work studying yeast cells relies on manually selecting the parameters or performing a grid search (Tyson and Novák, 2015; Jashnsaz et al., 2021). In contrast, we focus on estimating these parameters from data in this paper.

2.3 Existing ODE estimation methods

From the statistical perspective, significant effort has been devoted to estimating ODE systems and studying the theoretical properties of the resulting estimators. The general goal is to estimate a system of ODEs in the form:

$$\frac{dx(t)}{dt} = \begin{bmatrix} \frac{dx_1(t)}{dt} \\ \vdots \\ \frac{dx_p(t)}{dt} \end{bmatrix} = \begin{bmatrix} F_1(x(t)) \\ \vdots \\ F_p(x(t)) \end{bmatrix} = F(x(t)), \tag{1}$$

where t denotes the time index reparameterized to an interval $\mathcal{T} = [0, 1]$, and the unknown functionals $F = \{F_1, \ldots, F_p\}$ are the estimands. These functionals describe the regulatory dynamics among the variables and may either have an unknown form or be parameterized in a known form. In other words, this system models how the trajectories of all p variables collectively influence the instantaneous rate of change of each variable. Typically, for a single sample (i.e., cell), the unknown variable trajectories $x(t) = (x_1(t), \ldots, x_p(t)) \in \mathbb{R}^p$ are observed on n discrete time points $\{t_1, ..., t_n\}$ with measurement errors:

$$y_i = x(t_i) + \epsilon_i, \quad \text{for } i = 1, \ldots, n,$$

where $y_i = (y_{i1}, \ldots, y_{ip}) \in \mathbb{R}^p$ denotes the observed trajectories, and $\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{ip}) \in \mathbb{R}^p$ denotes the independent measurement errors with zero mean. Let $x(0) \in \mathbb{R}^p$ be the unknown initial conditions of this system.

Earlier studies primarily focused on estimating the functionals F_j's, assuming a known

regulatory structure, namely, which variables regulate which other variables. Various strategies have been developed to address the challenges of handling derivatives in the estimation process and ensure robust theoretical guarantees (Ramsay et al., 2007; Cao and Zhao, 2008; Liang and Wu, 2008; Brunel, 2008; Qi and Zhao, 2010; Xue et al., 2010; Gugushvili and Klaassen, 2012; Hall and Ma, 2014; Dattner and Klaassen, 2015). However, our study extends beyond estimating the exact regulatory dynamics. We aim to uncover the regulatory relationships among variables, emphasizing the reconstruction of a biologically interpretable network. This requires incorporating sparsity into the estimation, which adds complexity as we simultaneously model dynamics and infer a sparse regulatory structure. Recent advancements have introduced Lasso-type penalties in the objective function to encourage sparsity in the estimated functionals and have established the selection consistency (Wu et al., 2014; Zhang et al., 2015; Chen et al., 2017; Dai and Li, 2022), making them particularly relevant to our work. In this paper, we build upon the most sophisticated method called Kernel ODE (KODE, Dai and Li (2022)). The authors demonstrated its utility on yeast data in their paper, and the method is both flexible to capture nuanced regulatory dynamics due to its usage of the Reproducing Kernel Hilbert Space (RKHS). We review these details in the next section.

3 Methods

In this section, we describe the formal details of ODEinherit, our proposed method to measure how much a daughter cell inherits the mother cell's protein regulatory network, which we will denote as $\pi^{(M \to D)}$. ODEinherit's overall strategy involves first estimating the directed regulatory network for both mother and daughter cells by fitting an ODE system to the protein variables. Next, the inheritance measure is then defined as a percentage reflecting how well the mother network can explain the daughter trajectories. We summarize this workflow in Figure 2. The workflow is employed cell-wise to calculate an inheritance

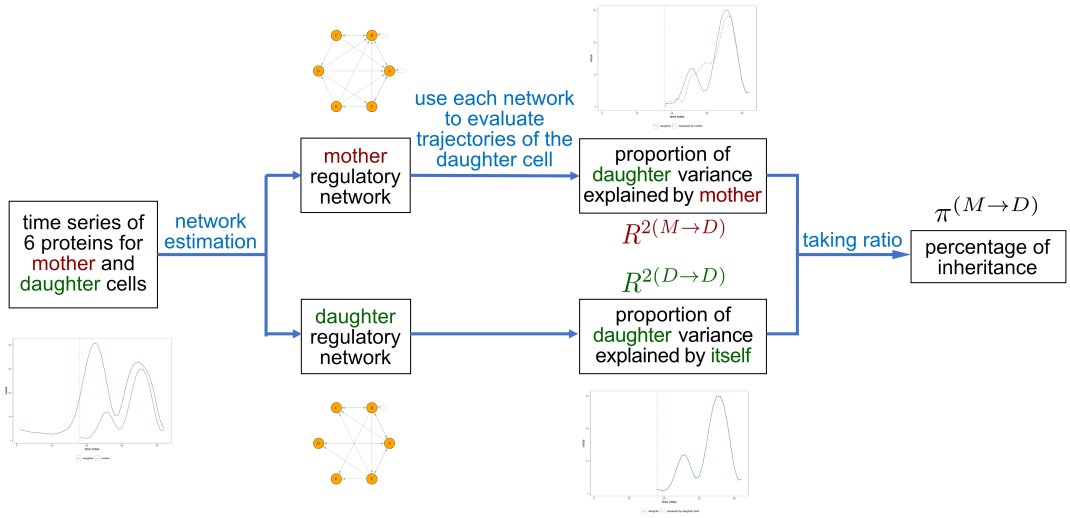

Figure 2: Workflow of measuring inheritance between a mother-daughter pair.

measure for each cell. For simplicity, we omit the cell index in the notations and denote the observed trajectories of a given cell as y_{ij} for $i = 1, \ldots, n$ and $j = 1, \ldots, p$.

3.1 ODE estimation and a review of Kernel ODE

We first introduce how we obtain a directed regulatory network of each cell by estimating the ODE system (1) from the observed time series (i.e., trajectories). We model our data using the commonly used additive model, where we assume that variables regulate a particular protein j via additive effects,

$$F_j(x(t)) = b_j + \sum_{k=1}^{p} \theta_{jk} F_{jk}(x_k(t)), \quad j = 1, \ldots, p, \tag{2}$$

where $b_j \in \mathbb{R}$ is the intercept, $\theta_{jk} \in \mathbb{R}$ and F_{jk} characterize the coefficient and dependency of variable j on each variable k, respectively. If $\theta_{jk} \neq 0$, we consider variable k to be a regulator of variable j and assign a directed edge from variable k to j to construct the regulatory network.

We review the estimation framework of Dai and Li (2022) (KODE) here, as our method will build upon this framework for inheritance analysis. For each variable j, KODE assumes

that the derivative function F_j resides in a RKHS $\mathcal{H} = \{1\} \oplus \sum_{k=1}^{p} \mathcal{H}_k$. Here, $\mathcal{H}_k$ is the RKHS generated by a given Mercer kernel K_k corresponding to variable k such that $F_{jk} \in \mathcal{H}_k$. In KODE, F_j is estimated using a two-step collocation strategy, which is known to be computationally effective for ODE estimation. The first step obtains smoothing estimates of the trajectories by

$$\widehat{z}_j(t) = \underset{z_j \in \mathcal{F}}{\arg\min} \left\{ \frac{1}{n} \sum_{i=1}^{n} \{y_{ij} - z_j(t_i)\}^2 + \lambda_{nj} \|z_j(t)\|_{\mathcal{F}}^2 \right\}, \quad j = 1, \ldots, p. \tag{3}$$

where $\mathcal{F}$ is a given space of smooth functions. Denote the smoothing estimates as $\widehat{z}(t) = (\widehat{z}_1(t), \ldots, \widehat{z}_p(t)) \in \mathbb{R}^p$. The second step estimates each functional $F_j \in \mathcal{H}$, the coefficients $\theta_j = (\theta_{j1}, \ldots, \theta_{jk}) \in \mathbb{R}^p$, and the initial condition $\theta_{j0} = X_j(0) \in \mathbb{R}$ by solving the following penalized optimization problem,

$$\min_{\theta_{j0}, \theta_j, F_j} \frac{1}{n} \sum_{i=1}^{n} \left\{ y_{ij} - \theta_{j0} - \int_0^{t_i} F_j(\widehat{z}(t)) dt \right\}^2 + \tau_{nj} \Big(\sum_{k=1}^{p} \|\theta_{jk} F_{jk}\|_{\mathcal{H}} \Big), \quad j = 1, \ldots, p. \tag{4}$$

An iterative optimization algorithm is used for estimation, where sparsity in $\widehat{\theta}_j$ is induced by a Lasso regularization. Let $\widehat{F}_j$ and $\widehat{\theta}_{j0}$ be the optimal solution from (4), then the trajectories are recovered by

$$\widehat{x}_j(t) = \widehat{\theta}_{j0} + \int_0^{t_i} \widehat{F}_j(\widehat{z}(t)) dt, \quad j = 1, \ldots, p. \tag{5}$$

The integration is evaluated using a first-order approximation over a fine grid on $\mathcal{T}$. The estimated regulatory network is constructed using the estimates $\{\widehat{\theta}_j, j = 1, \ldots, p\}$. We use $M_j = \{k : \widehat{\theta}_{jk} \neq 0\} \subset \{1, \ldots, p\}$ to denote the selected regulators of variable j for $j = 1, \ldots, p$, then the estimated network is given by $\{M_j : j = 1, \ldots, p\}$. We review additional details of KODE, such as data-driven strategies to tune parameters, in Appendix S2.

3.2 Measuring the goodness-of-fit of a network

We now introduce a statistic that measures the goodness-of-fit of this network based on how well it recovers the observed trajectories. This statistic will be critical to how we define inheritance in the following sections. For each variable j, we apply the KODE estimation algorithm to refit $\widehat{F}_j$ based on the smoothing estimates $\widehat{z}(t)$ by removing the Lasso regularization on $\widehat{\theta}_j$ and restricting $\widehat{\theta}_j \in M_j$. Let $\widehat{x}(t) = (\widehat{x}_1(t), \ldots, \widehat{x}_p(t))$ be the recovered trajectories from (5). We obtain the following variable-specific R^2 statistic,

$$R_j^2 = \max\left\{1 - \frac{\mathrm{MSS}_{j,\mathrm{res}}}{\mathrm{MSS}_{j,\mathrm{tot}}},\ 0\right\}, \quad \text{for } j = 1, \ldots, p,$$

where $\mathrm{MSS}_{j,\mathrm{res}} = \frac{1}{n}\sum_{i=1}^n \left\{y_{ij} - \widehat{x}_j(t_i)\right\}^2$ and $\mathrm{MSS}_{j,\mathrm{tot}} = \frac{1}{n}\sum_{i=1}^n \left\{y_{ij} - \bar{y}_j\right\}^2$ are the mean sums of squares of the residuals and of the observations, respectively, and $\bar{y}_j = \frac{1}{n}\sum_{i=1}^n y_{ij}$. Intuitively, R_j^2 measures the proportion of variance in the observations of variable j that is explained by the recovered trajectory $\widehat{x}_j(t)$. In cases where a negative R_j^2 value arises, the mean of the observations provides a better fit than our estimation, and we set $R_j^2 = 0$. To obtain an overall statistic for a cell, we use

$$R^2 = \frac{1}{p}\sum_{j=1}^p R_j^2, \tag{6}$$

namely, the average of R_j^2 across all variables.

Unlike prior studies that often focus on the asymptotics of the error in estimating trajectories such as $\int_{\mathcal{T}} \|\widehat{x}(t) - x(t)\|_2^2 dt$ (Dai and Li, 2022), we introduce a single-cell goodness-of-fit measure for R^2. This is because in many previous studies, there was no necessity to compute R^2 at the individual-cell level (i.e., refitting a cell's trajectory using its own network results in $R^2 = 1$ if the true F_j resides in the correct model space). In contrast, we define our goodness-of-fit metric (6) to serve two crucial purposes: (1) guiding the network pruning process by evaluating the explanatory power of pruned networks and (2) assess-

ing how well the mother cell's regulatory network explains the daughter cell's trajectories, which is central to our investigation of regulatory inheritance.

3.3 Measuring the protein regulatory inheritance between mother and daughter cells

We are now ready to describe how we quantify the daughter cell's inheritance of the mother's protein regulatory network. Our proposed metric uses the R^2 statistic to evaluate how much the mother cell's regulatory network explains the daughter cell's trajectories. Consider one mother-daughter pair, where we denote the mother and daughter cell's estimated regulatory networks as M and D, respectively. We refit the daughter trajectories using the pruned mother network and daughter network and obtain their R^2 values, $R^{2(M\to D)}$ and $R^{2(D\to D)}$. These represent the proportions of variance in daughter trajectories that are explained by the mother and daughter, respectively. We take a ratio to obtain a percentage of inheritance in this mother-daughter relationship,

$$\pi^{(M\to D)} = \min\left\{\frac{R^{2(M\to D)}}{R^{2(D\to D)}},\ 1\right\}. \tag{7}$$

Here, $R^{2(D\to D)}$ reflects the intrinsic explainability of the daughter trajectories, while $R^{2(M\to D)}$ reflects how well the mother's network explains the daughter trajectories. The quantity $\pi^{(M\to D)} \in [0, 1]$, which we call the *inheritance metric*, is the primary output of ODEinherit, where 0 and 1 represent the daughter cell inherits none or all of the mother cell's protein regulatory network, respectively.

The rationale for this metric stems from our hypothesis that the regulatory dynamics, represented by the F_j's, are not preserved in their exact form as they transition from the mother to the daughter cell. Instead, these functionals undergo modifications in the daughter cell as it progresses through its cycle, limiting the feasibility of directly predicting

the daughter's trajectories based on the mother's. Instead, we focus on the qualitative properties of the system, specifically, the regulatory network structure. By constraining the model space to the mother's regulator sets, we quantify the extent to which these sets can explain the daughter's trajectories, thereby capturing the inheritance of regulations without assuming stationarity in the dynamics.

3.4 Pruning the network

In this section, we describe a key quality that we empirically observed to adversely impact our inheritance metric and our procedure to remove this quality. Specifically, observe our inheritance metric $\pi^{(M\to D)}$ defined in (7). If M were a fully dense network (i.e., every protein regulates all proteins), then $\pi^{(M\to D)} = 1$ regardless of the daughter cell's trajectories. This means that our inheritance metric is not meaningful for dense networks. However, our empirical results suggest that KODE tends to select variables as regulators in the network when they are actually not (false positives) rather than overlook the actual regulators (false negatives). This leads to an overly dense estimated network, which hinders us from discerning the primary regulations between the variables. As we will show later in the simulations, having excessively dense networks will dramatically hinder our ability to measure the inheritance of regulatory networks. Hence, we describe in this section our strategy to refine the network in such a way that sparsifies each network without sacrificing the model fit.

Our strategy is to use the regulator sets selected by KODE as potential candidates and refine the network through an iterative pruning procedure. Our method leverages the R^2 statistic to quantify the importance of each regulator in terms of trajectory recovery and eliminates those with minor contributions. Specifically, for a given variable j, the pruning procedure evaluates each of its selected regulators by assessing the change in R_j^2 after its removal. A regulator is pruned if its removal does not reduce R_j^2 beyond a pre-

specified threshold (by default, 5%). Detailed steps are provided in Appendix S2. This approach significantly reduces the false positive rate at the cost of a slight increase in the false negative rate, resulting in sparser, more interpretable networks while ensuring no important regulators are omitted.

While sparser networks can also be obtained by increasing the Lasso regularization in KODE, the selection consistency depends on sufficiently strong regulatory effects in true edges and negligible effects in non-edges (Dai and Li, 2022; Chen et al., 2017). These assumptions can be challenging to satisfy under model misspecification or in the RKHS space, where the functional estimands take more complex forms. Simply thresholding the number of selected edges in Lasso is suboptimal, as it requires prior knowledge of the network and provides little insight into the significance of selected regulators. Our approach provides an explicit quantification of each of the six protein regulatory contributions while remaining computationally feasible. Empirically, this strategy demonstrates superior performance. However, in high-dimensional settings beyond the scope of this paper, imposing larger Lasso regularization likely remains a practical and effective choice.

4 Simulation study

In this section, we describe a suite of simulations to assess the reliability of our network estimation, even in misspecified settings, and to demonstrate that ODEinherit can meaningfully estimate the protein regulatory inheritance between simulated mother and daughter cells. We provide the overview in this section and defer additional details to Appendix S3.

4.1 Network estimation

We first evaluate the empirical performance of the network estimation strategy for both additive and non-additive ODE systems, where the assumed additive form (2) is violated in the latter case. It is important to verify this before we examine the performance of the

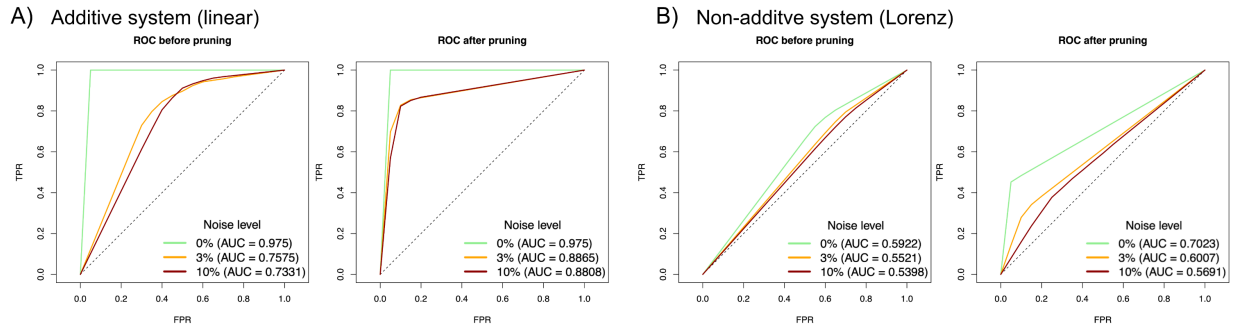

Figure 3: ROC curves of the network estimation in the additive **(A)** and non-additive system **(B)** under various noise levels, before (left) and after (right) our pruning procedure. The noise level is defined as the ratio of the noise variance, σ_j^2, to the variance of the true trajectory $x_j(t)$. The black dashed line represents the performance of a random classifier (baseline $AUC = 0.5$).

proposed inheritance measure under these ODE systems.

4.1.1 Additive ODE system

We first demonstrate that our network estimation procedure is accurate for a simple linear ODE system where the true model $F_j(x(t))$ is additive. This model is within the correct specification of the assumed generative model (2). Our simulated system consists of two triplets of variables, with solution trajectories being combinations of sine and cosine functions, defined over the interval $\mathcal{T} = [0, 1]$. It is specified as follows: for $k = 1$ or $k = 2$,

$$\begin{aligned} \frac{dx_{3k-2}(t)}{dt} &= -2A^{(k)} + A^{(k)}x_{3k-1}(t) - A^{(k)}x_{3k}(t), \\ \frac{dx_{3k-1}(t)}{dt} &= 2A^{(k)} - A^{(k)}x_{3k-2}(t) + A^{(k)}x_{3k}(t), \\ \frac{dx_{3k}(t)}{dt} &= A^{(k)}x_{3k-2}(t) - A^{(k)}x_{3k-1}(t), \end{aligned} \tag{8}$$

for time $t \in [0, 1]$. where $A^{(1)}$ and $A^{(2)}$ are chosen to make the triplets exhibit 3 and 10 periods on the interval $[0, 1]$, respectively, minimizing the correlation between the two triplets. We draw $n = 200$ data observations from the solution trajectories at the evenly-spaced time grid $\{1/n, 2/n, \ldots, 1\}$ with identical and independent Gaussian measurement errors $\epsilon_{ij} \overset{i.i.d.}{\sim} N(0, \sigma_j^2)$.

Figure 3A demonstrates that our network estimation is accurate after pruning. Specif-

Table 1: R^2 values for the original and pruned estimated networks in simulations of additive and non-additive systems at various noise levels.

System	Type	Noise level		
		0%	3%	10%
Additive	Original R^2	1.00	0.85	0.81
	Pruned R^2	1.00	0.87	0.83
Non-additive	Original R^2	0.99	0.98	0.96
	Pruned R^2	0.98	0.97	0.95

ically, we plot the ROC curves for the estimated networks before and after pruning, where the noise variance σ_j^2 is set to 0% (i.e., no noise), 3%, and 10% of the sample variance of the true trajectory values $\{x_j(t_i), i = 1, \ldots, n\}$. For each noise level, we perform 100 simulation runs, calculating the false positive rate (FPR) and true positive rate (TPR) for each run. The ROC curve and the area under the curve (AUC) are computed using these (FPR, TPR) pairs. It is seen that, in the presence of noise, the estimated network without pruning tends to be dense with a high FPR. The pruned network achieves a sparser graph, reducing the FPR while maintaining a similar level of explanatory power for the trajectories, as indicated by the high R^2 values in Table 1. We provide additional results in Appendix S3.

4.1.2 Non-additive ODE system

We next demonstrate that we can still estimate meaningful networks from much more complex trajectories, even if the generative model is not within assumed model space (2). This simulation uses a non-additive ODE system, where the additivity assumption is violated. Specifically, we consider the Lorenz system, a non-linear and aperiodic ODE system with interaction terms. The generating ODE model includes two triplets from this

system with different sets of parameters: for $k = 1$ or $k = 2$,

$$\begin{aligned}
\frac{dx_{3k-2}(t)}{dt} &= \sigma^{(k)} x_{3k-1}(t) - \sigma^{(k)} x_{3k-2}(t), \\
\frac{dx_{3k-1}(t)}{dt} &= \rho^{(k)} x_{3k-2}(t) - x_{3k-2}(t) x_{3k}(t) - x_{3k-1}(t), \\
\frac{dx_{3k}(t)}{dt} &= x_{3k-2}(t) x_{3k-1}(t) - \beta^{(k)} x_{3k}(t),
\end{aligned} \tag{9}$$

for time $t \in [0, 100]$ before rescaling time to range from 0 to 1. We set the parameters $\{\sigma^{(1)} = 10, \rho^{(1)} = 28, \beta^{(1)} = \frac{8}{3}\}$ and $\{\sigma^{(2)} = 5, \rho^{(2)} = 45, \beta^{(2)} = \frac{3}{2}\}$ such that the trajectories oscillate indefinitely. We draw $n = 200$ data observations at an evenly-spaced time grid over the intervals $[40, 50]$ and $[40, 60]$ for the first and the second triplets, respectively. This ensures that the frequencies are different for the triplets and are low enough for the oscillations to be captured in the observations. The measurement errors are $\epsilon_{ij} \overset{i.i.d.}{\sim} N(0, \sigma_j^2)$, where σ_j^2's reflect the noise levels as described in Section 4.1.1.

Despite analyzing misspecified data, our results nonetheless demonstrate that pruning enhances the accuracy of our estimated networks. For simplicity in the analysis, an edge is assigned in the regulatory network whenever one variable affects another, whether through an additive effect or an interaction effect. For example, we assign an edge $x_{3k} \to x_{3k-1}$ even though x_{3k} affects x_{3k-1} indirectly by interacting with x_{3k-2}. Figure 3(B) shows the ROC curves for the estimated networks, and the R^2 values are presented in Table 1. We can see that, although the selection accuracy is generally compromised under model misspecification, the pruning procedure still enhances overall accuracy while preserving a similar level of explanatory power for the trajectories.

4.2 Inheritance measures

With a reliable network estimate, we now move to evaluate ODEinherit's ability to reliably measure the amount of protein regulatory inheritance. To do this, we need to extend our simulation framework in order to generate mother-daughter pairs under each of the two

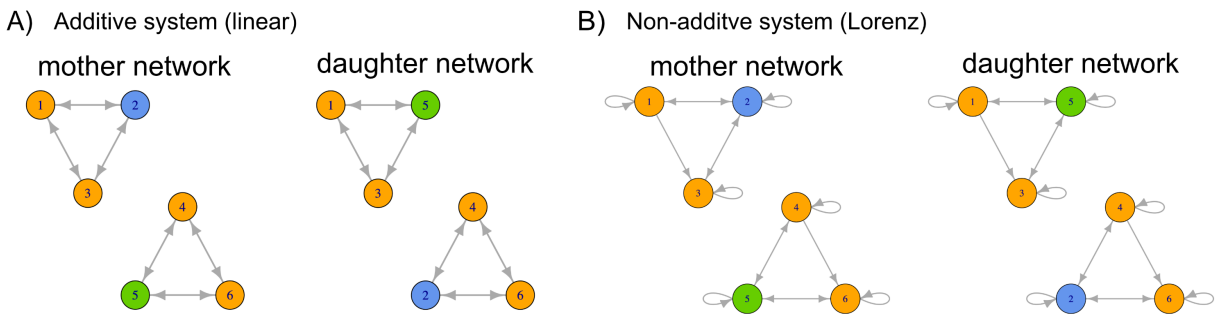

Figure 4: True mother and daughter regulatory networks in the additive system (**A**) and the non-additive system (**B**). Variables 2 and 5 are interchanged for the daughter, such that its network has an equivalent complexity as that of the mother.

ODE systems in Section 4.1. Our new simulation framework involves two components: a mother system and a pseudo-daughter system, which are used to generate the trajectories of the observed mother cell and an unobserved "pseudo-daughter" cell, respectively. The observed daughter trajectories are then generated as a convex combination of these two trajectories, where the combination weight controls the degree of inheritance between the mother and daughter cells. Intuitively, the pseudo-daughter trajectories represent the daughter's intrinsic dynamics without inheritance, while the actual daughter trajectories reflect partial inheritance from the mother.

We provide an overview of how we simulate mother-daughter pairs. Consider an ODE model with two triplets, either (8) or (9). In the mother system, we set the triplets $(1, 2, 3)$ and $(4, 5, 6)$ to adhere to the defined dynamics as before. However, in the pseudo-daughter system, variables 2 and 5 are interchanged, resulting in the triplets $(1, 5, 3)$ and $(4, 2, 6)$ following the same dynamics. The corresponding networks are shown in Figure 4. Notably, both systems have the same level of complexity, with each variable regulated by the same number of regulators across the two systems. The mother cell trajectories, denoted by $x^{(M)}(t) \in \mathbb{R}^6$, are generated from the mother system over $n = 200$ evenly-spaced time points on the standardized interval $[0, 1]$ at a noise level of 1%. The daughter cell is assumed to be born at time $t = 0.3$. The pseudo-daughter trajectories, denoted as $x^{(D,\text{pseudo})}(t) \in \mathbb{R}^6$,

are generated on the interval $[0.3, 1]$ with initial conditions $x^{(D,\text{pseudo})}(0.3)$ set to $x^{(M)}(0.3)$ at its time of birth. Let α denote the weight of mother trajectories, with larger values indicating higher degrees of inheritance. The actual daughter trajectories $x^{(D)}(t) \in \mathbb{R}^6$ are then simulated as

$$x^{(D)}(t) = \alpha x^{(M)}(t) + (1-\alpha) x^{(D,\text{pseudo})}(t), \quad t \in [0.3, 1],$$

with data drawn on the same observation time grid, limited to the interval $[0.3, 1]$.

To assess if ODEinherit measures the inheritance metric accurately, we create a suite of simulated mother-daughter pairs where we vary the amount of inheritance. To do this, we generate $R = 20$ mother-daughter pairs for each $\alpha \in \{0, 0.1, 0.2, \ldots, 1\}$, as described above for both the additive and non-additive systems. Here, $\alpha = 0$ and $\alpha = 1$ denote no or full inheritance, and we are interested to see if ODEinherit can measure meaningful differences in inheritance between these two extremes. We employ the estimation strategy in Section 3.3 and obtain the inheritance measure $\pi^{(M \to D)}$ for each cell. We note here that α does not directly represent the percentage of inheritance, and therefore $\pi^{(M \to D)}$ is not a direct estimate of α. Additionally, due to the nature of our ODE estimation approach, even a random network can account for a nonzero portion of the variance in the trajectories. In an extreme case when no regulation exists at all (i.e., an empty network), the fitted functionals $\widehat{F}_j$ in (2) become constants, and the recovered trajectories are simply linear trends. Hence, these trends might still explain some variance in the observed trajectories, yielding a positive $\pi^{(M \to D)}$ inheritance measure even when no inheritance exists. To account for this phenomenon, we establish a baseline for how much a random mother network can explain the daughter trajectories. We do this by generating 10 random mother networks of equivalent complexity for each estimated mother network. These random networks are then used to refit the daughter trajectories and calculate the corresponding $\pi^{(M \to D)}$ inheritance metric.

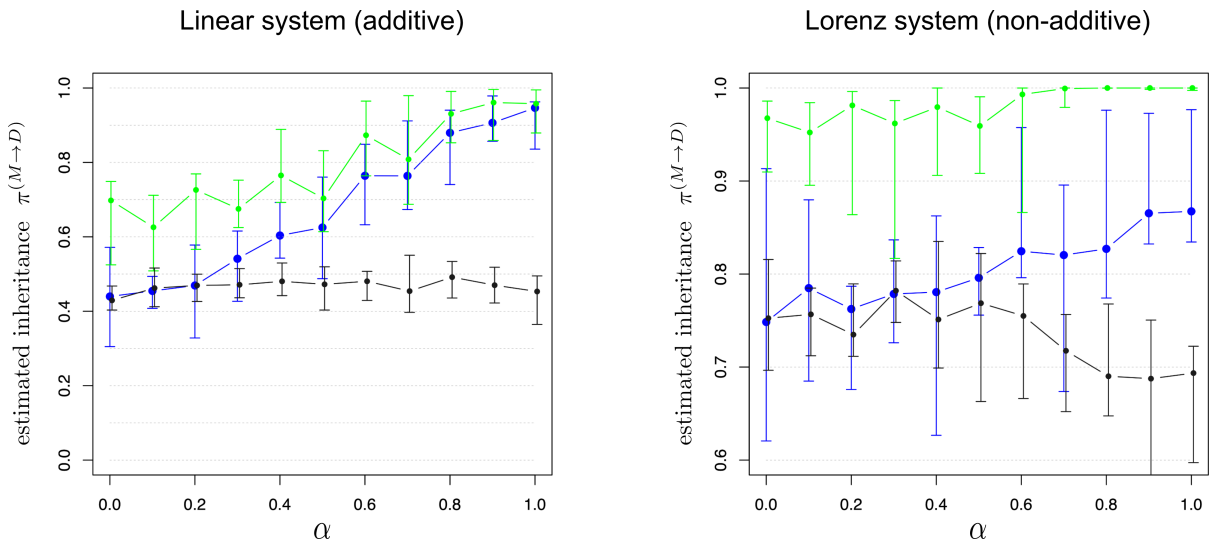

Figure 5: **A)** and **B)**: Estimated inheritance metric $\pi^{(M\to D)}$ against the weight of mother trajectories (α) in the additive and non-additive systems, respectively. Larger α values indicate higher degrees of inheritance. The green and blue curves show the measures calculated using networks before and after pruning, respectively. The black curve shows a baseline for how much a random mother network can explain the daughter trajectories by taking the input mother network to be random yet with equivalent complexity as the pruned mother network. Points indicate the median measures across simulations, and error bars represent the first and third quartiles.

Figure 5 shows that ODEinherit's protein regulatory inheritance is strongly correlated with the α, the amount of inheritance dictated in our simulation, when compared to alternative methods. This plot shows the inheritance metric $\pi^{(M\to D)}$ against α for each system, using the original, pruned, and random networks. We make three remarks about the results. First, the $\pi^{(M\to D)}$ estimate generally increases as α grows, indicating higher inheritance when the mother cell has more influence on the daughter cell. Notably, when there is no inheritance (i.e., $\alpha = 0$), the $\pi^{(M\to D)}$ measures are indistinguishable from those using random networks, validating ODEinherit's ability to detect a lack of inheritance. Second, even under model misspecification, the $\pi^{(M\to D)}$ estimate maintains an increasing trend with α but exhibits greater variability due to the reduced network estimation accuracy in this setting. The baseline is larger in this setting because the Lorenz system has a denser true network that inherently captures more variance. Third, without pruning, the $\pi^{(M\to D)}$ measures are largely inflated due to the high FPR in network estimation. The false positive edges inflate the $R^{2(M\to D)}$ values, which in turn leads to inflated $\pi^{(M\to D)}$ measures. The

issue becomes particularly severe in complex ODE systems under model misspecification, where $\pi^{(M\to D)}$ measures are all close to 1 regardless of the actual inheritance level (α), thereby obscuring meaningful inheritance quantification. Hence, the pruning procedure in ODEinherit is critical and helps correct this inflation, improving the accuracy of the inheritance metric $\pi^{(M\to D)}$.

5 Investigation of inheritance in yeast

We now return to the motivating data and analysis described in Section 2.1 and demonstrate how ODEinherit enhances our understanding of cellular dynamics.

5.1 Data details and preprocessing

We describe additional details about how we preprocessed the data before showcasing the results of ODEinherit. Our dataset includes 85 cells observed over 48 time points, consisting of 25 mother cells and 60 daughter cells. To filter out noise, we use a combination of Functional Principal Component Analysis (FPCA) and local polynomial regression to simultaneously smooth and interpolate the trajectories, achieving a fivefold resolution on the observed time grid. This analysis focuses exclusively on mother cells and first-generation daughter cells, as their longer time series provide more data for effective estimation. Due to heterogeneity across the cells, different mother cells may pass down different degrees of inheritance to their daughters. To address this, we again applied FPCA to all mother cells and grouped them into two main clusters using K-means clustering on their first five principal component scores. Figure 6A depicts the trajectories of a primary cell activity marker for each cluster. It is seen that cells in Cluster 1 exhibit fewer oscillations with larger amplitude, whereas cells in Cluster 2 undergo multiple cell cycles with greater consistency. We investigate the inheritance in each cluster separately. Further details are provided in Appendix S1.

5.2 Network estimation and inheritance measures

We apply the workflow described in Section 3 to each mother-daughter pair to infer their regulatory networks and calculate the corresponding inheritance measure. Since the true generating model is unknown, we use the first-order Matérn kernel with the same implementation as in Section 4.1.2, as it provides a more flexible function space to capture the complexity of the cell trajectories. We follow a similar strategy as in the preliminary analysis to validate that our method produces sensible inheritance measures for mother-daughter pairs. Specifically, each daughter cell is paired with 10 mother cells from the cluster that does not include its true mother cell, and the inheritance metrics are calculated on the "fake" pairs as a comparison.

Figure 6B shows histograms of the inheritance metrics for true and fake pairs in each cluster. We perform a Wilcoxon rank-sum test to test whether the median inheritance for true and fake pairs are identical, with the p-values displayed on the histograms. For highly variable mother cells (Cluster 1), the inheritance metric for true pairs are relatively low and close to those of fake pairs, suggesting weak inheritance. Conversely, for cells that undergo regular cell cycles (Cluster 2), the inheritance metric for true pairs are relatively high and clearly distinguishable at 0.05 level from those of fake pairs. These results indicate that the inheritance metric provided by our method are reasonable and biologically meaningful. We further demonstrate two examples of mother-daughter network pairs in Figure 6C, one with high inheritance and one with low inheritance. The difference between the mother and daughter networks is quantified using the Graph Edit Distance (GED), which is defined as the number of different edges between the networks. We find a negative correlation between GED and the inheritance measure ($\rho = -0.42$ for Cluster 1, $\rho = -0.20$ for Cluster 2). This relationship further supports the validity of our inheritance metrics, as smaller differences in the network structure are typically associated with stronger inheritance.

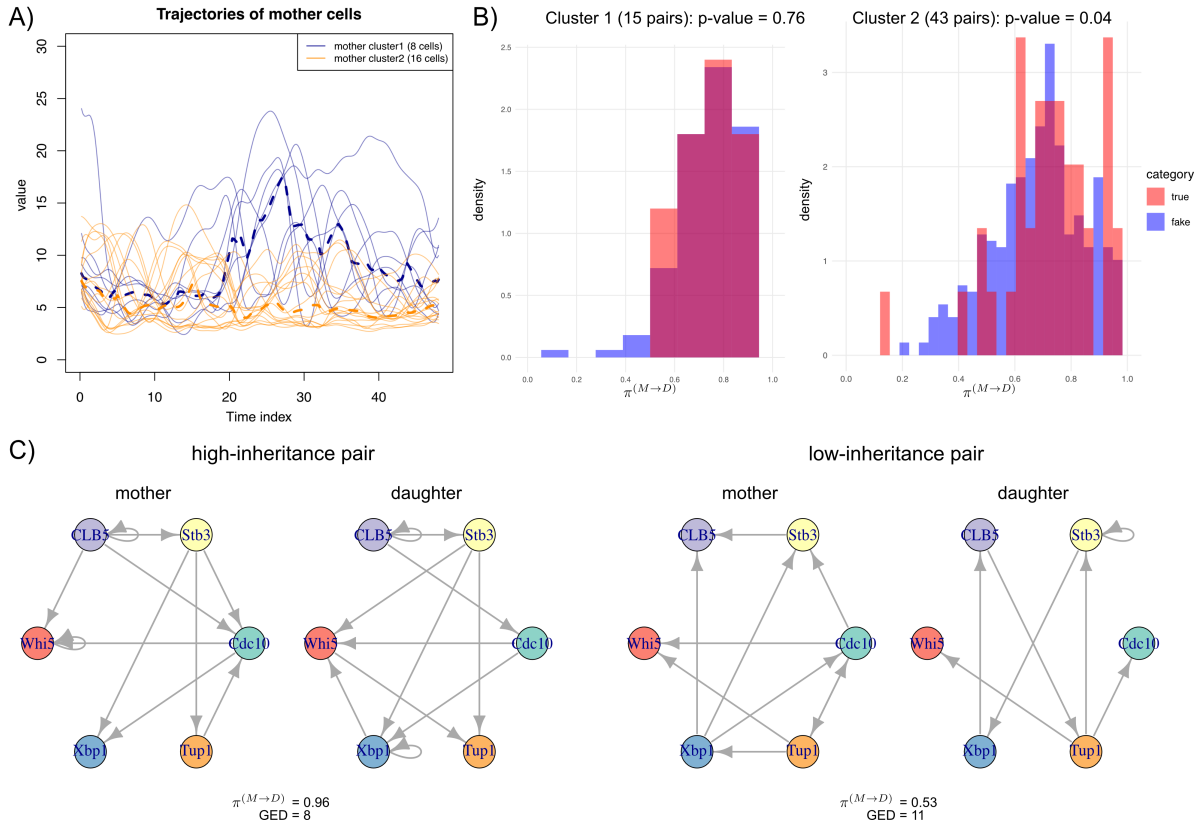

Figure 6: Inheritance metric in yeast cells. **A)**: The interpolated mother cells' trajectories of a primary cell activity marker in each cluster, with the median plotted as a dashed line. **B)**: Inheritance metric for mother-daughter pairs in each cluster, with p-values from the Wilcoxon rank-sum test. **C)**: Two examples of mother-daughter network pairs, one with a high inheritance metric and one with a low inheritance metric.

6 Conclusion

Understanding how regulatory information is transmitted across generations during cell divisions is a fundamental question in cell biology. Building on the hypothesis that a daughter cell inherits its regulatory machinery from its mother, we developed a novel statistical framework to quantify the extent of this inheritance. Our approach involves estimating an ODE system to model protein regulatory dynamics and measuring inheritance based on how effectively the mother's regulatory network predicts the daughter's trajectories. To ensure reliable results, our inheritance measure depends on a sparse regulatory network, which is achieved through a heuristic pruning procedure. Simulations validated the effectiveness of our method. We employed the method to investigate budding yeast cells

and revealed lineage-specific differences in inheritance rates, highlighting its potential to uncover biologically meaningful insights.

In this work, we primarily focused on ancestor cells and their first-generation daughter cells, as their longer observed trajectories provide more reliable data for modeling. Future research could explore how heritability influences cell fates across multiple generations, potentially uncovering long-term developmental trends in cell populations, as suggested by studies like (Mura et al., 2019). However, as later generations often exhibit shorter observed trajectories, additional statistical considerations will be needed to address these challenges. Another promising direction is to extend our framework to study cellular responses to experimental stimuli, leveraging ODE methods that have been adapted for such purposes (Dai and Li, 2022). This could potentially offer novel insights into how regulatory inheritance shifts under different stress conditions. These extensions underscore the versatility of our framework and highlight the need for continued development of statistical methodologies to harness the full potential of mother-daughter cell data.

Code availability

The code for ODEinherit is publicly available as R functions in `https://github.com/WenbinWu2001/ODEinherit`. For estimating the network, this codebase contains a further computationally optimized version of KODE, which was graciously provided by Lexin Li (which was originally in Matlab).

Acknowledgments

We thank Lexin Li for providing the original code for KODE, which we further adapted into R for this work. We also thank Ali Shojaie for further discussions that helped the ideas in this paper.

References

Argüello-Miranda, O., Liu, Y., Wood, N. E., Kositangool, P., and Doncic, A. (2018). Integration of multiple metabolic signals determines cell fate prior to commitment. *Molecular Cell*, 71(5):733–744.

Boczko, E. M., Gedeon, T., Stowers, C. C., and Young, T. R. (2010). ODE, RDE and SDE models of cell cycle dynamics and clustering in yeast. *Journal of Biological Dynamics*, 4(4):328–345.

Brunel, N. J.-B. (2008). Parameter estimation of ODE's via nonparametric estimators. *Electronic Journal of Statistics*, 2:1242 – 1267.

Cao, J. and Zhao, H. (2008). Estimating dynamic models for gene regulation networks. *Bioinformatics*, 24(14):1619–1624.

Chen, K. C., Calzone, L., Csikasz-Nagy, A., Cross, F. R., Novak, B., and Tyson, J. J. (2004). Integrative analysis of cell cycle control in budding yeast. *Molecular Biology of the Cell*, 15(8):3841–3862.

Chen, S., Shojaie, A., and Witten, D. M. (2017). Network reconstruction from high-dimensional ordinary differential equations. *Journal of the American Statistical Association*, 112(520):1697–1707.

Dai, X. and Li, L. (2022). Kernel ordinary differential equations. *Journal of the American Statistical Association*, 117(540):1711–1725.

Dattner, I. and Klaassen, C. A. J. (2015). Optimal rate of direct estimators in systems of ordinary differential equations linear in functions of the parameters. *Electronic Journal of Statistics*, 9(2):1939 – 1973.

Gneiting, T., Kleiber, W., and Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105(491):1167–1177.

Gugushvili, S. and Klaassen, C. A. (2012). $\sqrt{n}$-consistent parameter estimation for systems of ordinary differential equations: Bypassing numerical integration via smoothing. *Bernoulli*, 18(3):1061 – 1098.

Hall, P. and Ma, Y. (2014). Quick and easy one-step parameter estimation in differential equations. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 76(4):735–748.

Hamey, J. J. and Wilkins, M. R. (2023). The protein methylation network in yeast: A landmark in completeness for a eukaryotic post-translational modification. *Proceedings of the National Academy of Sciences*, 120(23):e2215431120.

Herrero, E., Stinus, S., Bellows, E., Berry, L. K., Wood, H., and Thorpe, P. H. (2020). Asymmetric transcription factor partitioning during yeast cell division requires the fact chromatin remodeler and cell cycle progression. *Genetics*, 216(3):701–716.

Higuchi-Sanabria, R., Pernice, W. M., Vevea, J. D., Alessi Wolken, D. M., Boldogh, I. R., and Pon, L. A. (2014). Role of asymmetric cell division in lifespan control in *Saccharomyces cerevisiae*. *FEMS Yeast Research*, 14(8):1133–1146.

Hindmarsh, A. C. (1983). ODEPACK, a systemized collection of ODE solvers. *Scientific Computing*.

Infant, T., Deb, R., Ghose, S., and Nagotu, S. (2021). Post-translational modifications of proteins associated with yeast peroxisome membrane: An essential mode of regulatory mechanism. *Genes to Cells*, 26(11):843–860.

Jashnsaz, H., Fox, Z. R., Munsky, B., and Neuert, G. (2021). Building predictive signaling models by perturbing yeast cells with time-varying stimulations resulting in distinct signaling responses. *Star Protocols*, 2(3):100660.

Kociemba, J., Jørgensen, A. C. S., Tadić, N., Harris, A., Sideri, T., Chan, W. Y., Ibrahim, F., Ünal, E., Skehel, M., Shahrezaei, V., et al. (2024). Multi-signal regulation of the GSK-3β homolog Rim11 controls meiosis entry in budding yeast. *The EMBO Journal*, pages 1–31.

Liang, H. and Wu, H. (2008). Parameter estimation for differential equation models using a framework of measurement error in regression models. *Journal of the American Statistical Association*, 103(484):1570–1583.

Lin, K. Z., Liu, H., and Roeder, K. (2021). Covariance-based sample selection for heterogeneous data: Applications to gene expression and autism risk gene detection. *Journal of the American Statistical Association*, 116(533):54–67.

Lin, Y. and Brown, L. D. (2004). Statistical properties of the method of regularization with periodic Gaussian reproducing kernel. *The Annals of Statistics*, 32(4):1723 – 1743.

Ma, C. and Gurkan-Cavusoglu, E. (2024). A comprehensive review of computational cell cycle models in guiding cancer treatment strategies. *NPJ Systems Biology and Applications*, 10(1):71.

Mukherjee, S. and Zhou, D.-X. (2006). Learning coordinate covariances via gradients. *Journal of Machine Learning Research*, 7(18):519–549.

Mura, M., Feillet, C., Bertolusso, R., Delaunay, F., and Kimmel, M. (2019). Mathematical modelling reveals unexpected inheritance and variability patterns of cell cycle parameters in mammalian cells. *PLoS Computational Biology*, 15(6):e1007054.

Otto, T. and Sicinski, P. (2017). Cell cycle proteins as promising targets in cancer therapy. *Nature Reviews Cancer*, 17(2):93–115.

Pomerening, J. R., Sontag, E. D., and Ferrell Jr, J. E. (2003). Building a cell cycle oscillator: Hysteresis and bistability in the activation of Cdc2. *Nature Cell Biology*, 5(4):346–351.

Qi, X. and Zhao, H. (2010). Asymptotic efficiency and finite-sample properties of the generalized profiling estimation of parameters in ordinary differential equations. *Annals of Statistics*, 38(1):435 – 481.

Radde, N. and Kaderali, L. (2009). Inference of an oscillating model for the yeast cell cycle. *Discrete Applied Mathematics*, 157(10):2285–2295.

Ramakanth, S., Kennedy, T., Yalcinkaya, B., Neupane, S., Tadic, N., Buchler, N. E., and Argüello-Miranda, O. (2024). Deep learning-driven imaging of cell division and cell growth across an entire eukaryotic life cycle. *bioRxiv*.

Ramsay, J. O., Hooker, G., Campbell, D., and Cao, J. (2007). Parameter estimation for differential equations: A generalized smoothing approach. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 69(5):741–796.

Shojaie, A. and Fox, E. B. (2022). Granger causality: A review and recent advances. *Annual Review of Statistics and Its Application*, 9(1):289–319.

Sible, J. C. and Tyson, J. J. (2007). Mathematical modeling as a tool for investigating cell cycle control networks. *Methods*, 41(2):238–247.

Soetaert, K. E., Petzoldt, T., and Setzer, R. W. (2010). Solving differential equations in R: Package deSolve. *Journal of Statistical Software*, 33(9).

Tyson, J. J. and Novák, B. (2015). Models in biology: Lessons from modeling regulation of the eukaryotic cell cycle. *BMC Biology*, 13:1–10.

Wu, H., Lu, T., Xue, H., and Liang, H. (2014). Sparse additive ordinary differential equations for dynamic gene regulatory network modeling. *Journal of the American Statistical Association*, 109(506):700–716.

Xue, H., Miao, H., and Wu, H. (2010). Sieve estimation of constant and time-varying coefficients in nonlinear ordinary differential equation models by considering both numerical error and measurement error. *Annals of Statistics*, 38(4):2351.

Yang, L., Lv, S., and Wang, J. (2016). Model-free variable selection in Reproducing Kernel Hilbert Space. *Journal of Machine Learning Research*, 17(82):1–24.

Zhang, X., Cao, J., and Carroll, R. J. (2015). On the selection of ordinary differential equation models with application to predator-prey dynamical models. *Biometrics*, 71(1):131–138.