Network Reconstruction from Perturbation Time Course Data

Gregory R Smith[1], Mehdi Bouhaddou[1,2], Alan D. Stern[1], Caitlin M. Anglin[2], Orrod M. Zadeh[2], Jake Erskine[2], and Marc R Birtwistle[1,2,#]

[1]Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY

[2]Department of Chemical and Biomolecular Engineering, Clemson University, Clemson, SC

[#]To whom all correspondence should be addressed: mbirtwi@clemson.edu

Summary

Quantitative, directional network structure inference remains challenging even for small systems, particularly when loops and cycles are present. We report a method that robustly infers direct, signed connections between network nodes from noisy, sparse perturbation time course data requiring only one perturbation per node. We find good sensitivity and specificity for classification, as well as quantitative agreement in randomized 2- and 3-node systems having varied and complex dynamics. Experimental application of the method to the ERK and AKT pathways, widely important in mammalian signaling, reveals evidence of bi-directional cross-talk coupled with strong negative feedback on both pathways, consistent with prior knowledge. Systematic application of this method can reduce important subnetwork structural uncertainty, enabling better prediction of dynamics, response to perturbations such as drugs, and understanding of biological networks. The method is general and can be applied to any network inference problem where perturbation time course experiments are possible.

Introduction

Networks underlie much cellular and biological behavior, including transcriptional, protein-protein interaction, signaling, metabolic, cell-cell, ecological, and social networks, among many others. As such, identifying and then representing their structure has been the focus of many for decades now. This is both from experimental perspectives alone, but predominantly computational with a variety of statistical methodologies that integrate new data with prior knowledge from interaction databases and experimental data (Angulo et al., 2017; Barabási and Albert, 1999; Califano et al., 2012; Calvano et al., 2005; Hein et al., 2015; Hill et al., 2017, 2017, Ideker et al., 2001, 2002; Liu et al., 2013; Ma'ayan et al., 2005; Margolin et al., 2006; Mazloom et al., 2011; Mehla et al., 2015; Molinelli et al., 2013; Pe'er et al., 2001; Pósfai et al., 2013; Shannon et al., 2003; Stein et al., 2015; Wynn et al., 2018). Network structure can be represented as either undirected or directed graphs with edges between nodes, and the links between nodes can be weighted or not. The edges in undirected graphs are often non-causal, and even directed graphs may have edges that are purely correlational, rather than causal. A complete and predictive understanding of the biological networks that dictate cellular behavior often requires information relating to causality, which are not typically guaranteed by commonly employed correlational and Bayesian methods. This is particularly the case for networks with cycles such as feedback or feedforward loops, which are nearly ubiquitous in biological systems. Moreover, dynamic and quantitative behavior of a network is important for prescribing biological function (e.g. circadian or p53 oscillators (Bell-Pedersen et al., 2005; Stewart-Ornstein et al., 2017)). To describe and predict such behavior, networks must often also be weighted. Identifying such networks largely requires perturbation time course data to gain information related to causality,

but how to uniquely identify such networks and what data are needed to do so is not as well explored.

Modular Response Analysis (MRA) approaches, first pioneered by Kholodenko and colleagues in 2002 (Kholodenko et al., 2002; Santra et al., 2018) inherently deal with cycles and causality by prescribing systematic perturbation experiments. The original instantiations struggled with noise, but total least squares MRA and monte carlo sampling helped to improve performance (Andrec et al., 2005; Santos et al., 2007). Incomplete and prior knowledge can be handled as well using both maximum likelihood and Bayesian approaches (Halasz et al., 2016; Klinger et al., 2013; Santra et al., 2013). However, most of these approaches are based on steady-state data, or fixed time point data, limiting abilities to deal with dynamic systems and infer causality. One MRA variant, dynamic Modular Response Analysis (DMRA) (Sontag et al., 2004), uses time-series perturbation data to uniquely infer a weighted, directed network that in principle can deal with cycles, and can also predict time-dependent network behavior. However, two separate perturbation experiments are reported to be required for each node, which can be experimentally challenging both in terms of scale and finding suitable distinct perturbations for a node (e.g. both production and degradation of a transcript, or phosphorylation and dephosphorylation of a protein). Moreover, as is often the case, noise in the experimental data can severely limit the inference accuracy. This drawback has largely precluded any widespread application of DMRA. Thus, there remains a need for inference methods that can reconstruct weighted, directed networks from realistic perturbation time course experiments and that function in the presence of typical noise levels.

Here we describe a novel network inference approach that builds on MRA and DMRA, and that addresses these challenges of experimental noise, sparse time courses, and correctly

inferring cycles. We test our approach on a series of simulated time-series perturbation data with known network topology under increasing levels of simulated noise. The approach has good accuracy and precision for identifying network structure in random two and three node networks that contain a wide variety of cycles. We also demonstrate the approach works well for a family of 16 non-linear feedforward loop network models. We then apply our approach to experimental data to infer pathway interaction strengths between ERK and AKT signaling from ELISA-based perturbation time-course experiments, revealing both previously known and new network connections. While challenges remain for expanding to larger systems with more parameters and dynamic network connection strengths, the proposed algorithm accurately reconstructs networks with feasible time course experiments.

Results

Formulation

Consider a two-node network with four directed, weighted edges (Fig. 1a), where a stimulus induces measurable dynamic changes in the nodes (Fig. 1b). Let $x_i(k)$ be the activity of node i at time point t_k. The network dynamics can cast as a system of ordinary differential equations (ODEs)

$$\frac{dx_1}{dt} \equiv f_1(x_1(k), x_2(k)) \equiv f_1(k); \frac{dx_2}{dt} \equiv f_2(x_1(k), x_2(k)) \equiv f_2(k) \quad (1)$$

The network edge weights can be connected to the system dynamics through the so-called Jacobian matrix (Santra et al., 2018; Sontag et al., 2004),

$$\mathbf{J} \equiv \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \equiv \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} \quad (2)$$

Thus, the network edge weights (F_{ij}'s) describe how the activity of one node affects the dynamics of another node in a causal and direct sense (though not necessarily in a physical sense).

Consider three time course experiments that each measure x_1 and x_2 dynamics in response to a common stimulus (Fig. 1c,d), where one is in the presence of no perturbation, one has a perturbation of x_1, and one has a perturbation of x_2. Consider further that the perturbations are reasonably specific, such that the perturbation of x_1 has small direct effects on x_2, and vice versa. Experimentally, this would assume that a shRNA or gRNA is specific to a particular considered node, or that a kinase inhibitor is used at low enough dose to predominantly inhibit the targeted node. A well-posed estimation problem can be formulated (see Methods) that, in principal,

allows for unique estimation of the Jacobian elements as a function of time from such data with the following set of linear algebra relations:

$$\begin{bmatrix} y_1(t_{k+1}) \\ y_1^{(2)}(t_k) \end{bmatrix} = \begin{bmatrix} \Delta_t x_1(t_{k+1}) & \Delta_t x_2(t_{k+1}) \\ \Delta_p^{(2)} x_1(t_k) & \Delta_p^{(2)} x_2(t_k) \end{bmatrix} \begin{bmatrix} F_{11}(t_k) \\ F_{12}(t_k) \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} y_2(t_{k+1}) \\ y_2^{(1)}(t_k) \end{bmatrix} = \begin{bmatrix} \Delta_t x_1(t_{k+1}) & \Delta_t x_2(t_{k+1}) \\ \Delta_p^{(1)} x_1(t_k) & \Delta_p^{(1)} x_2(t_k) \end{bmatrix} \begin{bmatrix} F_{21}(t_k) \\ F_{22}(t_k) \end{bmatrix} \quad (4)$$

Here, y_i refers to a measured first time derivative of node *i*, superscripts refer to a particular perturbation, and Δ to a difference with respect to perturbation (subscript *p*) or time (subscript *t*) (see Methods for details). Importantly, this formulation is generalizable to any *n* dimensional network. With n^2 unknown parameters in the Jacobian matrix, *n* equations originate from the vehicle perturbation and *n-1* equations originate from each of the *n* perturbations (following the above requirement, we discard equations from node *i* with perturbation *i*). This results in $n + n*(n-1) = n + n^2 - n = n^2$ independent equations.

Directly Solving for Edge Weights

As an initial test of the above formulation, we used a simple two-node, single activator network where Node x_1 activates Node x_2, and both nodes have first-order degradation (-1 diagonal elements), all with time-independent Jacobian elements (Fig. 2a—see Methods for equations). Following a stimulus that increases the activity of each node, we simulated an 11-point time course experiment (chosen for experimental feasibility—despite being on the high side of time points typically gathered). This procedure was done for the vehicle control and for each perturbation case (Node x_1 and Node x_2) to generate the necessary simulated data per the theoretical considerations above (Fig. 2b). Here, we modeled perturbations as complete

inhibition (although incomplete perturbations are explored later). After solving Eqs. 3-4 to infer the Jacobian elements at each time point, we saw excellent agreement with the median estimates and the ground truth values (Fig. 2c, "Analytic Solution").

How does this approach fare when data are noisy? We added a relatively small amount of simulated noise (10:1 signal-to-noise, typical for high precision experimental data) to the time course data and reapplied the inference algorithm. The resulting parameter predictions over the simulated time course became considerably less accurate, varying on a scale more than ten times greater than each parameter's magnitude with predictions both positive and negative regardless of the ground truth value (Fig. 2d).

Least Squares Estimation Leveraging Data Across Time Points

Although the aforementioned theoretical approach suggests the sufficiency of the perturbation time course datasets to uniquely estimate the edge weights, in practice even small measurement noise corrupts such estimates. Therefore, we considered an alternative representation by employing a least squares estimation approach to predict edge weights (Andrec et al., 2005), rather than solve the linear equations directly. For a given set of edge weight parameters, one can integrate Eq. 1 to obtain a solution to the dynamic behavior of the resulting model, which can be directly compared to data. Thus, our problem reduces to finding a set of time-invariant parameters F_{ij} that best fit simulation to data (see Methods). We applied this new approach to the single activator model, 10:1 signal-noise ratio case above where the analytic approach failed. This new estimation approach was able to infer the network structure accurately and precisely (Figure 2e).

Random Two and Three Node Networks

To investigate the robustness of the least squares estimation approach, we began testing it on increasingly complex networks with larger amounts of measurement noise. We generated 50 randomized two-node models where each edge weight is randomly sampled from a uniform distribution over the interval [-2,2] (Fig. 3a and Fig. S1, left). Each random network was checked for stability (see Methods), many (21/50) displayed potential for oscillatory behavior (non-zero imaginary parts of eigenvalues), and multiple networks with positive feedback were present (16/50). For each random two-node model, we generated a simulated dataset based on the prescribed time series perturbation experimental design, using complete inhibition as the perturbation. The stimulus to each node was also randomly sampled from a uniform distribution, but over the interval [0,2], which we also attempted to estimate. We added 10:1 signal-to-noise, 5:1 signal-to-noise, and 2:1 signal-to-noise (see Fig. 3b as an example) to the data.

For each randomized model and noise level, we applied least squared estimation and found reasonable agreement between inferred and ground truth values, even at the higher noise levels (Fig. 3c). The algorithm was also able to estimate the contributions of stimulus to each node; we refer to these quantities as F_{i0}, where i is the node the stimulus is acting on (see Methods). The no-noise case shows an almost exact match between prediction and ground truth for the edge weights. While the predictability of edge weights decreases as noise increases (especially as the magnitude of the ground truth value increases), in general, larger ground truth values translate to larger predicted values and vice versa. Importantly, we find incorrect classification for an edge (i.e. predicts an edge as positive/activator, when it is negative/repressor, or vice versa) is extremely rare (Fig. 3d).

We next expanded to random three node models (Fig. 4a). We similarly generated 50 randomized three-node models and the same noise levels were applied. As above, each random network was checked for stability (see Methods) and many (23/50) displayed potential oscillatory behavior (non-zero imaginary parts of eigenvalues), with multiple types of motifs present (Fig. S1, right). We generated simulated noisy perturbation time course data using complete inhibition. After applying the least squares estimation approach to the simulated noisy data from each random model, we again find high consistency between the predicted and ground truth parameter values (Fig. 4b). Directionality prediction accuracy is still strong (Fig. 4c), although performance is noticeably worse with the extreme case of 2:1 signal-to-noise.

Non-Linear Feed Forward Loop Networks with Time-Varying Jacobians

The previous applications were idealized cases, where the parameters F_{ij} were time-invariant and the model used to generate simulated data was the same as the one use to fit the data. Here, we challenged the proposed algorithm further with a series of non-linear feed forward loop (FFL) network models that have time-varying Jacobian elements (Fig. 5a, Table 1). Such FFL motifs are strongly enriched in multiple organisms and are important for signaling functions such as integrative control, persistence detection, and fold-change responsiveness (Goentoro and Kirschner, 2009; Goentoro et al., 2009; Nakakuki et al., 2010).

The FFL network has three nodes (x_1, x_2, and x_3). In the original set of models, the stimulus to this network acts only on x_1 (F_{10}). However, we wanted to challenge the algorithm by potentially allowing the stimulus to act on nodes independently, as may be the case in real situations. Each node exhibits first-order decay (F_{ii}=-1). The parameters F_{12}, F_{13}, and F_{23} represent connections that do not exist in the model; we call these null edges, but we allow them to be estimated. The relationship between x_1 and x_2 (F_{21}), between x_1 and x_3 (F_{31}), or between x_2

and x_3 (F_{32}) can be either activating or inhibitory. Furthermore, x_1 and x_2 can regulate x_3 through an "AND" gate (both needed) or an "OR" gate (Fig. 5a). These permutations give rise to 16 different FFL structures (Table 1).

For each of these structures, we integrated the system of ODEs to generate time series perturbation data consistent with the proposed reconstruction algorithm, again with varying levels of signal-to-noise (No noise, 10, 5, 2) and using complete perturbations, and then applied the estimation algorithm. We did not consider any prior knowledge in the network; we allowed the algorithm to choose values for all nine edge weights plus stimuli strengths F_{i0}. To assess the performance of the algorithm and account for uncertainty in estimated edge weights, we generated 50 bootstrapped datasets for each network structure/signal-to-noise pair, and thus obtained 50 sets of estimates, similar to previous work (Santos et al., 2007). Edges are classified as activators (>0), repressors (<0), or null (~0).

We first noticed that even in the absence of added noise, some inferences were incorrect (Fig. 5b). Model #4 (Table 1) is used as an example, where F_{21} is an activating and F_{31} and F_{32} are repressors with an AND gate; F_{31} is incorrectly predicted as activating (Fig 5b—compare ground truth to complete inhibition). We reasoned this is due to the following. Since x_1 is required for the activation of x_2, and both x_1 and x_2 are required for the activation of x_3, complete inhibition of x_1 removes all regulated activity from x_2 and x_3. Completely inhibiting x_2 activity also removes all activity from x_3. Thus, given this experimental setup, it becomes difficult to discern if x_1 directly influences x_3 or if it acts solely through x_2. If instead of applying complete inhibition, we use partial (50%) inhibition, F_{31} is correctly identified as a repressor in this case, and more broadly, we obtain perfect classification from noise-free data across all 16 FFL networks (top row, Fig. 5e). Thus, we instead applied 50% perturbation to all simulation data and

proceeded with least squares estimation. Example time courses and fits for the no noise case are shown in Fig. S3 and S4.

Next, we sought a rationale for the classification of null, activating and repressing edges from the least squares estimation results. From the bootstrapped datasets, there are 50 estimates for each edge weight. We chose a simple approach of a strict, symmetric cutoff around zero based on the percentile of these 50 estimates. For example, if the percentile cutoff was 0.9, and if 90% of an edge's weight estimates were greater than zero, then the edge would be classified as an activator, and if 90% were below zero, the edge would be classified as a repressor. Anything between is classified as a null link, meaning that there is not sufficient evidence to support inference of a connection. We varied the percentile cutoff from the median only (0.5) to the entire range of estimated values (1) and calculated the true and false positive rates for all edges across all 16 FFL models (Fig. 5c). For each noise level, we chose the percentile cutoff that yielded a 5% false positive rate (0.69 for 1:10 noise, 0.66 for 1:5 noise, and 0.7 for 1:2 noise) to generate receiver operating characteristic (ROC) curves (Fig. 5d). We observed excellent classification performance across all noise levels (10:1 AUC=0.996, 5:1 AUC=0.992, 2:1 AUC=0.959).

To evaluate the performance for each of the 16 FFL cases, we calculated the fraction of the 12 links in each FFL model that was classified correctly as a function of signal-to-noise, given the cutoffs determined above (Fig. 5e and S2). Perfect classification is a value of one, which is the case for no noise, and for most cases with 10:1 signal-to-noise. In general, as noise level increases, prediction accuracy decreases, as expected. This is particularly the case for the edges that are distinct across models (F_{21}, F_{31}, F_{32}), which are difficult to infer with 2:1 signal-to-noise (Fig. S2). Although for some models, performance at 2:1 signal-to-noise is quite poor, in

some cases it is surprisingly good and even perfect. This suggests that the proposed method is very accurate in low-to-medium noise cases and, even when the data are essentially unacceptably noisy (2:1) and thus more difficult to analyze, the non-zero predictions that are made have a surprisingly low likelihood of error.

Application to Experimental Data

We wanted to test the approach with experimental data. We decided to investigate the interaction strengths (i.e. crosstalk) between the ERK and AKT pathways, two broadly important mammalian signaling pathways regulating cell proliferation, death, migration and differentiation (Manning and Toker, 2017; Shaul and Seger, 2007). We used In-Cell ELISA to measure the dynamic response of ERK and AKT phosphorylation in serum- and growth-factor starved MCF10A cells to an Epidermal Growth Factor (EGF) and Insulin combination stimulus, which robustly activates both ERK and AKT over ~hour time scales to drive the cell cycle (Bouhaddou et al., 2018). We did this in the presence of vehicle (Fig. 6a, top), MEK inhibition (Fig. 6a, middle, directly upstream of ERK), or AKT inhibition (Fig. 6a, bottom).

After applying least squares estimation, we obtained the Jacobian which characterizes the network (Fig. 6b). There are two aspects that are corroborated by prior knowledge. First, F_{10} and F_{20} are both positive; it is well known that EGF+Insulin can activate ERK and AKT with independent mechanisms of Ras and PI-3K activation (Manning and Toker, 2017; Shaul and Seger, 2007). Second, negative feedback has been widely reported for both pathways through post-translational and transcriptional mechanisms (Efeyan and Sabatini, 2010; Fritsche-Guenther et al., 2011; Sturm et al., 2010). F_{11} and F_{22} are both negative and less than -1, which implies negative auto-regulation in excess of simple first-order decay. Cross-talk between these pathways is context dependent, with reports of none, positive and negative links (Dent, 2014;

Mendoza et al., 2011; Wang et al., 2009). In this context of serum-starved MCF10A cells stimulated with EGF+Insulin, our results suggest AKT positively regulates ERK, whereas ERK negatively regulates AKT, and both have similar magnitudes. These results can form the basis for future mechanistic investigations into the biochemistry and physical reactions comprising these putative links, and also confirm that our approach can be applied to real experimental time course data to uncover both known and imply new biological network structure.

Discussion

Despite intensive research focus on network reconstruction, there is still room to improve discrimination between direct and indirect edges (i.e. causality), particularly when biologically-ubiquitous feedback and feedforward cycles are present that stymie many statistical or correlation-based methods, and also given that experimental noise is inevitable. The presented method prescribes a realistic experimental design for causal edge weight inference with experimental noise. For small 2 and 3 node networks, the method can successfully handle both random networks with constant edge weights over a time course as well as time-variant edge weights in more complex, non-linear feed forward loop structures. Prediction accuracy was strong in many cases even with simulated noise that exceeds typical experimental variability (2:1 signal-to-noise). The method presented here is applied to molecular biology examples, but they are in fact quite general, and could be applied to vastly different fields where perturbation time course experiments are possible, and where network structures are important to determine.

Although complete inhibition is often used for practical perturbations (e.g. CRISPR), we found that partial inhibition is important to fully deconvolve the explored FFL networks. This is not likely to be the only case when partial inhibition will be important, so we are inclined to speculate it will be a generally important experimental design criterion moving forward. However, there is likely to be a sharp tradeoff between perturbation strength and feasibility, since the effects of small perturbations may be masked by noise. Partial inhibition is often "built-in" to certain assay types, such as si/shRNA or pharmacological inhibition that are titratable to certain magnitudes.

One major remaining challenge is scaling to larger networks. Here, we limited our analysis to 2 and 3 node networks. Conveniently, the number of necessary perturbation time courses needed grows linearly with the number of considered nodes. Furthermore, as long as system-wide or omic-scale assays are available, so to does the experimental workload. This is routine for transcriptome analyses, and is becoming even more commonplace for proteomic assays (e.g. mass cytometry, cyclic immunofluorescence, mass spectrometry, RPPA) (Lin et al., 2016).

Increasing the network size will quadratically increase the number of unknown parameters. Reducing this parameter space and obtaining good initial guesses (such as with MRA as we did here—see Methods) will be important. Imposing prior knowledge can also reduce the parametric space, such as in Bayesian Modular Response Analysis (Santra et al., 2013). As network size grows, sparseness of the Jacobian will increase, so judicious allocation of non-zero elements will be important. Checking estimated Jacobians for emergent properties such as degree distributions for scale-free networks (Barabási and Albert, 1999) can provide additional important constraints. Lastly, large estimation problems may be broken into several smaller problems to be merged subsequently, which is likely to yield large computational speed up by allowing parallelization of much smaller tasks.

Acknowledgements

We would like to thank Achla Gupta and Gomathi Jayaraman for help with ELISA experiments.

Author Contributions

MRB and GRS conceived of the work. GRS, MB, MRB, CMA, OMZ and JE performed analyses. MB and GRS made the figures. GRS, MB, and MRB wrote the manuscript. MB and ADS performed the ELISA experiments.

Declaration of Interests

The authors declare no competing interests.

Funding

MRB acknowledges funding from Mount Sinai, Clemson University, the National Institutes of Health Grants R01GM104184 and R21CA196418, and an IBM faculty award. MB and ADS were supported by a National Institute for General Medical Sciences-funded Integrated Pharmacological Sciences Training Program grant (T32GM062754).

References

Andrec, M., Kholodenko, B.N., Levy, R.M., and Sontag, E. (2005). Inference of signaling and gene regulatory networks by steady-state perturbation experiments: Structure and accuracy. J. Theor. Biol. *232*, 427–441.

Angulo, M.T., Moreno, J.A., Lippner, G., Barabási, A.-L., and Liu, Y.-Y. (2017). Fundamental limitations of network reconstruction from temporal data. J. R. Soc. Interface *14*.

Barabási, A.L., and Albert, R. (1999). Emergence of scaling in random networks. Science (80-.). *286*, 509–512.

Bell-Pedersen, D., Cassone, V.M., Earnest, D.J., Golden, S.S., Hardin, P.E., Thomas, T.L., and Zoran, M.J. (2005). Circadian rhythms from multiple oscillators: lessons from diverse organisms.

Nat. Rev. Genet. *6*, 544–556.

Bouhaddou, M., Barrette, A.M., Stern, A.D., Koch, R.J., DiStefano, M.S., Riesel, E.A., Santos, L.C., Tan, A.L., Mertz, A.E., and Birtwistle, M.R. (2018). A mechanistic pan-cancer pathway model informed by multi-omics data interprets stochastic cell fate responses to drugs and mitogens. PLoS Comput. Biol. *14*.

Califano, A., Butte, A.J., Friend, S., Ideker, T., and Schadt, E. (2012). Leveraging models of cell regulation and GWAS data in integrative network-based association studies. Nat. Genet. *44*, 841–847.

Calvano, S.E., Xiao, W., Richards, D.R., Felciano, R.M., Baker, H. V, Cho, R.J., Chen, R.O., Brownstein, B.H., Cobb, J.P., Tschoeke, S.K., et al. (2005). A network-based analysis of systemic inflammation in humans. Nature *437*, 1032–1037.

Dent, P. (2014). Crosstalk between ERK, AKT, and cell survival. Cancer Biol. Ther. *15*, 245–246.

Efeyan, A., and Sabatini, D.M. (2010). mTOR and cancer: many loops in one pathway. Curr. Opin. Cell Biol. *22*, 169–176.

Fritsche-Guenther, R., Witzel, F., Sieber, A., Herr, R., Schmidt, N., Braun, S., Brummer, T., Sers, C., and Blüthgen, N. (2011). Strong negative feedback from Erk to Raf confers robustness to MAPK signalling. Mol. Syst. Biol. *7*, 489.

Goentoro, L., and Kirschner, M.W. (2009). Evidence that Fold-Change, and Not Absolute Level, of β-Catenin Dictates Wnt Signaling. Mol. Cell *36*, 872–884.

Goentoro, L., Shoval, O., Kirschner, M.W., and Alon, U. (2009). The Incoherent Feedforward

Loop Can Provide Fold-Change Detection in Gene Regulation. Mol. Cell *36*, 894–899.

Halasz, M., Kholodenko, B.N., Kolch, W., and Santra, T. (2016). Integrating network reconstruction with mechanistic modeling to predict cancer therapies. Sci. Signal. *9*.

Hein, M.Y., Hubner, N.C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I.A., Weisswange, I., Mansfeld, J., Buchholz, F., et al. (2015). A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell *163*, 712–723.

Hill, S.M., Nesser, N.K., Johnson-Camacho, K., Jeffress, M., Johnson, A., Boniface, C., Spencer, S.E.F., Lu, Y., Heiser, L.M., Lawrence, Y., et al. (2017). Context Specificity in Causal Signaling Networks Revealed by Phosphoprotein Profiling. Cell Syst. *4*, 73–83.e10.

Ideker, T., Thorsson, V., Ranish, J.A., Christmas, R., Buhler, J., Eng, J.K., Bumgarner, R., Goodlett, D.R., Aebersold, R., and Hood, L. (2001). Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science (80-.). *292*, 929–934.

Ideker, T., Ozier, O., Schwikowski, B., and Siegel, A.F. (2002). Discovering regulatory and signalling circuits in molecular interaction networks. Bioinformatics *18 Suppl 1*, S233-40.

Kholodenko, B.N., Kiyatkin, A., Bruggeman, F.J., Sontag, E., Westerhoff, H. V, and Hoek, J.B. (2002). Untangling the wires: a strategy to trace functional interactions in signaling and gene networks. Proc. Natl. Acad. Sci. U. S. A. *99*, 12841–12846.

Klinger, B., Sieber, A., Fritsche-Guenther, R., Witzel, F., Berry, L., Schumacher, D., Yan, Y., Durek, P., Merchant, M., Sch??fer, R., et al. (2013). Network quantification of EGFR signaling unveils potential for targeted combination therapy. Mol. Syst. Biol. *9*.

Lin, J.-R., Fallahi-Sichani, M., Chen, J.-Y., and Sorger, P.K. (2016). Cyclic Immunofluorescence

(CycIF), A Highly Multiplexed Method for Single-cell Imaging. Curr. Protoc. Chem. Biol. *8*, 251–264.

Liu, Y.-Y., Slotine, J.-J., and Barabási, A.-L. (2013). Observability of complex systems. Proc. Natl. Acad. Sci. U. S. A. *110*, 2460–2465.

Ma'ayan, A., Jenkins, S.L., Neves, S., Hasseldine, A., Grace, E., Dubin-Thaler, B., Eungdamrong, N.J., Weng, G., Ram, P.T., Rice, J.J., et al. (2005). Formation of regulatory patterns during signal propagation in a Mammalian cellular network. Science *309*, 1078–1083.

Mangan, S., and Alon, U. (2003). Structure and function of the feed-forward loop network motif. Proc. Natl. Acad. Sci. U. S. A. *100*, 11980–11985.

Manning, B.D., and Toker, A. (2017). AKT/PKB Signaling: Navigating the Network. Cell *169*, 381–405.

Margolin, A.A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Dalla Favera, R., and Califano, A. (2006). ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics *7 Suppl 1*, S7.

Mazloom, A.R., Dannenfelser, R., Clark, N.R., Grigoryan, A. V, Linder, K.M., Cardozo, T.J., Bond, J.C., Boran, A.D.W., Iyengar, R., Malovannaya, A., et al. (2011). Recovering protein-protein and domain-domain interactions from aggregation of IP-MS proteomics of coregulator complexes. PLoS Comput. Biol. *7*, e1002319.

Mehla, J., Caufield, J.H., and Uetz, P. (2015). The yeast two-hybrid system: a tool for mapping protein-protein interactions. Cold Spring Harb. Protoc. *2015*, 425–430.

Mendoza, M.C., Er, E.E., and Blenis, J. (2011). The Ras-ERK and PI3K-mTOR pathways: cross-

talk and compensation. Trends Biochem. Sci. *36*, 320–328.

Molinelli, E.J., Korkut, A., Wang, W., Miller, M.L., Gauthier, N.P., Jing, X., Kaushik, P., He, Q., Mills, G., Solit, D.B., et al. (2013). Perturbation biology: inferring signaling networks in cellular systems. PLoS Comput. Biol. *9*, e1003290.

Nakakuki, T., Birtwistle, M.R., Saeki, Y., Yumoto, N., Ide, K., Nagashima, T., Brusch, L., Ogunnaike, B.A., Okada-Hatakeyama, M., and Kholodenko, B.N. (2010). Ligand-specific c-fos expression emerges from the spatiotemporal control of ErbB network dynamics. Cell *141*.

Pe'er, D., Regev, A., Elidan, G., and Friedman, N. (2001). Inferring subnetworks from perturbed expression profiles. Bioinformatics *17 Suppl 1*, S215-24.

Pósfai, M., Liu, Y.-Y., Slotine, J.-J., and Barabási, A.-L. (2013). Effect of correlations on network controllability. Sci. Rep. *3*, 1067.

Santos, S.D.M., Verveer, P.J., and Bastiaens, P.I.H. (2007). Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. Nat. Cell Biol. *9*, 324–330.

Santra, T., Kolch, W., and Kholodenko, B.N. (2013). Integrating Bayesian variable selection with Modular Response Analysis to infer biochemical network topology. BMC Syst. Biol. *7*.

Santra, T., Rukhlenko, O., Zhernovkov, V., and Kholodenko, B.N. (2018). Reconstructing static and dynamic models of signaling pathways using Modular Response Analysis. Curr. Opin. Syst. Biol. *9*, 11–21.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

Shaul, Y.D., and Seger, R. (2007). The MEK/ERK cascade: from signaling specificity to diverse functions. Biochim. Biophys. Acta *1773*, 1213–1226.

Sontag, E., Kiyatkin, A., and Kholodenko, B.N. (2004). Inferring dynamic architecture of cellular networks using time series of gene expression, protein and metabolite data. Bioinformatics *20*, 1877–1886.

Stein, R.R., Marks, D.S., and Sander, C. (2015). Inferring Pairwise Interactions from Biological Data Using Maximum-Entropy Probability Models. PLoS Comput. Biol. *11*, e1004182.

Stewart-Ornstein, J., Cheng, H.W.J., and Lahav, G. (2017). Conservation and Divergence of p53 Oscillation Dynamics across Species. Cell Syst. *5*, 410–417.e4.

Sturm, O.E., Orton, R., Grindlay, J., Birtwistle, M., Vyshemirsky, V., Gilbert, D., Calder, M., Pitt, A., Kholodenko, B., and Kolch, W. (2010). The mammalian MAPK/ERK pathway exhibits properties of a negative feedback amplifier. Sci. Signal. *3*.

Wang, C.-C., Cirit, M., and Haugh, J.M. (2009). PI3K-dependent cross-talk interactions converge with Ras as quantifiable inputs integrated by Erk. Mol. Syst. Biol. *5*, 246.

Wynn, M.L., Egbert, M., Consul, N., Chang, J., Wu, Z.-F., Meravjer, S.D., and Schnell, S. (2018). Inferring Intracellular Signal Transduction Circuitry from Molecular Perturbation Experiments. Bull. Math. Biol. *80*, 1310–1344.

Figure Legends

Figure 1. **(a)** Two-node network with Jacobian elements labeled. Green arrows are stimulus. **(b)** Typical noisy time series experiment following t=0 stimulus. **(c)** Perturbation time course experimental design. Each experiment consists of a perturbation (vehicle, inhibit x_1, inhibit x_2) followed by a stimulus and time series measurements of each node. **(d)** Example perturbation time courses and how Taylor series reference (R) and evaluation (E) points are defined (see Methods) with respect to time (t) and perturbation (p), where $x_1^{(0)}$ is no perturbation (vehicle) and $x_1^{(2)}$ is x_2 perturbation. This sample result would imply activation of x_1 by x_2. Analogous data would be collected for node x_2 as well.

Figure 2. **(a)** Single activator model schematic. **(b)** Simulated time course perturbation data from the network in (a). Top row: no added noise; bottom row: 10:1 signal-to-noise added. Columns correspond to different perturbation cases. **(c-e)** Actual versus inferred Jacobian elements (edge weights) for direct solution of Eq. 3-4 in the absence (c) or presence (d) of noise or with 10:1 signal-to-noise and the least squares approach. In c-d, error bars are standard deviation across time points.

Figure 3. **(a)** Random two node network. **(b)** Example signal-to-noise effects on time courses. **(c)** Ground truth versus inferred edge weights across all 50 random networks and noise levels. **(d)** Accuracy of edge weight directionality prediction (activator is >1, repressor is <1) by noise level.

Figure 4. **(a)** Random three node network. **(b)** Ground truth versus inferred edge weights across all 50 random networks and noise levels. **(c)** Accuracy of edge weight directionality prediction (activator is >1, repressor is <1) by noise level.

Figure 5. **(a)** Feedforward loop (FFL) network models. F_{11}, F_{22}, and F_{33} values are fixed at -1; F_{12}, F_{13}, and F_{23} values are fixed at 0. F_{21}, F_{31}, and F_{32} values can be positive or negative depending on the model. The combined effect of x_1 and x_2 on x_3 is described by either an AND gate or an OR gate. There are 16 possible model structures (Table 1). Stimulus is to all nodes as before (green arrows). Parameters in each model are varied such that the nodes are responsive to the stimulus and perturbations (see supplementary code for values). **(b)** Partial inhibition increases inference accuracy. In model structure #4, F_{31} is negative (ground truth) but inference is positive using complete inhibition. Partial inhibition (50%) allows correct inference. **(c)** True (TPR; blue lines) and false positive rates (FPR; red lines) of directionality inference percentile cutoff is changed. FPR of 0.05 is the light blue horizontal line. S:N is signal-to-noise. **(d)** ROC curves for directionality inference across all 16 FFL models given percentile cutoffs chosen from (c). **(e)** Fraction of correctly inferred edge weight directionalities across the 16 FFL models and noise strengths.

Figure 6. **(a)** Quantified ELISA data from perturbation time course experiments (dots, three biological replicate per condition) with model-generated fits (lines; except ERK data under MEK inhibition and AKT data under AKT inhibition which are experimental data directly used in estimations). Phospho-ERK and phospho-AKT was measured in response to EGF+Insulin stimulations in MCF10A cells with no perturbation (top), MEK inhibition (0.1nM, middle), and

AKT inhibition (10nM, bottom). Activities were scaled normalized to serum-starved controls not treated with EGF+Insulin. **(b)** Edge weight estimates for this two-node network. AKT is predicted to positively regulate ERK, whereas ERK is predicted to negatively regulate AKT. Both pathways exhibit evidence of strong negative feedback.

Methods

Deriving Sufficiency Conditions for Unique Estimation of Jacobian Elements

The first-order partial derivatives comprising **J** (Eq. 2) can be defined by a first-order Taylor series expansion of Eq. 1 about a time point *k*

$$f_1(k+1) \approx f_1(k) + \frac{\partial}{\partial x_1}\left(f_1(k)\right)\cdot\left(x_1(k+1)-x_1(k)\right) + \frac{\partial}{\partial x_2}\left(f_1(k)\right)\cdot\left(x_2(k+1)-x_2(k)\right) \qquad (5)$$

$$f_2(k+1) \approx f_2(k) + \frac{\partial}{\partial x_1}\left(f_2(k)\right)\cdot\left(x_1(k+1)-x_1(k)\right) + \frac{\partial}{\partial x_2}\left(f_2(k)\right)\cdot\left(x_2(k+1)-x_2(k)\right) \qquad (6)$$

Eq. 3 may be written more succinctly as

$$\begin{aligned} y_1(k+1) &\approx F_{11}(k)\cdot\Delta_t x_1(k+1) + F_{12}(k)\cdot\Delta_t x_2(k+1) \\ y_2(k+1) &\approx F_{21}(k)\cdot\Delta_t x_1(k+1) + F_{22}(k)\cdot\Delta_t x_2(k+1) \end{aligned} \qquad (7)$$

where

$$y_i(k+1) \equiv f_i(k+1) - f_i(k); \Delta_t x_i(k+1) \equiv x_i(k+1) - x_i(k)\,. \qquad (8)$$

We refer to time point *k* as the reference point with respect to time (RP_t) and time point *k+1* as the evaluation point with respect to time (EP_t) (Fig. 1d). The approximation in Eq. 7 becomes more accurate as RP's and EP's become closer; however, in practice experimentally the perturbations must induce measurable changes. Also, it is evident that the edge weights are potentially time-dependent, although this is rarely considered when describing biological networks.

How do we estimate the edge weights in Eq. 7 and thus reconstruct the network? Time series data can inform x_i's and f_i's as a function of time, following application of a stimulus.

Given such stimulus-response data, however, for each time point there are only two equations for four unknowns, an underdetermined system for which more data are needed.

Consider now stimulus-response time course data in the presence of single perturbations (Fig. 1c). Let $p^{(i)}$ be a variable that reflects the strength and/or presence of different potential perturbations: $p^{(1)}$ represents inhibition of x_1 and $p^{(2)}$ represents inhibition of x_2. If $p^{(j)}$ is not explicitly written, its value is zero and/or it has no effect. Now, the ODEs also become a function of the perturbation variables

$$f_i^{(j)}(k) \equiv f_i(k, p^{(j)}) = f_i(x_1(k), x_2(k), p^{(j)}) \qquad (9)$$

The Taylor series expansions for cases with perturbations become

$$y_1^{(1)}(k) \approx F_{11}(k) \cdot \Delta_p^{(1)} x_1(k) + F_{12}(k) \cdot \Delta_p^{(1)} x_2(k) + \frac{\partial}{\partial p^{(1)}}\left(f_1(k)\right) \cdot p^{(1)} \qquad (10)$$

$$y_1^{(2)}(k) \approx F_{11}(k) \cdot \Delta_p^{(2)} x_1(k) + F_{12}(k) \cdot \Delta_p^{(2)} x_2(k) + \frac{\partial}{\partial p^{(2)}}\left(f_1(k)\right) \cdot p^{(2)} \qquad (11)$$

$$y_2^{(1)}(k) \approx F_{21}(k) \cdot \Delta_p^{(1)} x_1(k) + F_{22}(k) \cdot \Delta_p^{(1)} x_2(k) + \frac{\partial}{\partial p^{(1)}}\left(f_2(k)\right) \cdot p^{(1)} \qquad (12)$$

$$y_2^{(2)}(k) \approx F_{21}(k) \cdot \Delta_p^{(2)} x_1(k) + F_{22}(k) \cdot \Delta_p^{(2)} x_2(k) + \frac{\partial}{\partial p^{(2)}}\left(f_2(k)\right) \cdot p^{(2)} \qquad (13)$$

where

$$y_i^{(j)}(k) \equiv f_i^{(j)}(k) - f_i(k); \Delta_p^{(j)} x_i(k) \equiv x_i(k, p^{(j)}) - x_i(k) \qquad (14)$$

Here, we have chosen the reference point with respect to perturbations (RP_p) to be the same as RP_t, but the evaluation point with respect to perturbations (EP_p) to be defined by the perturbation

time course (Fig. 1d). Since the reference point is the same, the Jacobian elements remain identical in these equations. Thus, now we have six potential equations with which to estimate the four Jacobian elements. It is interesting to note that the Jacobian elements, or network, may be affected by the perturbation, but we do not necessarily have to know those effects mathematically, since the reference point is the same. However, we must make some determination as to how the perturbations $p^{(1)}$ and $p^{(2)}$ directly affect Node 1 and Node 2 dynamics f_1 and f_2 to account for the perturbation variable partial derivatives.

By design, the Node 1 perturbation has significant direct effects on Node 1 dynamics, and similarly for the Node 2 perturbation on Node 2 dynamics. Using equations including $\partial f_1/\partial p^{(1)}$ and $\partial f_2/\partial p^{(2)}$ require precise definition of perturbation strength and their effects on dynamics, which is difficult, as well as using perturbation data that is likely to be extremely far away from the reference point, and therefore we do not employ equations involving these terms here. On the other hand, if the Node 1 perturbation has negligible direct effect on Node 2 dynamics, that is, the only effects on Node 2 dynamics are through its effects on Node 1 (i.e. $p^{(1)}$ is not explicit in f_2), and similarly the Node 2 perturbation has negligible direct effect on Node 1 dynamics, then $\partial f_2/\partial p^{(1)}$ and $\partial f_1/\partial p^{(2)}$ can be neglected from the above equations. This mild condition is often the case experimentally. From this, the main set of linear equations presented in Eq. 3-4 are obtained.

ODE Model Equations and Simulation

The two-node single activator model is described by

$$\frac{dx_1}{dt} = f_1(x_1,x_2) = 1 - x_1$$
$$\frac{dx_2}{dt} = f_2(x_1,x_2) = 1 + 3x_1 - x_2 \qquad (15)$$

The random two node network is described by

$$\frac{dx_1}{dt} = f_1(x_1,x_2) = F_{10} + F_{11}x_1 + F_{12}x_2$$
$$\frac{dx_2}{dt} = f_2(x_1,x_2) = F_{20} + F_{21}x_1 + F_{22}x_2 \qquad (16)$$

Values for F_{10} and F_{20} are sampled from a uniform distribution over the range [0,2] and values for F_{11}, F_{12}, F_{21}, and F_{22} are sampled from a uniform distribution over the range [-2,2] using the MATLAB function `rand`.

The random three node networks use the same sampling rules with the following equations

$$\frac{dx_1}{dt} = f_1(x_1,x_2,x_3) = F_{10} + F_{11}x_1 + F_{12}x_2 + F_{13}x_3$$
$$\frac{dx_2}{dt} = f_2(x_1,x_2,x_3) = F_{20} + F_{21}x_1 + F_{22}x_2 + F_{23}x_3 \qquad (17)$$
$$\frac{dx_3}{dt} = f_3(x_1,x_2,x_3) = F_{30} + F_{31}x_1 + F_{32}x_2 + F_{33}x_3$$

The feedforward loop models (Mangan and Alon, 2003) are described by:

$$\frac{dx_1}{dt} = f_1(x_1, x_2, x_3) = 1 - x_1$$

$$\frac{dx_2}{dt} = f_2(x_1, x_2, x_3) = f(x_1, K_{x_1x_2}) - x_2 \qquad (18)$$

$$\frac{dx_3}{dt} = f_3(x_1, x_2, x_3) = G(x_1, K_{x_1x_3}, x_2, .K_{x_2x_3}) - x_3$$

When an AND gate is present

$$G(x_1, K_{x_1x_3}, x_2, K_{x_2x_3}) = f(x_1, K_{x_1x_3}) * f(x_2, K_{x_2x_3}) \qquad (19)$$

When an OR gate is present

$$G(x_1, K_{x_1x_3}, x_2, K_{x_2x_3}) = fc(x_1, K_{x_1x_3}, K_{x_2x_3}, x_2) + fc(x_2, K_{x_2x_3}, K_{x_1x_3}, x_1) \qquad (20)$$

For a given u, v ϵ {x_1, x_2, x_3} and K, K_u, K_v ϵ { $K_{x_1x_2}$, $K_{x_1x_3}$, $K_{x_2x_3}$ }:

If u activates its target, then:

$$f(u,K) = \frac{\left(\frac{u}{K}\right)^2}{1+\left(\frac{u}{K}\right)^2} \;;\; fc(u,K_u,K_v,v) = \frac{\left(\frac{u}{K_u}\right)^2}{1+\left(\frac{u}{K_u}\right)^2+\left(\frac{v}{K_v}\right)^2} \qquad (21)$$

If u represses its target, then:

$$f(u,K) = \frac{1}{1+\left(\frac{u}{K}\right)^2} \;;\; fc(u,K_u,K_v,v) = \frac{1}{1+\left(\frac{u}{K_u}\right)^2+\left(\frac{v}{K_v}\right)^2} \qquad (22)$$

Models were integrated using `ode15s` in MATLAB with default settings and zero initial conditions.

Time Point Selection

Unless otherwise noted, we sample 11 evenly spaced time points across a time course. This number is chosen to be on the high end of experimental feasibility. To determine the length of a time course, we calculate the time necessary for each element to reach within 1% of steady-state across perturbation conditions. If not all elements reach a steady state (e.g. oscillations), then a default time of T = 10 arbitrary units is used.

Modeling Perturbations

We model perturbations by scaling the differential equation for the perturbed node. For example, a complete inhibition of x_1 multiplies dx_1/dt by zero, and a 50% inhibition multiplies by 0.5. Because simulations in this study start from a zero initial condition, this perturbation model suffices, but in other scenarios, different considerations might be important for perturbation models.

Simulated Noise

Normally distributed white (zero mean) noise is added to simulated time courses point-wise with

$$y = x + N(0, d \cdot x) \qquad (23)$$

where x is the simulation data point, y is the noisy data point, and d represents the noise level (standard deviation of the normal random variable). Signal-to-noise ratio of 10:1, 5:1 and 2:1 are, respectively d = 0.1, 0.2, and 0.5. Normally distributed samples are obtained using `randn` in MATLAB.

Parameter Estimation

Consider experimental data matrices X_{p0}, X_{pk} , where $k \in \{1 \ldots m\}$ is a perturbation index, describing the activity level of each node at each time point t_i, $i \in \{1 \ldots n\}$ for each perturbation condition. For a given set of parameters $\hat{F}_{uv}$, $u \in \{1 \ldots m\}$, $v \in \{0 \ldots m\}$ in a system with m nodes, we want to generate a set of simulated data matrices $\hat{X}_{p0}, \hat{X}_{pk}$. To that end, we design a system of differential equations by declaring initial values $\frac{d\hat{x}_u(t_0)}{dt} = \hat{F}_{u0}$ and for a given time point t_i, $i \in \{1 \ldots n\}$:

$$\frac{d\hat{x}_u(t_i)}{dt} = \frac{d\hat{x}_u(t_{i-1})}{dt} + \sum_{v=1}^{m} \hat{F}_{uv} * (\hat{x}_u(t_i) - \hat{x}_u(t_{i-1})) \qquad (24)$$

We integrate this system of ODEs at each experimental time point by the MATLAB function `ode15s`. We model perturbations by replacing simulated output for the variable being perturbed with the corresponding experimental value at each time point. That is,

$$\hat{X}_{pk}(k, t_i) = X_{pk}(k, t_i) \forall k \in \{1, \ldots, m\}, i \in \{1, \ldots, n\} \qquad (25)$$

This ensures we are accurately capturing the change in activity as a result of the perturbation. We make this correction prior to performing `ode15s` on each time interval so that changes in other nodes reflect the perturbed values. This allows us to efficiently calculate simulated time courses that reflect each perturbation experiment.

We then define the objective function as the sum of squared errors y

$$y = \sum_{i=1}^{n} \left(\left(\sum_{j=1}^{m} \left(X_{p0}(i,j) - \hat{X}_{p0}(i,j) \right)^2 \right) + \sum_{k=1}^{m} \sum_{j \neq k} \left(X_{pk}(i,j) - \hat{X}_{pk}(i,j) \right)^2 \right) \qquad (26)$$

Note here that we do not use data from node k when perturbation k was used (as explained above).

The MATLAB function `fmincon` is used to minimize y by changing edge weights and stimulus terms (F_{i0}) within the range [-100,100]. To generate initial guesses for each of the edge weights, we apply steady-state modular response analysis (MRA) (Kholodenko *et al.*, 2002) using the final values of the time course. Briefly, for each node in the system x_i, we calculate the ith column of the global response matrix $\mathbf{R_P}$ by the central fractional differences in the final time point levels for each node j:

$$R_{Pij} = \frac{2(x_j^{(i)} - x_j^{(0)})}{(x_j^{(i)} + x_j^{(0)})} \qquad (27)$$

Here $x_j^{(0)}$ is the value of node j with the vehicle perturbation and $x_j^{(i)}$ is the value of node j with the perturbation of node i. The initial edge weight estimates are contained in the local response matrix $\mathbf{r} = -[dg(\mathbf{R_P}^{-1})]^{-1} \cdot \mathbf{R_P}^{-1}$. If an edge weight is predicted to be identically zero by MRA, it is held fixed in the estimation. The remaining parameters (effects of stimulus) are initially set to one.

Cell Culture

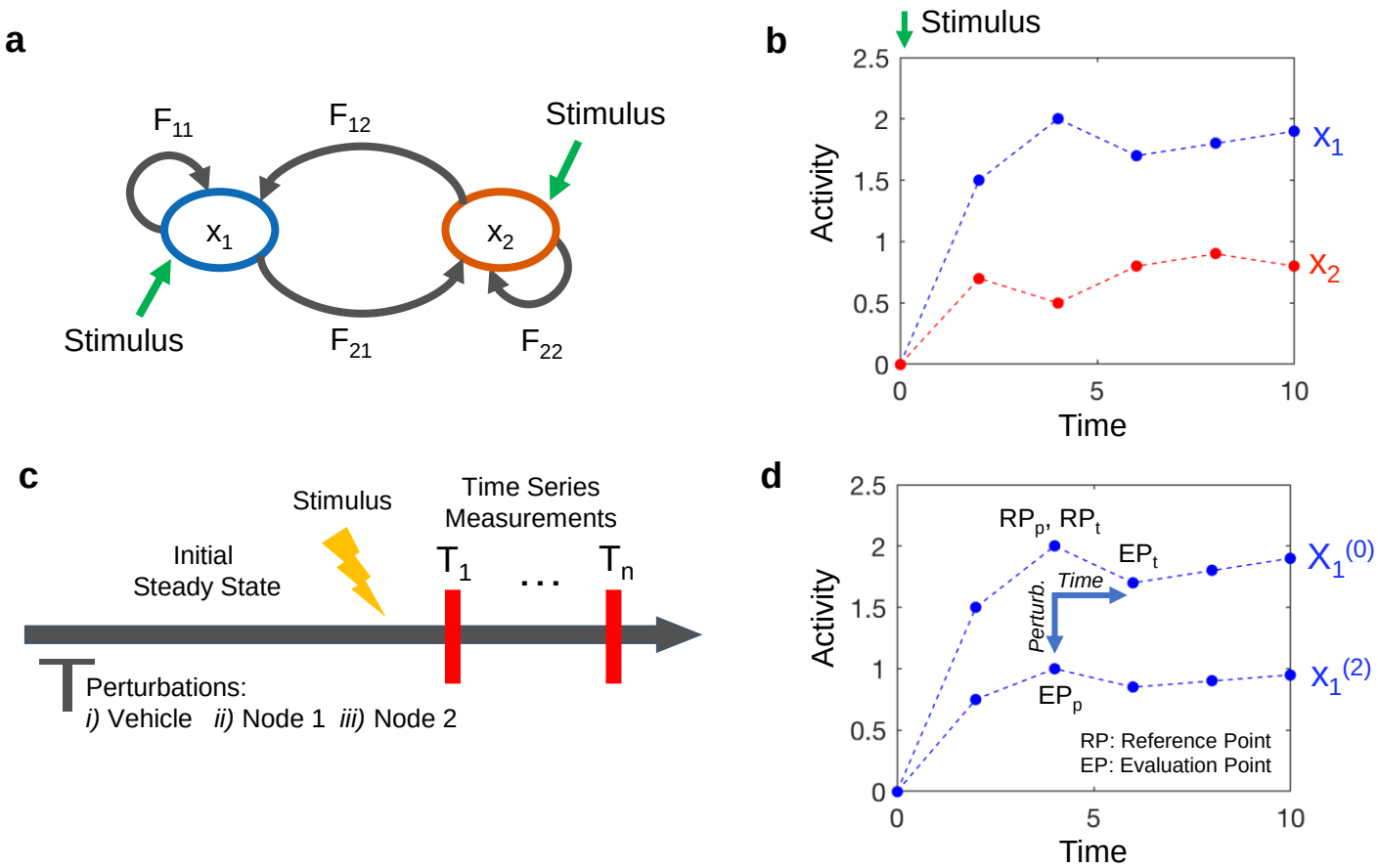
MCF10A cells were cultured in DMEM/F12 (Gibco# 11330032) medium supplemented with 5% (v/v) horse serum (Gibco# 16050–122), 20ng/mL EGF (PeproTech# AF-100-15), 0.5 mg/mL hydrocortisone (Sigma# H-0888), 10μg/mL insulin (Sigma# I-1882), and 100ng/mL cholera toxin (Sigma# C-8052). Cells were cultured at 37°C in 5% CO_2 in a humidified incubator, and passaged every 2–3 days with 0.25% trypsin (Corning# 25-053-CI) to maintain subconfluency. Starvation medium was purely DMEM/F12 medium. Cell line identity was confirmed by STR profiling.

In-Cell ELISA Perturbation Time-Course Experiments

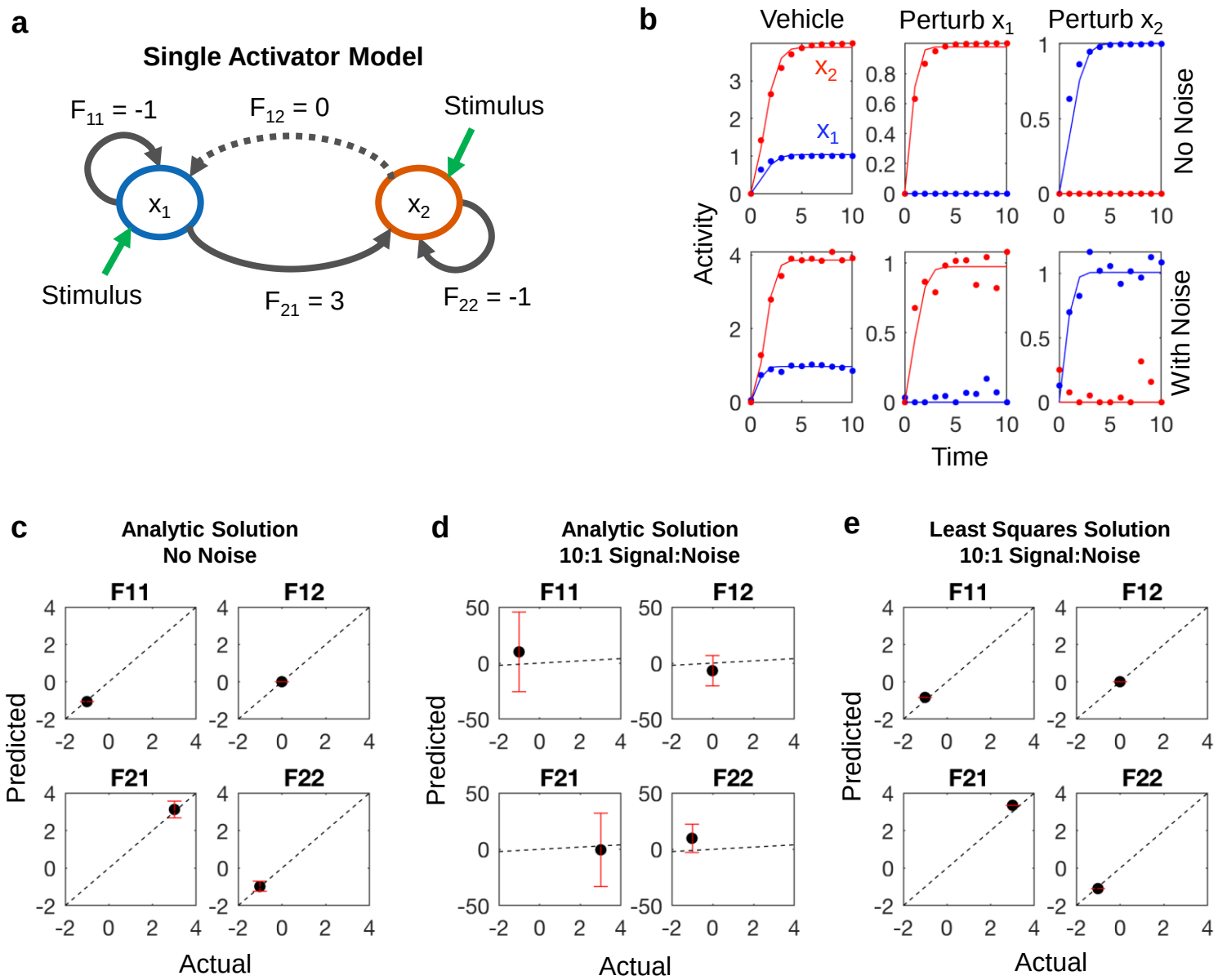
MCF10A cells were seeded in 96-well plates (Corning# 3603), at 10,000 cells per well, in serum containing medium (see above). Cells were serum- and growth factor-starved overnight (~18 hrs) and stimulated with 20ng/mL EGF (PeproTech# AF-100-15) and 10μg/mL insulin (Sigma# I-1882) for several different time periods (5, 30, 60, 90, 120, 240, and 360 minutes). For conditions receiving inhibitors, cells were pre-incubated with either 0.1 nM MEK- (PD0325901, Sigma# PZ0162) or 10 nM AKT-inhibitor (MK2206, ChemieTek# CT-MK2206) for 30 minutes prior to stimulation with growth factors. Cells were then fixed with 4% paraformaldehyde (EMS# 15710-S) for 15 minutes at room temperature, washed 3x with 200uL of 1x PBS, then permeabilized with 200uL of 1:100 dilution of Triton X-100 (Sigma# X100-500ML) in 1x PBS for 30 minutes at room temperature. Permeabilization buffer was then removed and each well was incubated with 200uL of Odyssey Blocking Buffer (TBS; LICOR #927-5000) for 2 hours at room temperature. The blocking buffer was removed and cells were incubated overnight at 4°C with 100 uL of a primary antibody solution composed of either anti-ppERK (Cell Signaling# 4370, 1:1000) or anti-pAKT(S473) (Cell Signaling# 4060, 1:1000) antibodies diluted directly in the Odyssey Blocking Buffer. The next day, primary antibody solution was removed, cells were washed 3x with 250uL of 1x wash buffer (625uL of Tween-20, Sigma# P1379-500ML, was diluted in 250mL of 1x PBS to make a stock solution). Cells were incubated with 100uL of secondary antibody solution (Goat Anti-Rabbit IgG Horseradish Peroxidase Thermo# 31460, 1:1000 dilution in Odyssey Blocking Buffer then 1:10 in 1X PBS), was added for 2 hours at room temperature. Secondary antibody solution was removed and cells were washed 4x with 250uL of 1x wash buffer. Next, 100uL of OPD substrate solution (5mg of o-Phenylenediamine, Sigma# P-2903, diluted in 10mL citric phosphate buffer pH 5 with 10uL of 30% hydrogen peroxide added immediately prior to use) was added to each well and the plate was immediately

put into the plate reader (PerkinElmer Ensight Model# H3570) for kinetic reading. Absorbance was measured at 450nm every two minutes, with shaking in-between each reading, for a total of 30 minutes. Absorbance values obtained during the linear range of the reaction were used. Serum-starved (control) absorbance values were subtracted from all values (separately for pERK and pAKT), which normalized baseline (the zero time point) signal to zero. The 60-minute time point was removed due to the presence of outliers and unacceptable quality control metrics.

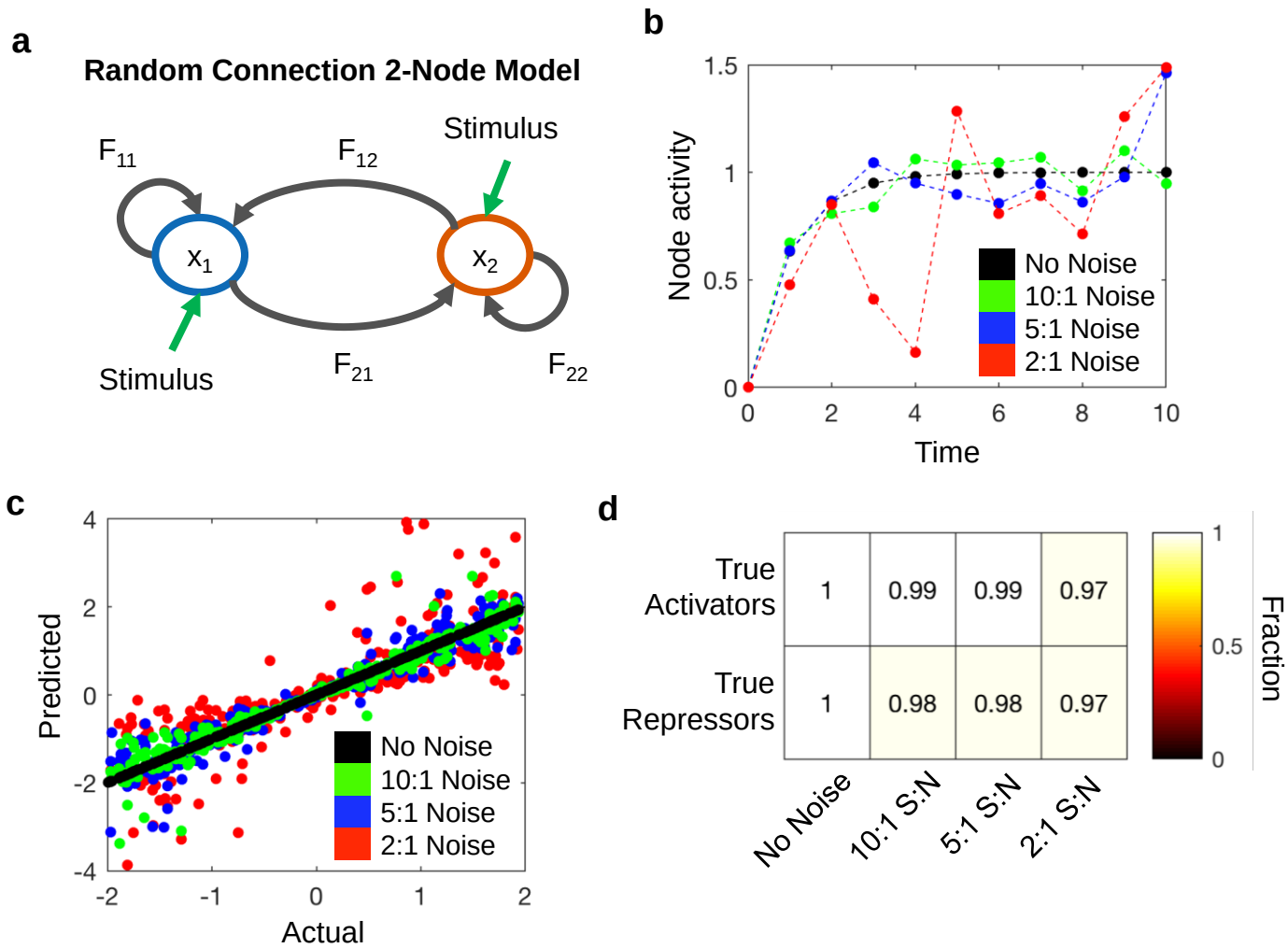
Smith et al., 2018. Figure 1



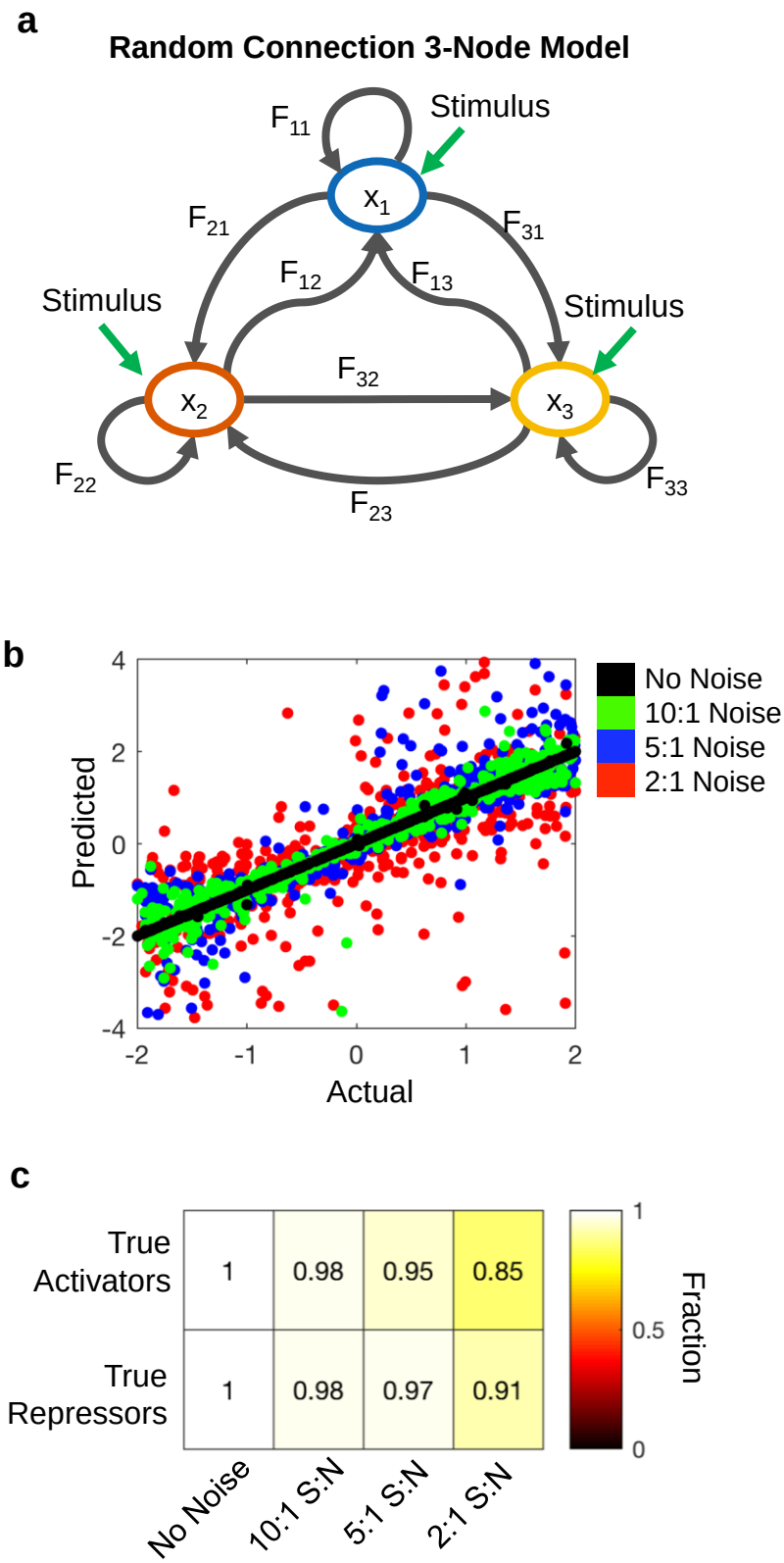
Smith et al., 2018. Figure 2



Smith et al., 2018. Figure 3



Smith et al., 2018. Figure 4

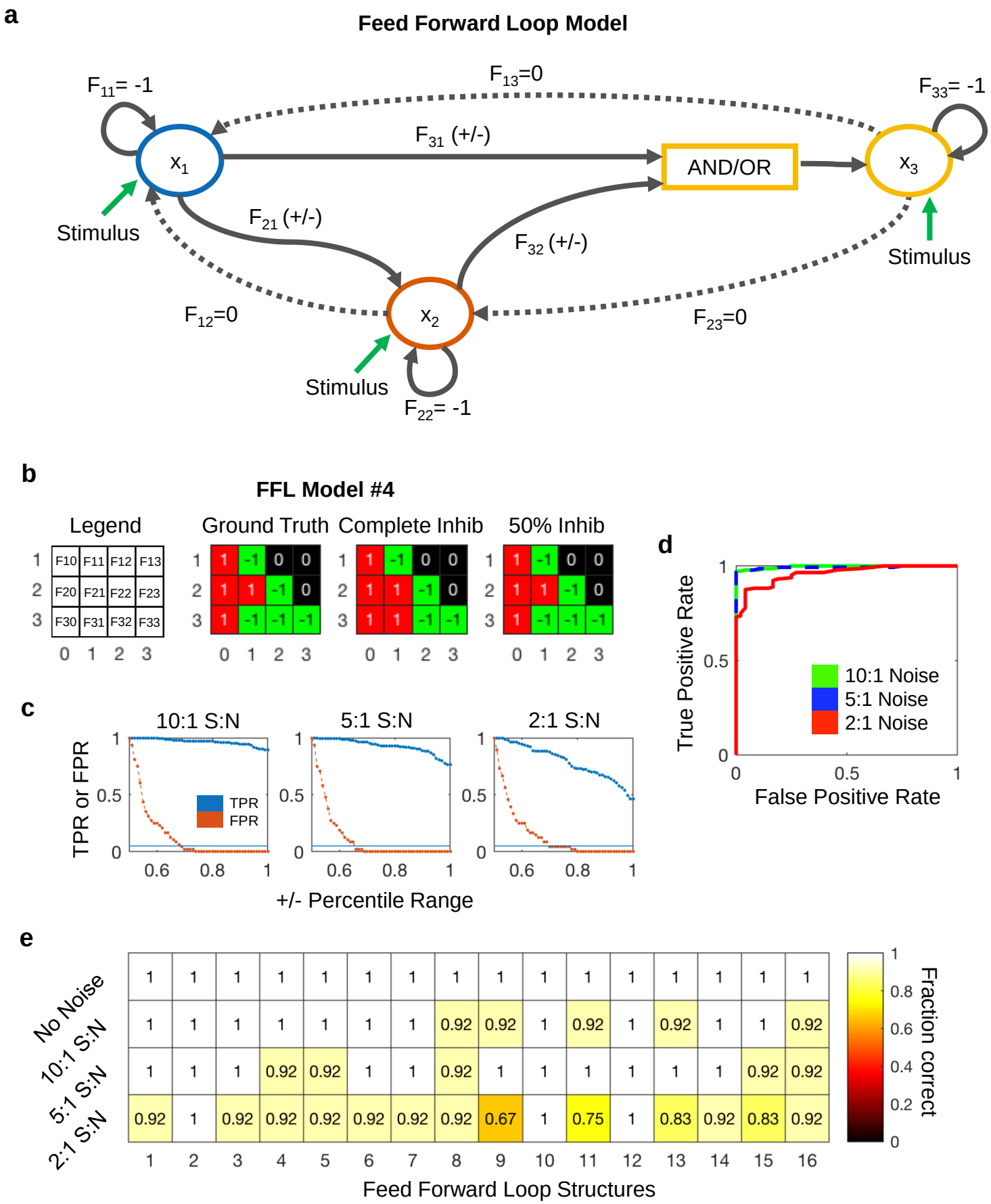


Smith et al., 2018. Table 1

Index	F_{21}	Gate	F_{31}	F_{32}
1	Activator	AND	Activator	Activator
2	Activator	AND	Activator	Repressor
3	Activator	AND	Repressor	Activator
4	Activator	AND	Repressor	Repressor
5	Activator	OR	Activator	Activator
6	Activator	OR	Activator	Repressor
7	Activator	OR	Repressor	Activator
8	Activator	OR	Repressor	Repressor
9	Repressor	AND	Activator	Activator
10	Repressor	AND	Activator	Repressor
11	Repressor	AND	Repressor	Activator
12	Repressor	AND	Repressor	Repressor
13	Repressor	OR	Activator	Activator
14	Repressor	OR	Activator	Repressor
15	Repressor	OR	Repressor	Activator
16	Repressor	OR	Repressor	Repressor

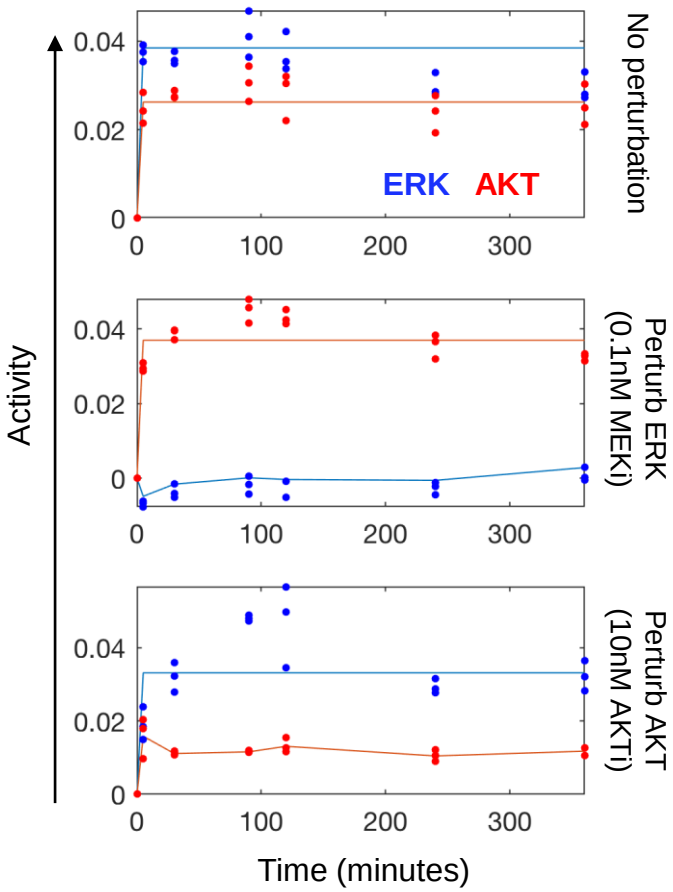
Table 1. The 16 Feedforward Loop Network Structures.

Smith et al., 2018. Figure 5

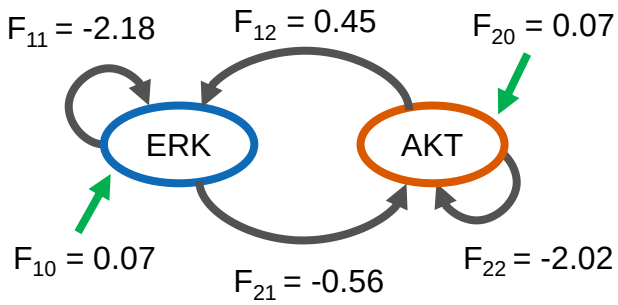


Smith et al., 2018. Figure 6

a



b



Smith et al., 2018. Figure S1

Jacobian Values for Random Models

Random 2-Node Models

	F10	F11	F12	F20	F21	F22
1	0.35	-1.14	1.01	1.7	-0.71	-0.24
2	1.71	-0.71	-0.47	0.47	1.74	-1.34
3	0.98	-0.06	0.14	1.8	-0.07	-1.71
4	1.68	-1.39	1.26	0.19	0.11	-1.32
5	1.53	-0.96	-0.54	0.54	-0.04	-0.18
6	1.71	-0.64	-0.42	1.54	1.78	-1.5
7	1.36	-1.13	0.14	1.23	1.45	-1.89
8	1.69	-1.23	0.68	1.07	-0.04	-0.79
9	0.59	-0	-0.39	1.25	0.61	-1.74
10	0.65	-0.52	0.78	0.78	-0.29	-1.11
11	1.17	-1.02	1.74	0.84	0.6	-1.82
12	0.74	-1.21	0.18	0.81	1.55	-0.7
13	0.59	-0.6	0.01	1.8	-0.03	-1.78
14	0.47	-1.2	0.05	1.28	-1.45	-0.95
15	0.13	-1.05	0.43	1.66	0.75	-1.81
16	1.58	-1.97	1.93	0.02	1.25	-1.56
17	1.73	-0.67	-1.16	0.92	0.37	-1.96
18	0.74	-1.6	1.72	1.51	-1.64	-0.62
19	0.86	-1.17	0.84	1.07	1.11	-1.42
20	0.75	-1.77	-0.4	0.41	1.22	-1.29
21	0.63	-0.13	-1.45	0.19	0.16	-1.45
22	1.75	-0.67	-0.52	1.57	0	-1.64
23	0.01	-1.35	1.67	0.17	-0.58	-0.9
24	0.01	-0.73	0.8	0.51	-0.75	-0.82
25	1.66	-0.8	0.39	1.03	0.15	-0.55
26	1.59	-0.96	-0.35	1.88	-0.57	-1.75
27	1.05	-0.38	0.21	0.54	1.17	-1.87
28	1.29	-1.34	0.31	1.14	-0.25	-0.93
29	0.41	-1.65	-0.17	0.24	1.23	-1.43
30	1.13	-1.88	1.49	0.88	0.86	-1.73
31	1.18	-1.22	0.45	1.27	-0.92	-0.18
32	1.87	-0.71	1.91	1.03	0.05	-1.53
33	0.47	-0.88	0.26	0.9	1.38	-1.85
34	1.42	-0.75	0.57	0.84	0.53	-1.03
35	0.09	-1.08	0.31	0.81	-1.35	-0.52
36	0.81	-1.08	-0.11	0.9	-0.08	-1.29
37	1.18	-1.46	0.92	0.52	0.39	-1.81
38	0.58	-1.42	1.65	1.85	-1.29	-1.85
39	1.46	-0.48	-0.7	0.04	0.4	-0.66
40	1.92	-1.31	-0.44	1.62	-0.77	-1.68
41	0.91	-0.75	-0.01	1.12	1.57	-1.29
42	1.02	-1.77	0.48	0.52	-0.75	-1.64
43	0.29	-1.97	0.52	1.71	0.42	-1.09
44	1.91	-1.82	1.28	1.61	-0.78	-0.59
45	0.97	-1.03	-0.04	1.74	1.25	-0.73
46	0.91	-0.52	-0.3	0.19	-0.03	-0.2
47	0.05	-1.99	0.77	0.83	0.95	-0.61
48	0.94	-1.42	-0.81	0.13	1.15	-0.71
49	1.62	-1.75	-1.56	0.29	0.67	-1.54
50	1.08	-1.51	0.17	1.79	-0.36	-0.55

Random 3-Node Models

	F10	F11	F12	F13	F20	F21	F22	F23	F30	F31	F32	F33
1	0.09	-1.34	0.04	1.47	1.53	-1.46	-1.74	0.69	0.96	-0.02	-0.75	-0.71
2	1.12	-1.21	-0.42	-0.64	0.71	-0.51	-1.54	0	1.63	-0.63	-0.01	-1.53
3	0.92	-0.59	-1.71	0.02	0.55	0.27	-1.89	0.45	0.85	0.58	-0.78	-0.84
4	0.76	-1.64	0.93	0.73	0.87	-1.43	-0.77	0.72	1.01	0.94	0.59	-1.73
5	1.24	-1.86	1.4	-1.44	0.46	-0.5	-1.19	0.67	0.25	1.35	0.81	-0.82
6	1.95	-1.18	0.17	-1.92	1.16	-0.26	-1.23	1.72	1.62	0.67	-0.8	-1.77
7	1.73	-1.75	0.43	-0.34	0.83	-0.51	-1.03	0.03	1.96	-0.17	1.91	-1.83
8	0.91	-0.1	-0.09	-0.49	2	0.6	-1.76	-1.28	0.77	0.02	1.58	-0.86
9	0.93	-1.67	0.72	-1.64	0.11	1.95	-1.06	1.25	0.28	1.15	0.15	-1.06
10	0.61	-1.5	-0	1.05	1.33	0.49	-1.24	-0.31	1.91	-1.46	0.24	-1.81
11	0.03	-1	0.62	0.01	0.83	1.06	-1	-1	0.99	1.92	0.91	-1.45
12	0.23	-1.56	0.4	-0.52	0.4	0.18	-0.38	1.55	0.65	-0.52	-0.24	-1.56
13	1.16	-1.5	1.77	-0.71	1.44	1.32	-1.75	-1.05	1.21	1.32	0.75	-1.81
14	1.13	-1.47	1	0.46	1.49	-1.34	-1.26	1.73	0.53	1.48	0.48	-1.72
15	0.73	-0.29	-1.06	0.52	0.16	1.49	-0.88	-0.01	0.23	0.66	0.29	-1.43
16	1.3	-1.04	1.02	-0.87	0.82	-0.09	-1.93	0.77	1.14	1.72	-0.77	-1.87
17	0.96	-0.9	0.95	1.29	0.54	0.05	-2	-0.54	0.6	-0.25	-0.21	-0.61
18	0.55	-1.54	-0.05	1.22	1.94	-0.43	-1.48	1.09	0.24	0.49	0.41	-1.07
19	1.18	-0.83	-1.04	-0.13	0.81	1.68	-1.88	0.96	1.32	0.77	-1.82	-1.05
20	1.42	-1.82	-1.22	0	0.23	0.11	-0.94	0.44	0.95	0.78	-0.3	-1.2
21	0.4	-0.14	-0.16	-0.15	0.92	0.68	-1.08	-0.02	1.74	-0.24	0.23	-1.9
22	1.17	-1.27	1.31	0.71	0.78	-0.2	-1.88	-0.2	1.7	-0.75	1.63	-0.47
23	1.36	-1.57	1.36	0.29	1.37	0.02	-1.75	1.39	1.06	-0.86	0.85	-1.66
24	0.98	-1.86	-0.25	-0.13	1.97	1.91	-1.01	0.42	1.6	-1.02	-0.2	-0.56
25	0.4	-1.82	-0.27	1.52	0.69	0.1	-1.01	-0.16	1.76	0.06	1.51	-1.46
26	1.43	-0.82	1.17	-1.08	1.81	-0.64	-1.43	0.94	0.15	0.28	0.41	-1.39
27	1.23	-1.51	1.44	-0.2	1.51	0.6	-1.15	-0.76	0.9	0.93	1.5	-1.55
28	1.27	-1.03	-0.11	0.5	1.41	1.65	-0.52	-1.43	0.62	1.82	1.88	-1.78
29	0.68	-1.8	0.65	0.2	1.91	-1.49	-1.68	-0.52	1.13	-1.25	-0.3	-1.17
30	1.34	-0.82	-0.86	-1.86	0.76	0.8	-1.49	-1.7	0.18	0.5	0.47	-0.65
31	1.57	-1.83	1.38	0.15	0.48	-0.35	-1.41	1.16	1.85	-0.55	-1.41	-1
32	0.87	-0.52	-0.2	0.16	1.74	-0.68	-0.6	0.3	1.97	1.17	0.04	-1.31
33	0.11	-1.05	1.21	-0.08	1.5	-1.11	-0.28	0.34	0.08	1.34	0.29	-0.65
34	0.57	-1.23	0.53	1.77	0.23	0.95	-1.02	1.46	1.18	-0.05	0.1	-1.87
35	0.39	-1.56	1.99	-1.68	0.11	0.48	-1.38	1.83	0.41	0.19	0.72	-1.89
36	0.25	-1.72	-0.09	-0.13	1.35	-0.47	-0.88	-1.9	0.12	0.1	1.22	-0.92
37	0.45	-1.53	0.97	0.07	0.12	-0.35	-0.75	0.09	0.66	0.62	1.36	-1.92
38	0.32	-1.85	0.36	1.28	0.5	-0.03	-1.87	1.67	1.72	0.74	0.28	-1.14
39	1.77	-0.7	-1.09	-0.2	0.25	0.61	-1.96	0.22	1.76	1.23	-1.74	-0.88
40	0.97	-0.97	0.36	-0.95	1.45	-0.22	-0.49	-0.36	0.34	0.02	0.87	-1.15
41	0.45	-1.7	-0.04	0.66	1.21	-1.36	-1.24	0.72	1.28	-1.16	0.32	-1.55
42	0.47	-0.45	0.29	-0.25	1.18	-0.84	-0.82	-0.24	0.37	1.31	0.05	-1.51
43	1.88	-1.55	0.77	1.95	1.24	-0.06	-1.8	-0.06	1.65	-0.37	-0.61	-1.45
44	1.09	-0.63	1.46	-1.63	0.86	0.57	-1.66	-0.03	0.8	1.9	0.13	-1.36
45	1.6	-1.34	0.2	-0.09	1.76	-0.56	-1.9	1.7	1.96	-0.08	-0.63	-1.23
46	0.95	-1.25	0.14	1	1.32	0.73	-1.46	-0.32	1.56	-0.76	1.4	-1.62
47	0.35	-0.67	0.62	0.71	0.25	1.24	-1.86	1.63	1.04	0.81	-1.27	-1.32
48	0.64	-1.57	0.64	-1.55	0.08	0.22	-1.39	1.6	0.12	1.36	0.66	-1.61
49	0.27	-1.84	0.99	1.97	1.82	-1.58	-1.31	1.56	1.91	0.41	-0.12	-1.08
50	0.41	-1.62	0.02	1.67	0.62	1	-1.84	0.42	1.87	-1.19	-0.86	-0.31

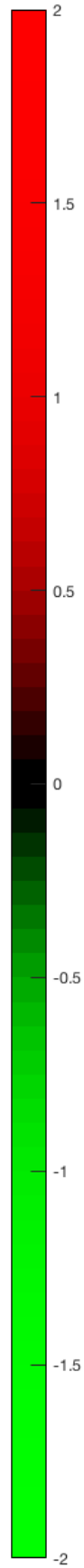


Figure S1. The randomly sampled theoretical (ground truth) Jacobian values used for each of the 50 iterations for the random 2-node (left) and 3-node (right) model.

Smith et al., 2018. Figure S2

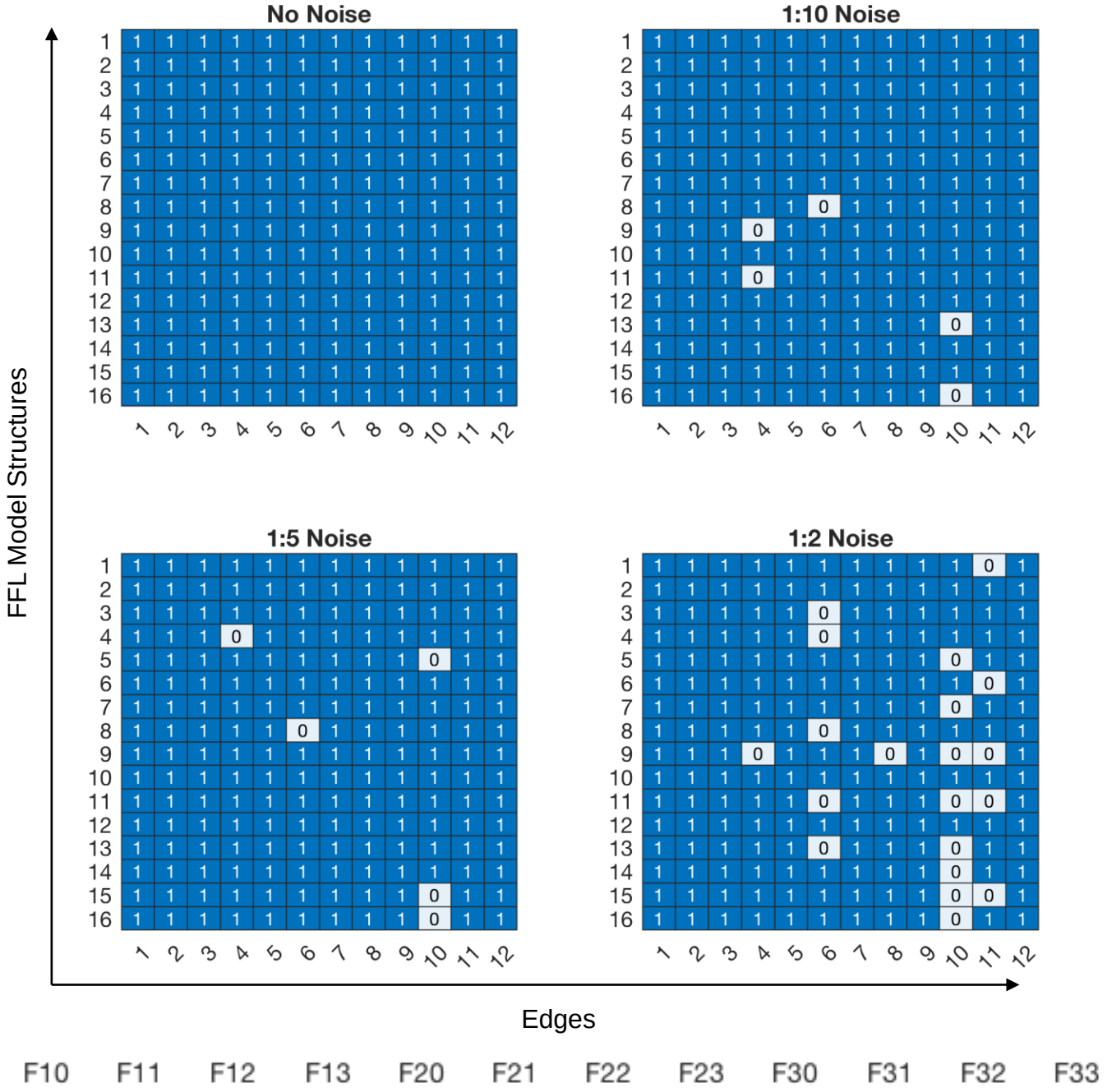


Figure S2. Detailed results from the Feed Forward Loop models from Figure 5 depicting whether an edge (1-12, columns) was correctly (1) or incorrectly (0) predicted for each model structure (1-16, columns) under each noise level. Correct predictions were classified based on the optimal percentile cutoffs identified for each noise level.

Smith et al., 2018. Figure S3

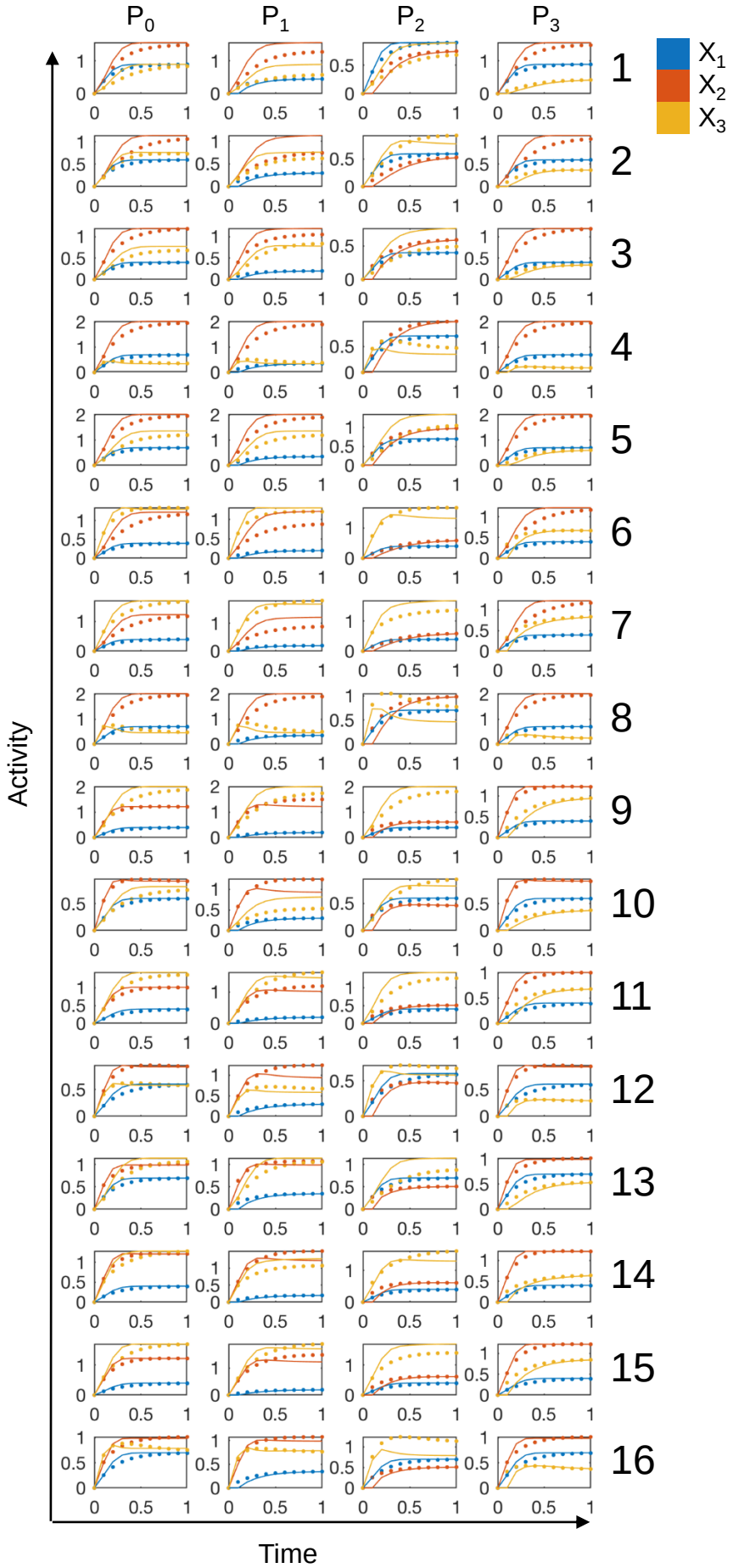


Figure S3. Simulated experimental data (dots) and model-generated fits (lines) for each model structure. Different perturbations—vehicle (P0), perturb x1 (P1), perturb x2 (P2), perturb x3 (P3)—are across the columns and different model structures (1-16) are down the rows. Each node (1-3) is a different color as indicated in the legend.

Smith et al., 2018. Figure S4

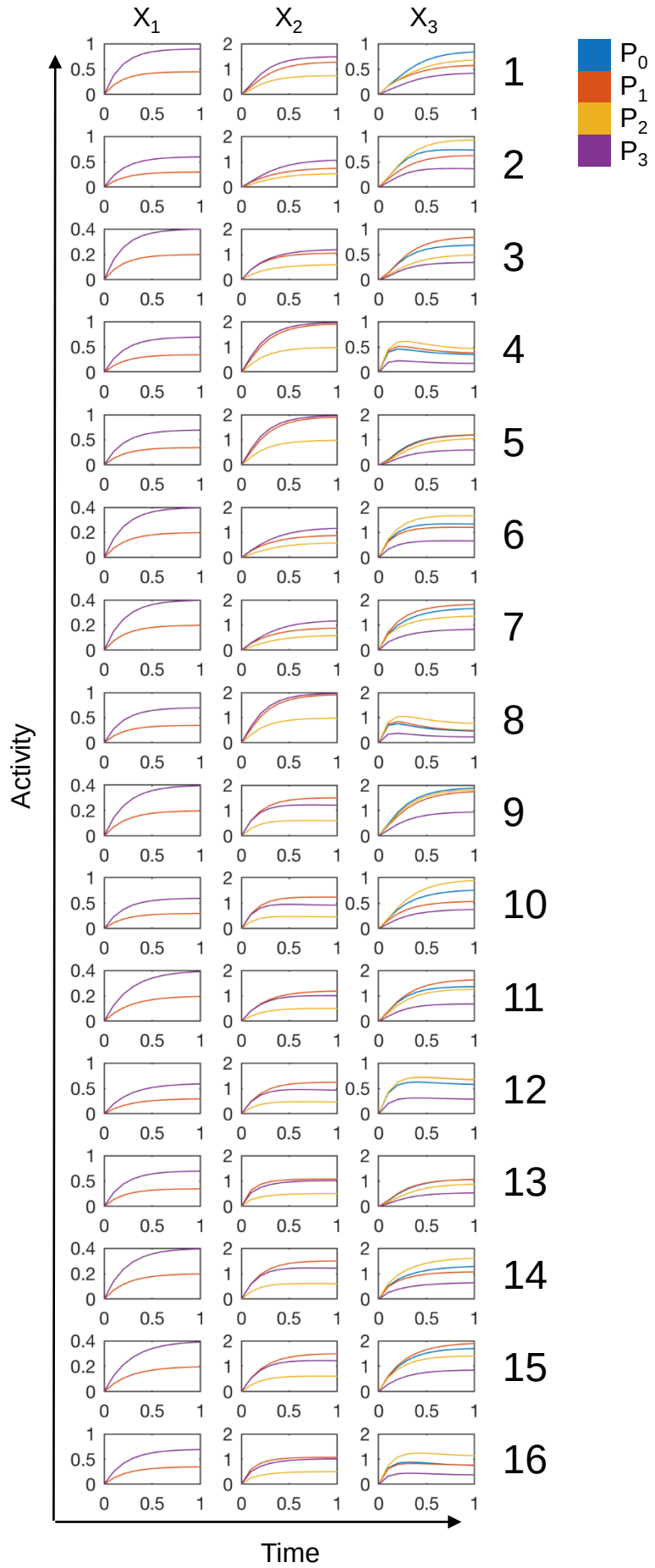


Figure S4. Simulated experimental data (lines) broken down by node (1-3) along the columns and model structure (1-16) down the rows. Different colors represent different perturbation schemes as indicated by the legend. The parameters of the models are selected based on their generation of responsive time courses for each node and perturbation type.