Moment-Based Inference Predicts Bimodality in Transient Gene Expression

C. Zechner, J. Ruess, P. Krenn, S. Pelet, M. Peter, J. Lygeros, H. Koeppl

S.1 Population Dynamics and Extrinsic Variability

Consider a chemical reaction system of m different species in a constant random environment, which is given by a $\mathcal{D}$-valued random variable Z that is distributed according to the probability measure P_Z, where $\mathcal{D}$ is some set. Denote by $X(t) = (X_1(t), \ldots, X_m(t))$ the stochastic process that describes the time evolution of the number of molecules of the m species and by $p(x,t|z)$ the conditional probability $P(X(t) = x|Z = z)$, where $x = (x_1, \ldots, x_m) \in \mathcal{X}$ and $\mathcal{X} \subset \mathbb{N}^m$ is the set of all reachable states. The time evolution of the conditional probability distribution of this system then follows a conditional chemical master equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}p(x,t|z) = \sum_{k=1}^{K} -p(x,t|z)a_k(x,z) + p(x-\nu_k,t|z)a_k(x-\nu_k,z),$$

where $a_k(x,z),\ k = 1,\ldots,K$, are the propensity functions and $\nu_k,\ k = 1,\ldots,K$, the stoichiometric transition vectors of the K reactions of the system.

Multiplying both sides of the conditional CME by the i-th component x_i of the state vector x, summing over all states $x \in \mathcal{X}$ and integrating over all possible values of Z with respect to the probability measure P_Z yields:

$$\int_{\mathcal{D}} \sum_{x\in\mathcal{X}} \frac{\mathrm{d}}{\mathrm{d}t}p(x,t|z)x_i dP_Z = \int_{\mathcal{D}} \sum_{x\in\mathcal{X}} \sum_{k=1}^{K} [-p(x,t|z)a_k(x,z)x_i + p(x-\nu_k,t|z)a_k(x-\nu_k,z)x_i]dP_Z$$

The left hand side is then the time derivative of the marginal mean of X_i, whereas the right hand side can be simplified by a change of variables as follows:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[X_i(t)] &= \int_{\mathcal{D}} \sum_{k=1}^{K} \sum_{x\in\mathcal{X}} [-p(x,t|z)a_k(x,z)x_i + p(x,t|z)a_k(x,z)(x_i+\nu_{k_i})]dP_Z \\
&= \sum_{k=1}^{K} \int_{\mathcal{D}} \sum_{x\in\mathcal{X}} p(x,t|z)a_k(x,z)(-x_i + x_i + \nu_{k_i})dP_Z \\
&= \sum_{k=1}^{K} \nu_{k_i} \int_{\mathcal{D}} \sum_{x\in\mathcal{X}} p(x,t|z)a_k(x,z)dP_Z \\
&= \sum_{k=1}^{K} \nu_{k_i} \int_{\mathcal{D}} \mathbb{E}[a_k(X(t),z)|z]dP_Z \\
&= \sum_{k=1}^{K} \nu_{k_i} \mathbb{E}[a_k(X(t),Z)],
\end{aligned}$$

where ν_{k_i} is the i-th component of ν_k.

Multiplying both sides with $x_i x_j$ instead of x_i and performing the same calculations yields the time derivatives of the second order moments:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[X_i(t)X_j(t)] &= \int_{\mathcal{D}} \sum_{k=1}^{K} \sum_{x\in\mathcal{X}} \left[-p(x,t|z)a_k(x,z)x_i x_j + p(x,t|z)a_k(x,z)(x_i+\nu_{k_i})(x_j+\nu_{k_j})\right] dP_Z \\ &= \sum_{k=1}^{K} \int_{\mathcal{D}} \sum_{x\in\mathcal{X}} p(x,t|z)a_k(x,z)(-x_i x_j + x_i x_j + x_i\nu_{k_j} + x_j\nu_{k_i} + \nu_{k_i}\nu_{k_j}) dP_Z \\ &= \sum_{k=1}^{K} \left(\nu_{k_i}\mathbb{E}[X_j(t)a_k(X(t),Z)] + \nu_{k_j}\mathbb{E}[X_i(t)a_k(X(t),Z)] + \nu_{k_i}\nu_{k_j}\mathbb{E}[a_k(X(t),Z)]\right)\end{aligned}$$

Through the terms $\mathbb{E}[a_k(X(t),Z)]$ and $\mathbb{E}[X_i(t)a_k(X(t),Z)]$ cross moments of the species and the extrinsic variable influence the time evolution of the species moments. Even if the extrinsic variable Z is independent of the species initial conditions, over time it will become correlated with $X(t)$. This requires computing the time evolution of these correlations along with the species moments.

The time derivatives of the cross moments $\mathbb{E}[X_i(t)Z]$ can be obtained with the same calculations as above by multiplying both sides by $x_i z$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[X_i(t)Z] &= \int_{\mathcal{D}} \sum_{k=1}^{K} \sum_{x\in\mathcal{X}} [-p(x,t|z)a_k(x,z)x_i z + p(x,t|z)a_k(x,z)(x_i+\nu_{k_i})z] dP_Z \\ &= \sum_{k=1}^{K} \int_{\mathcal{D}} \sum_{x\in\mathcal{X}} p(x,t|z)a_k(x,z)(-x_i z + x_i z + \nu_{k_i} z) dP_Z \\ &= \sum_{k=1}^{K} \nu_{k_i}\mathbb{E}[Z a_k(X(t),Z)]\end{aligned}$$

If we close the moment equations and replace all the higher order moments of the species and all the higher order cross moments of the species and the extrinsic variable by functions of the lower order moments, we obtain a dynamical system for the moments and cross moments of order up to two that depends only on some lower order moments of the distribution of the extrinsic variable. If this distribution is unknown, these moments can be included as unknown parameters in the parameter search. The specific moment equations and closure functions for the examples considered in this work are given in Section S.3.1 and Section S.4.1.

S.2 Separating extrinsic from intrinsic noise in a birth-death process with variable birth rate

To demonstrate that, in principle, the extrinsic statistics can be identified from measurements of species moments we consider a simple birth-death process, initialized at zero, with death rate b, where the birth rate Z is a random variable distributed according to some probability measure P_Z.

$$\emptyset \xrightarrow{Z} X \xrightarrow{b} \emptyset$$

Using the derivation of the previous section the time evolution of the moments of this system can be computed as

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[X\right] &= \mathbb{E}\left[Z\right] - b\mathbb{E}\left[X\right], \\
\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[X^2\right] &= \mathbb{E}\left[Z\right] + 2\mathbb{E}\left[XZ\right] - 2b\mathbb{E}\left[X^2\right] + b\mathbb{E}\left[X\right], \\
\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[XZ\right] &= \mathbb{E}\left[Z^2\right] - b\mathbb{E}\left[XZ\right].
\end{aligned}$$

This is a closed linear system and can be solved analytically. Denote by $\mathbb{E}\left[X\right]^{tr}$ and $\mathbb{E}\left[X\right]^{ss}$ two observations of the mean of X, taken at a transient time point t_{tr} and at stationarity, respectively. These two observations allow to compute the death rate and the mean of P_Z as

$$b = -\frac{1}{t_{tr}}\log\left(\frac{\mathbb{E}\left[X\right]^{ss} - \mathbb{E}\left[X\right]^{tr}}{\mathbb{E}\left[X\right]^{ss}}\right) \quad \text{and} \quad \mathbb{E}\left[Z\right] = b\mathbb{E}\left[X\right]^{ss}.$$

If also an observation $\mathbb{E}\left[X^2\right]^{ss}$ of the stationary second moment is available, the second moment of P_Z can be computed as

$$\mathbb{E}\left[Z^2\right] = b^2\left(\mathbb{E}\left[X^2\right]^{ss} - \mathbb{E}\left[X\right]^{ss}\right)$$

Therefore, for this process, mean and variance of the extrinsic variable can be computed from measurements of the mean at one transient time point and mean and second moment at stationarity.

S.3 A Simple Model of Transient Gene Activation

The simple mass-action model under consideration is defined by the four reactions

$$\begin{aligned}
A &\xrightarrow{c_1} \emptyset \\
A + B &\underset{c_3}{\overset{c_2}{\rightleftharpoons}} AB \\
AB &\xrightarrow{c_4} AB + C,
\end{aligned} \tag{1}$$

and the corresponding stochastic rate constants given in Tab. S.1.

S.3.1 Moment Dynamics

Denote by μ^1_A, μ^1_B, μ^1_D and μ^1_C the means of A, B, AB and C, respectively. Let μ^2_A, μ^2_{AB}, μ^2_{AD}, μ^2_{AC}, μ^2_B, μ^2_{BD}, μ^2_{BC}, μ^2_D, μ^2_{DC}, μ^2_C and μ^3_{ABD}, μ^3_{A2B}, μ^3_{ABC}, μ^3_{AB2} be the second and third order moments, respectively. Since we assumed that this system is unaffected by extrinsic variability, the moment dynamics are not affected by moments of an extrinsic variable. Several tools are available, e.g. [7], that are able to automatically compute moment dynamics and closure functions from the reaction system and allow fast implementation of our identification method. For the system in question the moment equations become

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mu^1_A &= c_3 \cdot \mu^1_D - c_2 \cdot \mu^2_{AB} - c_1 \cdot \mu^1_A \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^1_B &= c_3 \cdot \mu^1_D - c_2 \cdot \mu^2_{AB} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^1_D &= c_2 \cdot \mu^2_{AB} - c_3 \cdot \mu^1_D \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^1_C &= c_4 \cdot \mu^1_D \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_A &= c_1 \cdot \mu^1_A - 2 \cdot c_1 \cdot \mu^2_A + c_2 \cdot \mu^2_{AB} - 2 \cdot c_2 \cdot \mu^3_{A2B} + c_3 \cdot \mu^1_D + 2 \cdot c_3 \cdot \mu^2_{AD} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{AB} &= c_3 \cdot \mu^1_D - c_2 \cdot \mu^3_{AB2} - c_2 \cdot \mu^3_{A2B} + c_3 \cdot \mu^2_{AD} + c_3 \cdot \mu^2_{BD} - (c_1 - c_2) \cdot \mu^2_{AB} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{AD} &= c_2 \cdot \mu^3_{A2B} - c_2 \cdot \mu^2_{AB} - c_2 \cdot \mu^3_{ABD} - c_3 \cdot \mu^1_D + c_3 \cdot \mu^2_D - (c_1 + c_3) \cdot \mu^2_{AD} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{AC} &= c_3 \cdot \mu^2_{DC} - c_2 \cdot \mu^3_{ABC} - c_1 \cdot \mu^2_{AC} + c_4 \cdot \mu^2_{AD} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_B &= c_2 \cdot \mu^2_{AB} - 2 \cdot c_2 \cdot \mu^3_{AB2} + c_3 \cdot \mu^1_D + 2 \cdot c_3 \cdot \mu^2_{BD} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{BD} &= c_2 \cdot \mu^3_{AB2} - c_2 \cdot \mu^2_{AB} - c_2 \cdot \mu^3_{ABD} - c_3 \cdot \mu^1_D + c_3 \cdot \mu^2_D - c_3 \cdot \mu^2_{BD} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{BC} &= c_3 \cdot \mu^2_{DC} - c_2 \cdot \mu^3_{ABC} + c_4 \cdot \mu^2_{BD} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_D &= c_2 \cdot \mu^2_{AB} + 2 \cdot c_2 \cdot \mu^3_{ABD} + c_3 \cdot \mu^1_D - 2 \cdot c_3 \cdot \mu^2_D \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{DC} &= c_2 \cdot \mu^3_{ABC} - c_3 \cdot \mu^2_{DC} + c_4 \cdot \mu^2_D \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_C &= c_4 \cdot \mu^1_D + 2 \cdot c_4 \cdot \mu^2_{DC}
\end{aligned}
$$

A closed system is obtained by replacing the third order cumulants by zero, which is equivalent to replacing the third order moments by functions of the lower order moments as follows:

$$
\begin{aligned}
\mu^3_{ABD} &= \mu^1_A \cdot \mu^2_{BD} + \mu^1_B \cdot \mu^2_{AD} + \mu^1_D \cdot \mu^2_{AB} - 2 \cdot \mu^1_A \cdot \mu^1_B \cdot \mu^1_D, \\
\mu^3_{A2B} &= -2 \cdot \mu^1_B \cdot {\mu^1_A}^2 + 2 \cdot \mu^2_{AB} \cdot \mu^1_A + \mu^1_B \cdot \mu^2_A, \\
\mu^3_{ABC} &= \mu^1_A \cdot \mu^2_{BC} + \mu^1_B \cdot \mu^2_{AC} + \mu^1_C \cdot \mu^2_{AB} - 2 \cdot \mu^1_A \cdot \mu^1_B \cdot \mu^1_C, \\
\mu^3_{AB2} &= -2 \cdot \mu^1_A \cdot {\mu^1_B}^2 + 2 \cdot \mu^2_{AB} \cdot \mu^1_B + \mu^1_A \cdot \mu^2_B
\end{aligned}
$$

Note that the accuracy of moment closure methods depends not only on the network structure, but also on the unknown parameters. Therefore, it is a priori unclear if a certain closure method leads to accurate approximations of the moments. However, for a given set of parameters accuracy of a moment closure method can be checked by comparing the approximations with estimates computed from a large number of stochastic simulations. Therefore we arbitrarily picked a moment closure method, performed the parameter search with it and checked the accuracy of the approximations locally for the found parameters. Second order zero cumulant moment closure led to good agreement of the approximations with estimates obtained from stochastically simulating the system with the estimated parameters and was therefore sufficient for our model.

S.3.2 Model Calibration

We first validated the normality assumption of the empirical moment estimates as described in the Materials and Methods section in the main text. As graphical method we applied standard probability-probability ($P-P$) and quantile-quantile ($Q-Q$) plots, whereas to quantitatively assess normality we performed the Kolmogorov-Smirnov-Lilliefors test [4]. Results and p-values for the tests are demonstrated for the protein distribution at 10000s in Fig. S.1.

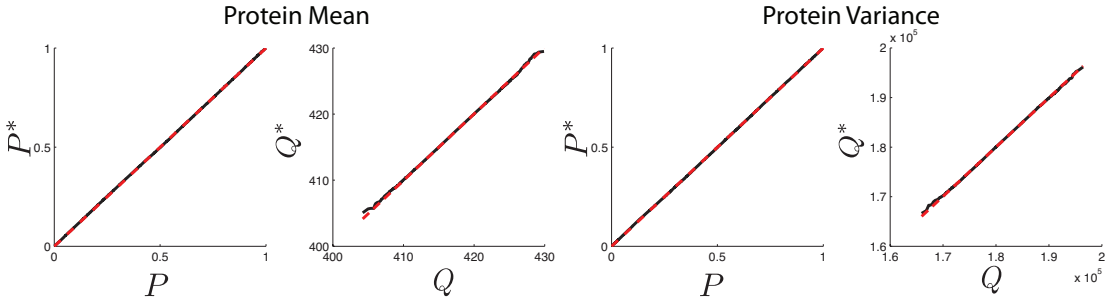

Figure S.1: Normality of empirical moments. Distributions of empirical moments were computed using bootstrapping. In particular, we randomly subsampled 10000 abundance values from the simulated target distribution and computed the sample mean and variance. This was repeated 50000 times to obtain approximations of the distributions of the moment estimates. For comparison, analytical normal distributions were fitted to the bootstrapped data and compared against each other using standard probability-probability ($P-P$) and quantile-quantile ($Q-Q$) plots [4]. In all four plots, the x-axes correspond to the analytical normal distributions, whereas the y-axes correspond to the bootstrapped empirical distributions. The low Kolmogorov-Smirnov test statistics D_K in conjunction with the large p-values indicate that the distribution of the empirical moment estimates can be well approximated by a normal distribution.

In order to calibrate the model to the reference data, we estimated the kinetic parameters from Tab. S.1. For simplicity, we assumed that the initial conditions of the species are known. If in some application the initial conditions are unknown, they can be included as unknown parameters in the parameter search and estimated along with the other parameters. For species A we set the initial amount to 50 molecules, B was initialized at 1 (corresponding to the gene being initially in the inactive state), whereas all other species were initialized at 0 molecules. The reference data was generated by stochastic simulation using $M = 20000$ sample paths of length $T = 10000s$. We computed first and second order moments of species C and their corresponding uncertainties (see main text, Materials and Methods) each $2000s$.

The obtained data was treated as measurements and used to calibrate the model by computing MAP estimates using a standard Metropolis-Hastings (M-H) sampler (see main text, Materials and Methods). The scaling parameters of the log-normal proposal densities were set to $v_j = 0.01$. After burn-in we recorded around 10000 samples of the M-H algorithm and determined MAP values. This inference was performed multiple times using random initial parameter values, drawn from a log-normal distribution $\mathcal{LN}(\ln 0.002, 2^2)$. Each time, the inference scheme ended up with equivalent MAP estimates (up to small random deviations, introduced by the randomized parameter search).

The inferred parameter set used for further analysis is given in Tab. S.1. The calibrated model was validated by comparing the distributions of species C to the distributions obtained from the reference model at the measurement time points (for better accuracy the comparison was based on 50000 sample paths for both models), as shown in Fig. S.2.

Table S.1: Inferred model parameters and Metropolis-Hastings setup. The reference and inferred MAP estimates are denoted as γ_j and $\gamma_{j,MAP}$, respectively.

Parameter	c_1	c_2	c_3	c_4
γ_j	$1.500 \cdot 10^{-2}$	$8.000 \cdot 10^{-4}$	$1.000 \cdot 10^{-3}$	$4.000 \cdot 10^{-1}$
$\gamma_{j,MAP}$	$1.380 \cdot 10^{-2}$	$7.050 \cdot 10^{-4}$	$9.865 \cdot 10^{-4}$	$3.988 \cdot 10^{-1}$
Unit	s^{-1}	s^{-1}	s^{-1}	s^{-1}

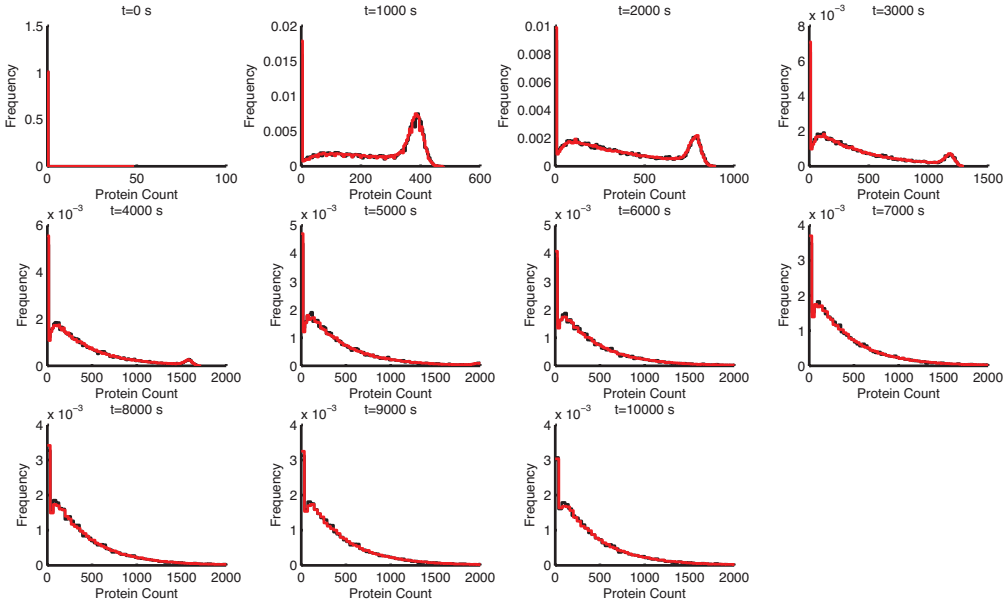

Figure S.2: Protein distributions for each measurement time point. Red: Calibrated model; Black: Reference model.

S.3.3 Second Order Moment Resolves Non-Identifiability

For the simple model of transient gene activation, the population mean alone does not provide enough information to uniquely determine the four model parameters. This is demonstrated in Fig. S.3, where we compared two different model configurations, each of them found by running the same MCMC algorithm with different initial conditions. In both cases, the estimated means fit well the reference mean. In contrast, the variances significantly differ from each other. Neither parameter set can reproduce the underlying distribution. For the second parameter configuration, the distribution is even unimodal (see Fig S.4).

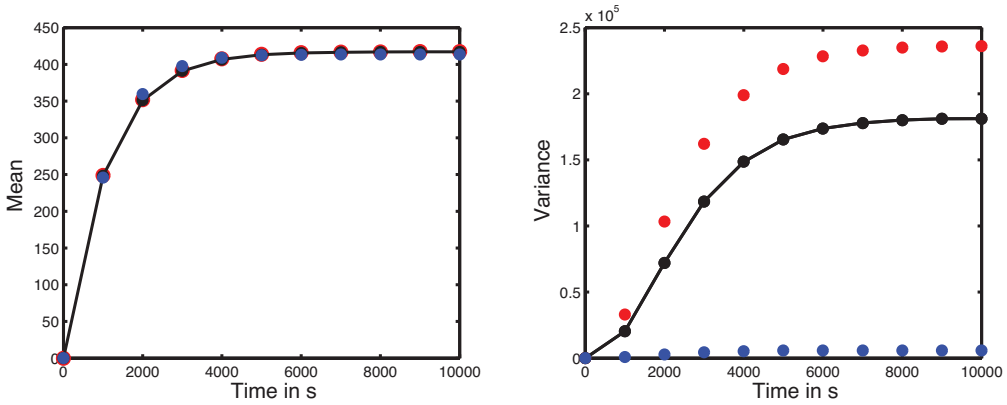

Figure S.3: The plot shows results for two parameter configurations that achieved almost equivalent mean values but strongly differ in the variance. Red, Blue: Calibrated models; Black: Reference model

S.4 Osmo-Stress induced MAPK Hog1 Activity and pSTL1-qV Protein Expression

To keep the notation simple, we use the following acronyms for the model species:

Table S.2: Acronyms for the chemical species.

Species	$pSTL1^{off}$	$pSTL1^{on}$	CR	$pSTL1^{on} \cdot CR$	RIB	$mRNA$	$pSTL1-qV$
Acronym	A	B	C	D	E	F	G

The Hog1 induced gene expression model is given by the reaction network

$$\begin{aligned} A &\underset{c_2}{\overset{c_1(t)}{\rightleftharpoons}} B \\ B + C &\underset{c_4}{\overset{c_3}{\rightleftharpoons}} D \\ D &\xrightarrow{c_5} D + F \\ F + E &\xrightarrow{c_6} F + E + G \\ G &\xrightarrow{c_7} \emptyset \\ F &\xrightarrow{c_8} \emptyset. \end{aligned} \tag{2}$$

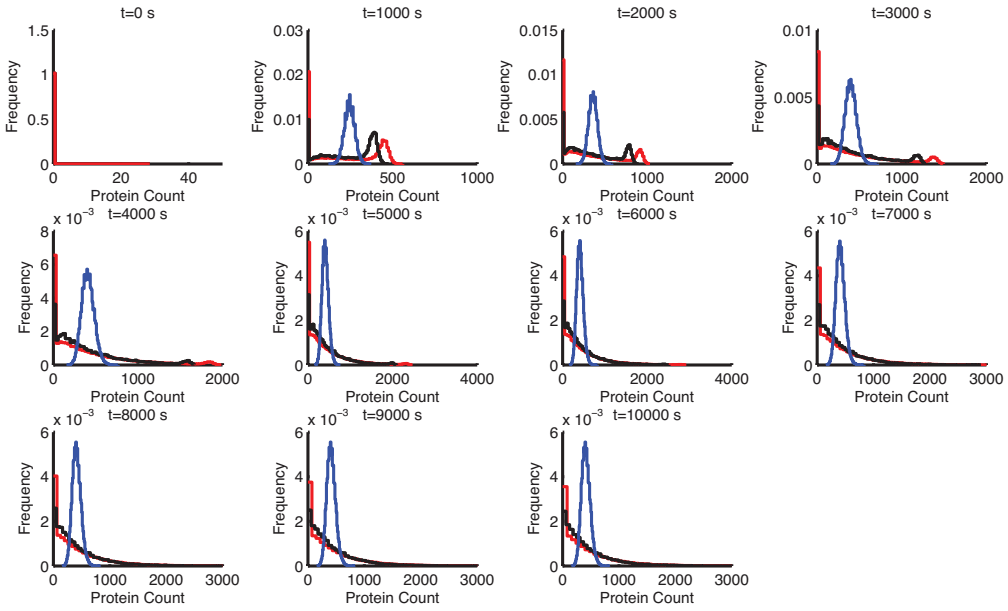

Figure S.4: Protein distributions for two parameter configurations. As the blue distribution is unimodal for all time points, it follows that the multi-modality cannot be predicted from the protein mean alone. Red, Blue: Calibrated models; Black: Reference model.

Note that $c_1(t)$ denotes a time-varying kinetic parameter as described in Section S.4.2. Note that the synthesis reaction for protein G can equivalently be written as the standard first-order reaction $F \xrightarrow{c_6 E} F + G$.

S.4.1 Moment Dynamics

Denote by μ^1_A, μ^1_B, μ^1_C, μ^1_D, μ^1_E, μ^1_F, and μ^1_G the means of A, B, C, D, E, F and G, respectively. Let μ^2_A, μ^2_{AB}, μ^2_{AC}, μ^2_{AD}, μ^2_{AE}, μ^2_{AF}, μ^2_{AG}, μ^2_B, μ^2_{BC}, μ^2_{BD}, μ^2_{BE}, μ^2_{BF}, μ^2_{BG}, μ^2_C, μ^2_{CD}, μ^2_{CE}, μ^2_{CF}, μ^2_{CG}, μ^2_D, μ^2_{DE}, μ^2_{DF}, μ^2_{DG}, μ^2_E, μ^2_{EF}, μ^2_{EG}, μ^2_F, μ^2_{FG}, μ^2_G and μ^3_{ABC}, μ^3_{AEF}, μ^3_{B2C}, μ^3_{BCD}, μ^3_{BCE}, μ^3_{BC2}, μ^3_{BCF}, μ^3_{BCG}, μ^3_{E2F}, μ^3_{DEF}, μ^3_{EFG}, μ^3_{EF2}, μ^3_{CEF}, μ^3_{BEF} be the second and third order moments, respectively. Then the moment equations for this system are given by:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{1}_{A} &= c_2 \cdot \mu^{1}_{B} - c_1 \cdot \mu^{1}_{A} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{1}_{B} &= c_1 \cdot \mu^{1}_{A} - c_2 \cdot \mu^{1}_{B} - c_3 \cdot \mu^{2}_{BC} + c_4 \cdot \mu^{1}_{D} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{1}_{C} &= c_4 \cdot \mu^{1}_{D} - c_3 \cdot \mu^{2}_{BC} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{1}_{D} &= c_3 \cdot \mu^{2}_{BC} - c_4 \cdot \mu^{1}_{D} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{1}_{E} &= 0 \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{1}_{F} &= c_5 \cdot \mu^{1}_{D} - c_8 \cdot \mu^{1}_{F} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{1}_{G} &= c_6 \cdot \mu^{2}_{EF} - c_7 \cdot \mu^{1}_{G} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{A} &= c_1 \cdot \mu^{1}_{A} - 2 \cdot c_1 \cdot \mu^{2}_{A} + c_2 \cdot \mu^{1}_{B} + 2 \cdot c_2 \cdot \mu^{2}_{AB} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{AB} &= c_1 \cdot \mu^{2}_{A} - c_1 \cdot \mu^{1}_{A} - c_2 \cdot \mu^{1}_{B} + c_2 \cdot \mu^{2}_{B} - c_3 \cdot \mu^{3}_{ABC} + c_4 \cdot \mu^{2}_{AD} - (c_1 + c_2) \cdot \mu^{2}_{AB} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{AC} &= c_2 \cdot \mu^{2}_{BC} - c_1 \cdot \mu^{2}_{AC} - c_3 \cdot \mu^{3}_{ABC} + c_4 \cdot \mu^{2}_{AD} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{AD} &= c_2 \cdot \mu^{2}_{BD} + c_3 \cdot \mu^{3}_{ABC} - (c_1 + c_4) \cdot \mu^{2}_{AD} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{AE} &= c_2 \cdot \mu^{2}_{BE} - c_1 \cdot \mu^{2}_{AE} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{AF} &= c_2 \cdot \mu^{2}_{BF} + c_5 \cdot \mu^{2}_{AD} - (c_1 + c_8) \cdot \mu^{2}_{AF} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{AG} &= c_2 \cdot \mu^{2}_{BG} + c_6 \cdot \mu^{3}_{AEF} - (c_1 + c_7) \cdot \mu^{2}_{AG} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{B} &= c_1 \cdot \mu^{1}_{A} + 2 \cdot c_1 \cdot \mu^{2}_{AB} + c_2 \cdot \mu^{1}_{B} - 2 \cdot c_2 \cdot \mu^{2}_{B} + c_3 \cdot \mu^{2}_{BC} - 2 \cdot c_3 \cdot \mu^{3}_{B2C} + c_4 \cdot \mu^{1}_{D} + 2 \cdot c_4 \cdot \mu^{2}_{BD} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{BC} &= c_1 \cdot \mu^{2}_{AC} - c_3 \cdot \mu^{3}_{B2C} - c_3 \cdot \mu^{3}_{BC2} + c_4 \cdot \mu^{1}_{D} + c_4 \cdot \mu^{2}_{CD} + c_4 \cdot \mu^{2}_{BD} - (c_2 - c_3) \cdot \mu^{2}_{BC} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{BD} &= c_1 \cdot \mu^{2}_{AD} - c_3 \cdot \mu^{2}_{BC} + c_3 \cdot \mu^{3}_{B2C} - c_3 \cdot \mu^{3}_{BCD} - c_4 \cdot \mu^{1}_{D} + c_4 \cdot \mu^{2}_{D} - (c_2 + c_4) \cdot \mu^{2}_{BD} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{BE} &= c_1 \cdot \mu^{2}_{AE} - c_2 \cdot \mu^{2}_{BE} - c_3 \cdot \mu^{3}_{BCE} + c_4 \cdot \mu^{2}_{DE} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{BF} &= c_1 \cdot \mu^{2}_{AF} - c_3 \cdot \mu^{3}_{BCF} + c_4 \cdot \mu^{2}_{DF} + c_5 \cdot \mu^{2}_{BD} - (c_2 + c_8) \cdot \mu^{2}_{BF} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{BG} &= c_1 \cdot \mu^{2}_{AG} - c_3 \cdot \mu^{3}_{BCG} + c_4 \cdot \mu^{2}_{DG} + c_6 \cdot \mu^{3}_{BEF} - (c_2 + c_7) \cdot \mu^{2}_{BG} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{2}_{C} &= c_3 \cdot \mu^{2}_{BC} - 2 \cdot c_3 \cdot \mu^{3}_{BC2} + c_4 \cdot \mu^{1}_{D} + 2 \cdot c_4 \cdot \mu^{2}_{CD}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ CD} &= c_3 \cdot \mu^3_{BC2} - c_3 \cdot \mu^2_{BC} - c_3 \cdot \mu^3_{BCD} - c_4 \cdot \mu^1_D + c_4 \cdot \mu^2_D - c_4 \cdot \mu^2_{CD} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ CE} &= c_4 \cdot \mu^2_{DE} - c_3 \cdot \mu^3_{BCE} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ CF} &= c_4 \cdot \mu^2_{DF} - c_3 \cdot \mu^3_{BCF} + c_5 \cdot \mu^2_{CD} - c_8 \cdot \mu^2_{CF} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ CG} &= c_4 \cdot \mu^2_{DG} - c_3 \cdot \mu^3_{BCG} + c_6 \cdot \mu^3_{CEF} - c_7 \cdot \mu^2_{CG} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ D} &= c_3 \cdot \mu^2_{BC} + 2 \cdot c_3 \cdot \mu^3_{BCD} + c_4 \cdot \mu^1_D - 2 \cdot c_4 \cdot \mu^2_D \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ DE} &= c_3 \cdot \mu^3_{BCE} - c_4 \cdot \mu^2_{DE} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ DF} &= c_3 \cdot \mu^3_{BCF} + c_5 \cdot \mu^2_D - (c_4 + c_8) \cdot \mu^2_{DF} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ DG} &= c_3 \cdot \mu^3_{BCG} + c_6 \cdot \mu^3_{DEF} - (c_4 + c_7) \cdot \mu^2_{DG} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ E} &= 0 \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ EF} &= c_5 \cdot \mu^2_{DE} - c_8 \cdot \mu^2_{EF} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ EG} &= c_6 \cdot \mu^3_{E2F} - c_7 \cdot \mu^2_{EG} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ F} &= c_5 \cdot \mu^1_D + 2 \cdot c_5 \cdot \mu^2_{DF} + c_8 \cdot \mu^1_F - 2 \cdot c_8 \cdot \mu^2_F \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ FG} &= c_5 \cdot \mu^2_{DG} - (c_8 + c_7) \cdot \mu^2_{FG} + c_6 \cdot \mu^3_{EF2} \\
\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\ 2}_{\ G} &= c_6 \cdot \mu^2_{EF} + 2 \cdot c_6 \cdot \mu^3_{EFG} + c_7 \cdot \mu^1_G - 2 \cdot c_7 \cdot \mu^2_G
\end{aligned}
$$

A closed system is then obtained by replacing the third order cumulants by zero, which is equivalent to replacing the third order moments by functions of the lower order moments as follows:

$$
\begin{aligned}
\mu^3_{ABC} &= \mu^1_A \cdot \mu^2_{BC} + \mu^1_B \cdot \mu^2_{AC} + \mu^1_C \cdot \mu^2_{AB} - 2 \cdot \mu^1_A \cdot \mu^1_B \cdot \mu^1_C, \\
\mu^3_{AEF} &= \mu^1_A \cdot \mu^2_{EF} + \mu^1_E \cdot \mu^2_{AF} + \mu^1_F \cdot \mu^2_{AE} - 2 \cdot \mu^1_A \cdot \mu^1_E \cdot \mu^1_F, \\
\mu^3_{B2C} &= -2 \cdot \mu^1_C \cdot {\mu^1_B}^2 + 2 \cdot \mu^2_{BC} \cdot \mu^1_B + \mu^1_C \cdot \mu^2_B, \\
\mu^3_{BCD} &= \mu^1_B \cdot \mu^2_{CD} + \mu^1_C \cdot \mu^2_{BD} + \mu^1_D \cdot \mu^2_{BC} - 2 \cdot \mu^1_B \cdot \mu^1_C \cdot \mu^1_D, \\
\mu^3_{BCE} &= \mu^1_B \cdot \mu^2_{CE} + \mu^1_C \cdot \mu^2_{BE} + \mu^1_E \cdot \mu^2_{BC} - 2 \cdot \mu^1_B \cdot \mu^1_C \cdot \mu^1_E, \\
\mu^3_{BC2} &= -2 \cdot \mu^1_B \cdot {\mu^1_C}^2 + 2 \cdot \mu^2_{BC} \cdot \mu^1_C + \mu^1_B \cdot \mu^2_C, \\
\mu^3_{BCF} &= \mu^1_B \cdot \mu^2_{CF} + \mu^1_C \cdot \mu^2_{BF} + \mu^1_F \cdot \mu^2_{BC} - 2 \cdot \mu^1_B \cdot \mu^1_C \cdot \mu^1_F, \\
\mu^3_{BCG} &= \mu^1_B \cdot \mu^2_{CG} + \mu^1_C \cdot \mu^2_{BG} + \mu^1_G \cdot \mu^2_{BC} - 2 \cdot \mu^1_B \cdot \mu^1_C \cdot \mu^1_G, \\
\mu^3_{E2F} &= -2 \cdot \mu^1_F \cdot {\mu^1_E}^2 + 2 \cdot \mu^2_{EF} \cdot \mu^1_E + \mu^1_F \cdot \mu^2_E, \\
\mu^3_{DEF} &= \mu^1_D \cdot \mu^2_{EF} + \mu^1_E \cdot \mu^2_{DF} + \mu^1_F \cdot \mu^2_{DE} - 2 \cdot \mu^1_D \cdot \mu^1_E \cdot \mu^1_F, \\
\mu^3_{EFG} &= \mu^1_E \cdot \mu^2_{FG} + \mu^1_F \cdot \mu^2_{EG} + \mu^1_G \cdot \mu^2_{EF} - 2 \cdot \mu^1_E \cdot \mu^1_F \cdot \mu^1_G, \\
\mu^3_{EF2} &= -2 \cdot \mu^1_E \cdot {\mu^1_F}^2 + 2 \cdot \mu^2_{EF} \cdot \mu^1_F + \mu^1_E \cdot \mu^2_F, \\
\mu^3_{CEF} &= \mu^1_C \cdot \mu^2_{EF} + \mu^1_E \cdot \mu^2_{CF} + \mu^1_F \cdot \mu^2_{CE} - 2 \cdot \mu^1_C \cdot \mu^1_E \cdot \mu^1_F, \\
\mu^3_{BEF} &= \mu^1_B \cdot \mu^2_{EF} + \mu^1_E \cdot \mu^2_{BF} + \mu^1_F \cdot \mu^2_{BE} - 2 \cdot \mu^1_B \cdot \mu^1_E \cdot \mu^1_F.
\end{aligned}
$$

As in the example of Section S.3, second order zero cumulant moment closure led to good agreement of the approximations with estimates obtained from stochastically simulating the system with the found parameters and was therefore sufficient for this model.

S.4.2 Nuclear Hog1 Activity as Model Input

We used fluorescence microscopy to measure nuclear enrichment of MAPK Hog1. The data set was recorded by co-authors, originally published in [9]. To extract the time evolution from the single-cell measurements (around 50-120 cells), we first removed outliers manually and normalized the data. In particular, we divided nuclear Hog1 (obtained by image segmentation) by the total Hog1 in the cell to reduce photo-bleaching artifacts. Note that the relocation data measures active and also inactive Hog1, whereas the latter cannot bind to the target gene. Thus, we also subtracted the mean basal levels, calculated from time-points before osmotic stress. The basal active Hog1 level - denoted u_0 is then inferred during model calibration, together with the other parameters. Note that this procedure is based on the assumption that export of nuclear Hog1 is fast once it has been deactivated. The resulting data was then used to perform a linear regression with radial basis functions (RBFs). The base functions were centered around the measurement time points and the width parameter was adjusted according to the increasing measurement intervals. This procedure was repeated for each salt concentration. Note that to properly capture the stress-dependent protein expression, we additionally measured the STL1 reporter for intermediate salt concentrations in the flow cytometry experiments. The corresponding deterministic Hog1 functions were obtained by performing a linear interpolation along the NaCl-time plane (see Fig S.5).

In the following, the extracted nuclear Hog1 function for a given salt concentration is denoted as $u(t)$. To obtain the corresponding time-varying gene activation intensity, we transformed the Hog1 abundance using a Hill-function, i.e.,

$$
c_1(t) = \frac{V_{max}(u(t)+u_0)^{n_H}}{K_d^{n_H} + (u(t)+u_0)^{n_H}}, \tag{3}
$$

with u_0, n_H, V_{max} and K_d as unknown parameters, estimated during model calibration.

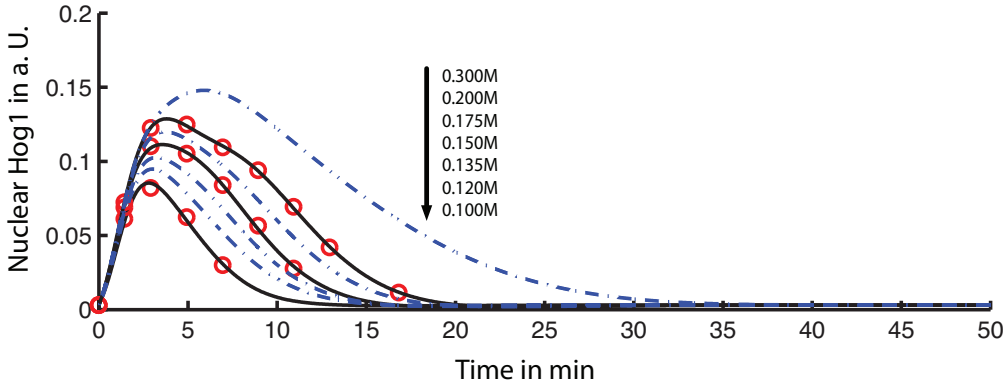

Figure S.5: Extracted nuclear Hog1 enrichment over time. Black: Fitted using linear RBF regression; blue: interpolated; red: mean abundance obtained from microscopy.

S.4.3 Extrinsic Variability

We model extrinsic variability at two different stages of the proposed model. First, we assume variability in chromatin remodeling, because it depends on a variety of different complexes which might be subject to cell-to-cell variations (such as RSC or the SAGA complex, [8]). Chromatin remodeling is modeled by recruitment of species C by the active gene B. Thus, variability in the total number of CR ($Z_1 = [C]+[D]$) leads to variability in the remodeling efficiency. Furthermore, we assume heterogeneity in the translation efficiency, which is reflected by a variability in the proxy species $Z_2 = [E]$. Mean and covariance matrix of the random vector $Z = [Z_1, Z_2]^T$ are defined as

$$\mathbb{E}[Z] = \begin{pmatrix} \alpha_1^1 \\ \alpha_2^1 \end{pmatrix} \tag{4}$$

and

$$\mathbb{E}\left[(Z - \mathbb{E}[Z])(Z - \mathbb{E}[Z])^T\right] = \begin{pmatrix} \alpha_{11}^2 & \alpha_{12}^2 \\ \alpha_{12}^2 & \alpha_{22}^2 \end{pmatrix} \tag{5}$$

The extrinsic statistics then enter the moment equations in terms of the initial conditions for the means $\mathbb{E}[C+D](0) = \mathbb{E}[C](0) = \alpha_1^1$ (as initially all CR is unbound), $\mathbb{E}[E](0) = \alpha_2^1$, the variances $Var(C+D)(0) = Var(C)(0) = \alpha_{11}^2$, $Var(E)(0) = \alpha_{22}^2$ and the covariance $Cov(C+D, E)(0) = Cov(C, E)(0) = \alpha_{12}^2$ and remain constant for all times, since $[C] + [D]$ and $[E]$ are conserved. The extrinsic statistics were assumed to be unknown and inferred from the measurements as explained in Section S.4.5. Note that within a moment-based approach, no assumptions on the distribution P_Z are required. However, once a comparison between the protein distributions is desired, realizations z have to be drawn from P_Z for each SSA run. As P_Z is not fully characterized by mean and variance only, further assumptions need to be made. Here, we restrict the shape of P_Z to be log-normal [3] and compute it's parameters from the inferred extrinsic statistics (i.e., first and second order moments).

S.4.4 Flow Cytometry Data for pSTL1-qV Reporter

The time-lapse pSTL1-qV distribution dataset (FC;TL) was measured using flow cytometry at different time points between zero and 48 minutes in six-minute intervals. Salt was added to the media straight after time $t = 0min$ by hand, which yielded a slightly delayed rather than an immediate exposition to the desired salt level. We accounted for this by correcting the later time points by $-3min$, and

consequently, we obtained the modified measurement time points $0min$, $3min$, $9min$,...,$45min$. We observed a strong bimodality in the logarithmic forward scatter (FSC) and side scatter (SSC) plots, indicating the presence of two phenotypic subpopulations. Consequently, we applied a weak cell gate on the FSC and SSC channel to exclude one of the two subpopulations. In analogy, this procedure was performed on the additional snapshot dataset (FC;SN).

S.4.5 Model Calibration

Analogous to Section S.3.2, we first validated the normality assumption of the empirical moment as described in the Materials and Methods section in the main text. As graphical method we applied standard probability-probability $(P-P)$ and quantile-quantile $(Q-Q)$ plots whereas to quantitatively assess normality we performed the Kolmogorov-Smirnov-Lilliefors test. The results and the corresponding p-values of the normality tests are shown in Fig. S.6. The described analysis was performed on the pSTL1-qV distribution for $0.1M$ NaCl at time $45min$, where the measured distribution is bimodal.

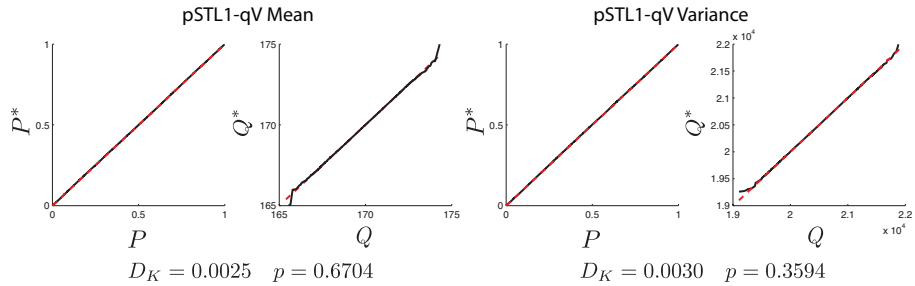

Figure S.6: Normality of empirical moments. Distributions of empirical moments were computed using bootstrapping. In particular, we randomly subsampled 10000 abundance values from the pSTL1-qV distribution and computed the sample mean and variance. This was repeated 50000 times to obtain approximations of the distributions of the moment estimates. For comparison, analytical normal distributions were fitted to the bootstrapped data and compared against each other using standard probability-probability $(P-P)$ and quantile-quantile $(Q-Q)$ plots. In all four plots, the x-axes correspond to the analytical normal distributions, whereas the y-axes correspond to the bootstrapped empirical distributions. The low Kolmogorov-Smirnov test statistics D_K in conjunction with the large p-values indicate that the distribution of the empirical moment estimates can be well approximated by a normal distribution.

The model comprises the parameters u_0, n_H, K_d, V_{max}, c_2 ,c_3, c_4, c_5, c_6, c_7 , c_8, α_1^1, α_2^1, α_{11}^2, α_{22}^2 and α_{12}^2. Note that the propensity of a translation event is proportional to the product $c_6 \cdot [E]$. Thus, the parameters c_6, α_2^1 and α_{22}^2 are structurally unidentifiable [1, 6]; hence we estimated the products $c_6\alpha_2^1$ and $c_6^2\alpha_{22}^2$. Given those products, statistics of the number of ribosomes could for instance be estimated by setting c_6 to values from literature and quantifying the remaining part.

Due to the high-dimensional state and parameter space and the fact that our flow cytometry experiments only captured distributions of a single protein, we expected the inference problem to be non-convex (i.e., characterized by a multi-modal posterior distribution). For this reason, a M-H scheme analogous to Section S.3.2 was performed 50 times, each time with randomly drawn initial parameter configurations. Note that in general, this is likely to give parameter configurations for which the nonlinear moment system is numerically unstable and thus, we first selected a stable parameter configuration γ_s (see Tab. S.3). Then, each parameter value was randomly initialized around those initial parameter values, i.e., $\gamma_j^0 \sim \mathcal{LN}(\ln \gamma_{j,s}, 0.5^2)$. The scaling parameters v_j of the log-normal proposal densities are given in Tab. S.3. We then sorted the 50 parameter sets according to their maximum a-posteriori probabilities and selected the best five parameter sets for further inspection.

To quantitively assess the goodness-of-fit we made use of a distance measure between probability distributions. Among the various metrics proposed in literature (see [5] for an overview) we chose the uniform or Kolmogorov metric (measuring the maximum deviation between the cumulative distributions). As it is scale-invariant and bounded by one it appears more amenable to interpretation than others distance measures. Moreover, it was already used in the context of stochastic simulations of chemical kinetics [2]. Thus, for each of the above parameter sets we computed the Kolmogorov distance for all NaCl concentrations and time-points between the empirical and the predicted pSTL1-qV distributions. The final parameter set was selected such as to minimize the total distance, i.e., the sum of individual Kolmogorov distances over all time-points and the three concentrations used for fitting (i.e., 0M, 0.12M and 0.2M of NaCl). Fig. S.7 depicts the Kolmogorov distance for the best performing parameter set which is given in Tab. S.3. The results suggest that the model provides a reasonable approximation of the experimental data; indeed we are not aware of any study in the literature reporting comparable accuracy in predicting time lapsed distribution data. We believe that a further decrease in distance D_K may be achievable by adding complexity to the transcriptional and translational steps. However, in the absence of any further experimental readout, such extensions are likely to introduce new structural identifiability problems.

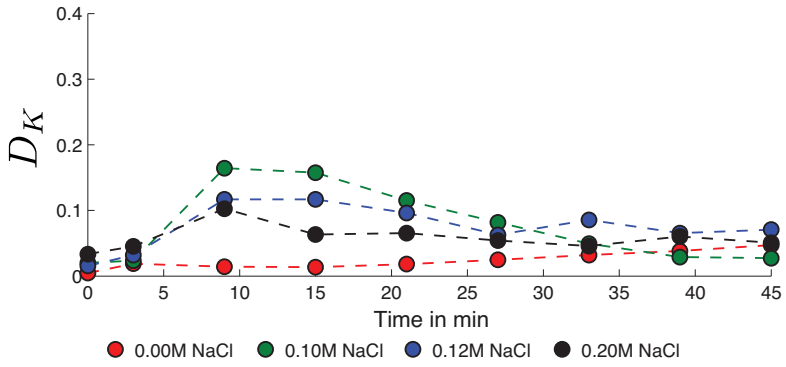

Figure S.7: Distance in distribution between model predictions and data. The Kolmogorov distance D_K was computed between the empirical and the predicted pSTL1-qV distribution for each time point and concentration. Note that the distance D_K is bounded by one.

S.4.6 Model Validation

Besides validating the predicted distributions with the time-lapsed dataset (FC;TL) for an intermediate NaCl concentration (i.e., 0.1M), we further validated the inferred model using the additional snapshot dataset (FC;SN). In particular, we computed the coefficient of variation and the percentage of protein producing cells at $t = 45min$ for different salt concentrations between $0M$ and $0.3M$. The former was computed as

$$CV = \frac{\hat{\sigma}_F}{\hat{\mu}_F}, \tag{6}$$

where $\hat{\mu}_F$ and $\hat{\sigma}_F$ are the empirical means and standard deviations obtained from the protein distributions. For comparison with the experimental data, we added the estimated mean and variance of the autofluorescence intensity to the means and variances of the simulated distributions. To estimate the number of responding cells for the different concentrations (dose-response) from the experimental data, we removed the autofluorescence by performing a deconvolution between the measured flow cytometry distributions and the autofluorescence distribution (obtained by flow cytometry for 0M NaCl, see main text, Materials and Methods). Subsequently, the percentage of non-expressing cells was given by the

Table S.3: Inferred model parameters and Metropolis-Hastings setup. The $\gamma_{j,s}$ denote the initial parameter values, chosen such as to obtain stable moment dynamics. The inferred MAP estimates are denoted as $\gamma_{j,MAP}$. The scaling parameters of the log-normal proposal distributions used in the Metropolis-Hastings algorithm are denoted as v_j.

Parameter	$\gamma_{j,s}$	$\gamma_{j,MAP}$	Unit	v_j
u_0	$3.000 \cdot 10^{-2}$	$1.581 \cdot 10^{-2}$	$a.U.$	0.01
n_H	3.000	6.130	1	0.01
K_d	$2 \cdot 10^{-1}$	$1.418 \cdot 10^{-1}$	$a.U.$	0.01
V_{max}	1.000	1.025	s^{-1}	0.01
c_2	1.000	1.384	s^{-1}	0.02
c_3	$4.000 \cdot 10^{-4}$	$6.669 \cdot 10^{-4}$	s^{-1}	0.02
c_4	$1.000 \cdot 10^{-3}$	$1.469 \cdot 10^{-2}$	s^{-1}	0.02
c_5	1.000	$2.825 \cdot 10^{-1}$	s^{-1}	0.02
c_7	$1.000 \cdot 10^{-3}$	$5.476 \cdot 10^{-4}$	s^{-1}	0.02
c_8	$1.000 \cdot 10^{-4}$	$1.283 \cdot 10^{-4}$	s^{-1}	0.02
α_1^1	$3.300 \cdot 10^{1}$	$2.250 \cdot 10^{2}$	1	0.01
$c_6\alpha_2^1$	$3.300 \cdot 10^{-2}$	$5.663 \cdot 10^{-3}$	s^{-1}	0.01
α_{11}^2	$1.900 \cdot 10^{3}$	$7.809 \cdot 10^{3}$	1	0.02
$c_6^2\alpha_{22}^2$	$1.900 \cdot 10^{-3}$	$3.098 \cdot 10^{-6}$	s^{-2}	0.02
$c_6\alpha_{12}^2$	$1.100 \cdot 10^{-1}$	$8.935 \cdot 10^{-2}$	s^{-1}	0.01

histogram value at bin zero. To interpolate between the measurement time points we fitted curves for the CV and the dose-response using polynomial base functions and a Hill-function, respectively. In both cases, the parameters were found by minimizing the L_2 - norm between the parametric model output and the CV and dose-response values.

S.4.7 Slow Transcriptional Activation Causes Transient Bimodality

The authors of [9] hypothesize that the induced pSTL1-qV bimodality originates from slow stochastic events in the transcriptional activation (i.e., interaction of the active gene with the chromatin remodeling complex, mRNA Polymerase II binding, etc.). Depending on the kinetic parameters, our model can induce bimodality either in the gene activation - or the subsequent CR binding step. To computationally verify the hypothesis from [9], we performed a statistical analysis of those events using stochastic simulation with the inferred model parameters from Tab.S.3. In particular, we found that for intermediate salt concentrations, all cells have active Hog1 bound to the promoter, whereas CR binding occurs in only a fraction of the cells (see Fig. S.8). This observation provides strong computational support for the hypothesis that chromatin remodeling is the key factor in bimodal pSTL1-qV induction and furthermore that the induction is governed by intrinsic stochasticity.

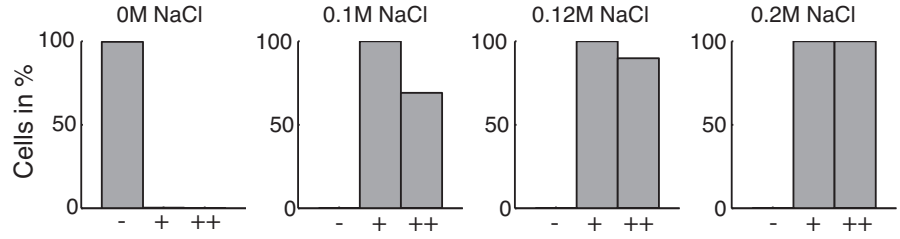

Figure S.8: Statistical analysis of the transcriptional activation in pSTL1-qV expression. Bars indicate the percentage of cells that never activated the gene (-), that activated the gene at least once (+) and cells that initiated transcription (++). Statistics were computed from 1000 traces obtained using stochastic simulation with the inferred parameters.

References

[1] R. Bellman and K. Astrom. On structural identifiability. *Mathematical Biosciences*, 7(3-4):329–339, 1970.

[2] Y. Cao and L. Petzold. Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems. *Journal of Computational Physics*, 212(1):6–24, Feb. 2006.

[3] C. Furusawa, T. Suzuki, A. Kashiwagi, T. Yomo, and K. Kaneko. Ubiquity of log-normal distributions in intra-cellular reaction dynamics. *Biophysics*, 1:25–31, 2005.

[4] J. D. Gibbons and S. Chakraborti. *Nonparametric statistical inference*. Marcel-Dekker, New York, 4 edition, 2003.

[5] A. L. Gibbs and F. E. Su. On choosing and bounding probability metrics. *International Statistical Review*, 70(3):419–435, 2002.

[6] S. Hengl, C. Kreutz, J. Timmer, and T. Maiwald. Data-based identifiability analysis of non-linear dynamical models. *Bioinformatics*, 23(19):2612–2618, 2007.

[7] J. P. Hespanha. `StochDynTools` — a MATLAB toolbox to compute moment dynamics for stochastic networks of bio-chemical reactions. Available at: `http://www.ece.ucsb.edu/~hespanha`, May 2007.

[8] F. C. Holstege, E. G. Jennings, J. J. Wyrick, T. I. Lee, C. J. Hengartner, M. R. Green, T. R. Golub, E. S. Lander, and R. a. Young. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell*, 95(5):717–28, Nov. 1998.

[9] S. Pelet, F. Rudolf, M. Nadal-Ribelles, E. de Nadal, F. Posas, and M. Peter. Transient activation of the HOG MAPK pathway regulates bimodal gene expression. *Science (New York, N.Y.)*, 332(6030):732–5, May 2011.