

## SUPPLEMENTAL MATERIAL

### Algorithm for model reduction

The process for reducing complex models described in the text begins with a statistical model  $P(\xi|\theta)$  and a set of parameters  $\theta$  as a starting point. The parameters  $\theta$  may originate from fitting the model to available data, in which case there will often be large uncertainties associated with them (see the discussion of sloppiness in the main text). Model reduction will then compress these uncertainties into the relevant parameter combinations. In other cases, the parameter values may originate from a systematic exploration of a parameter space that includes a variety of physical systems as specific realizations. In this case, model reduction will identify which physical systems have the same collective behavior. Whatever their origin, from the combination of model and parameter values, the Manifold Boundary Approximation Method (MBAM) proceeds as an iterated four step algorithm.

First, given a statistical model  $P(\xi|\theta)$  and set of parameters  $\theta$ , calculate the Fisher Information Matrix (FIM) and identify the eigendirection with smallest eigenvalue. Second, numerically calculate a geodesic as the solution of an initial value problem (described below) using the parameter values and eigendirection as the initial conditions. (Since the eigendirection defines two directions, in practice we select the direction in which the parameter space norm of the velocity vector is increasing, i.e. the direction most aligned with the geodesic acceleration.) Follow this path until a boundary is identified by the Fisher Information becoming singular. Third, analytically evaluate the limit associated with the boundary identified in the second step to produce a new model with one less parameter. Fourth, calibrate the new model by fitting it to the behavior of the original model (for example, by least squares fitting). Use this new model and fitted parameter values as a starting point to repeat the four steps until the reduced model is satisfactorily simple.

The FIM is given by  $g_{\mu\nu} = -\langle \partial_\mu \partial_\nu l \rangle = \langle \partial_\mu l \partial_\nu l \rangle$ , where  $\langle \cdot \rangle$  means expectation value with respect to  $P(\xi|\theta)$ ,  $l = \log P(\xi|\theta)$ , and  $\partial_\mu$  means differentiation with respect to  $\theta_\mu$ . For several classes of models, more explicit formulas can be given. For example, a special but important case is that of a deterministic model with additive random gaussian noise (such as would correspond to least squares data fitting). In this case, if the deterministic model makes  $M$  predictions  $y_m(\theta)$ , where  $m = 1 \dots M$ , each of which is measured with experimental uncertainty  $\sigma_m$ , then the Fisher Information is given by  $g = (J^T J)$ , where the Jacobian matrix  $J$  has components given by  $J_{m\mu} = (\partial y_m / \partial \theta_\mu) / \sigma_m$ . The exponential model from the text as well as the EGFR network model both fall within this class.

Another example is a Hamiltonian system distributed according to the Boltzmann distribution in which the Hamiltonian is linear in the parameters (such as the Ising model considered in the text),  $H(\xi, \theta) = \sum_\mu \theta_\mu \phi_\mu(\xi)$ , where  $\xi$  describes the system's configuration and  $\phi_\mu$  are some functions of the configuration. Then the probability that the system be found in state  $\xi$  is given by

$$P(\xi|\theta) = \exp(-H(\xi, \theta)) / Z, \quad (1)$$

where  $Z$  is a normalization constant which depends on  $\theta$ , but not  $\xi$ . In this case, the FIM is given by  $g_{\mu\nu} = \partial_\mu \partial_\nu \log Z = -\partial_\mu \partial_\nu F$ , where  $F$  is the Hemholtz free energy.

Because the eigenvalues of the Fisher Information are often spread over many orders of magnitude, the smallest eigenvalues and eigenvectors may be particularly susceptible to numerical noise. For the examples in the main text, we did not calculate the Fisher Information using finite differences. For the EGFR model, we numerically integrated the differential equations corresponding to the sensitivities of the model with respect to the parameters. Because most of the code used in our simulations was generated automatically with computer algebra systems, this process was not much more computationally intensive than evaluating finite differences but produced more accurate results. For the other models, explicit expressions could be found for the FIM.

Computationally, the most difficult aspect of this method is numerically identifying the manifold boundary via a geodesic. We summarize this processes here, more information is available in references[1, 2]. A geodesic is given as a parameterized curve through parameter space satisfying the differential equation

$$\ddot{\theta}^\mu + \Gamma_{\alpha\beta}^\mu \dot{\theta}^\alpha \dot{\theta}^\beta = 0, \quad (2)$$

where we have used the Einstein summation convention in which repeated indices are summed and dots represent derivatives with respect to the curve's parameterization. In Eq. (2), components of  $\Gamma$  are the connection coefficients which are given by

$$\Gamma_{\alpha\beta}^\mu = \frac{1}{2} g^{\mu\nu} (\partial_\alpha g_{\nu\beta} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta}) \quad (3)$$

and  $g^{\mu\nu}$  is the inverse Fisher Information. Note that  $\Gamma$  (just like the metric) can be expressed in terms of expectation values

$$\Gamma_{\alpha\beta}^\mu = g^{\mu\nu} \left\langle \left( \partial_\alpha \partial_\beta l + \frac{1}{2} \partial_\alpha l \partial_\beta l \right) \partial_\nu l \right\rangle. \quad (4)$$

It is noteworthy that although the expression for  $\Gamma$  involves second derivatives of the model with respect to the parameters, because the second derivatives are each contracted with the velocity  $\dot{\theta}$  in Eq. (2) it is in general only necessary to calculate a directional second derivative, which for large models is a major computational

simplification[1, 2]. We typically find that estimating this second derivative with finite difference is adequate. For the special case of least squares data fitting, the connection coefficients are given by

$$\Gamma_{\alpha\beta}^{\mu} = g^{\mu\nu} \sum_{m=1}^M \partial_{\nu} y_m \partial_{\alpha} \partial_{\beta} y_m / \sigma_m^2. \quad (5)$$

In practice, one need only calculate the solution to Eq. (2) until the limiting behavior is identified. Consequently, this equation need not be integrated very accurately. We find that numerically integrating the geodesic path with large error tolerances is often satisfactory. Furthermore, in many cases one can identify the limiting behavior by looking at the initial velocity  $\dot{\theta}$  (FIM eigendirection) without calculating the geodesic at all. For the protein signaling model described in the text, we were able to identify about half of the limits without resorting to geodesics.

### Examples of Identifying and Evaluating Geodesic Limits

Because the manifold boundary approximation customizes itself to the functional form the model, it is not immediately transparent how this process is accomplished in real-world scenarios. Typically, manifold boundaries correspond to limits in which one or more parameters approach physically limiting values (such as infinity or zero) in a coordinated way. A combination of relevant parameters then emerges based on the functional form of the model such that the model prediction remain finite in the limit. Identifying this combination, therefore involves inspecting the functional form of the model. We here consider several models of various functional forms to illustrate the process. The variety of models considered here further demonstrates the generality of the MBAM.

The first three examples we present here are drawn from the MINPACK-2 suite of test problems[3]. This is a standard ensemble of models drawn from a variety of different research contexts originally intended to benchmark data fitting algorithms. We then consider models of an enzyme-substrate reaction leading to Michaelis-Menten type approximation, a multi-state partition function leading to a step function approximation, and a particle-lattice model leading to thermodynamic and continuum limit approximations.

First we consider a model of a thermistor resistance as a function of temperature,  $T$ , which takes the form

$$y(T, \theta) = \theta_1 e^{\theta_2 / (T + \theta_3)}. \quad (6)$$

Details of this model's implementation (including temperature values and data points) can be found in reference[3] for this and subsequent models. A geodesic

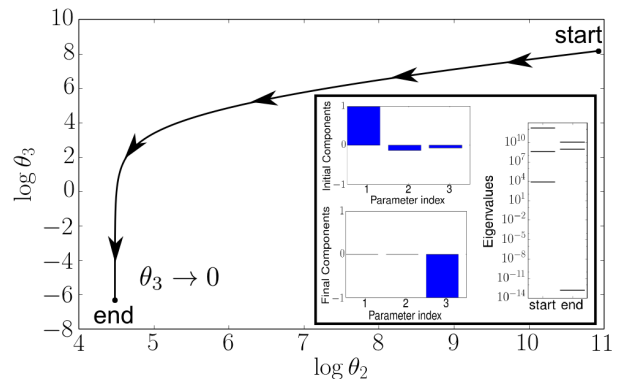


FIG. 1: **Geodesic limit for a model of thermistor resistance.** Although initial oriented primarily with the first parameter, the geodesic curves to approach a limit in which  $\theta_3 \rightarrow 0$ .

originating from the best fit parameters takes a path as illustrated in Figure 1. The eigenvalues of the Fisher Information indicate an initial insensitivity to fluctuations in the first parameter  $\theta_1$ , i.e. the normalization constant. The geodesic, however, reveals a limiting boundary to be  $\theta_3 \rightarrow 0$ . The resulting simplified model is found by evaluating this limit and takes the form

$$y(T, \phi) = \phi_1 e^{\phi_2 / T}, \quad (7)$$

where we have used  $\phi$  to denote the reduced model parameters, a convention we adopt throughout this section.

Because this limit involves only one parameter, it is not difficult to identify or evaluate. However, it does help to illustrate the relationship between the manifold boundaries and the model predictions. Specifically, this approximation will be accurate as long as the model does not explore the low-temperature behavior of the system. Indeed, if the original model were used to predict the system behavior at  $T = 0$ , the limit observed in Figure 1 would not be a manifold boundary because some of the model predictions would be unbounded in this limit. As we consider more subtle examples below, we repeatedly use the fact that model predictions approach a well-defined limit at the boundary in order to identify the appropriate combination.

We next consider a model to analyze the kinetics of an enzyme reaction. It has four parameters and takes the form

$$y(u, \theta) = \frac{\theta_1 (u^2 + \theta_2 u)}{u^2 + u\theta_3 + \theta_4}, \quad (8)$$

where  $u$  is substrate concentration of the reaction. A geodesic beginning from the best fit parameter values is given in Figure 2. In this case, the limit involves two parameters,  $\theta_1$  and  $\theta_2$  approaching 0 and  $\infty$  respectively. Considering the functional form in Eq. (8) suggests that the combination  $\theta_1 \times \theta_2$  is a likely candidate for the new

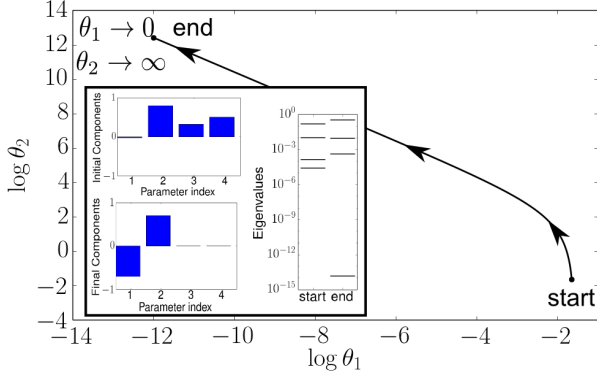


FIG. 2: **Geodesic limit for a model of enzyme kinetics.** This limit involves two parameters approaching *zero* and  $\infty$  in a coordinated way. The appropriate nonlinear combination,  $\phi_1 = \theta_1\theta_2$  is then deduced by considering the functional form of the model and the behavior of the model predictions as the limit is approached numerically.

parameter. Indeed, inspecting this parameter combination numerically in the limit confirms that it remains finite. Furthermore, all of the model predictions that depend only on this combination also remain finite. Therefore, we define a new parameters  $\phi_1 = \theta_1\theta_2$ , which is the nonlinear combination that is retained in the reduced model. (We take  $\phi_2 = \theta_3$  and  $\phi_3 = \theta_4$ .) The simplified functional form, after taking  $\theta_1 \rightarrow 0$  is then

$$y(u, \phi) = \frac{\phi_1 u}{u^2 + \phi_2 u + \phi_3}. \quad (9)$$

We now consider a model describing the chemical kinetics of thermal isomerization of  $\alpha$ -pinene. The parameters are the reaction coefficients and the model is constructed as a set of five coupled linear differential equations

$$\frac{d}{dt}y_1 = -(\theta_1 + \theta_2)y_1 \quad (10)$$

$$\frac{d}{dt}y_2 = \theta_1 y_1 \quad (11)$$

$$\frac{d}{dt}y_3 = \theta_2 y_1 - (\theta_3 + \theta_4)y_3 + \theta_5 y_5 \quad (12)$$

$$\frac{d}{dt}y_4 = \theta_3 y_3 \quad (13)$$

$$\frac{d}{dt}y_5 = \theta_4 y_3 - \theta_5 y_5. \quad (14)$$

Geodesics originating from the best fit are illustrated in Figure 3. In this limit the parameters  $\theta_4$  and  $\theta_5$  both become infinite. By considering the functional form of the equations together with numerical exploration of the limit, we find that the combination  $\theta_4/\theta_5$  remains finite, leading to a new parameters  $\phi_4 = \theta_4/\theta_5$  while  $\phi_i = \theta_i$  for  $i = 1, 2$ , and 3.

Dividing the equation for  $\frac{d}{dt}y_5$  by  $\theta_5$  and taking the

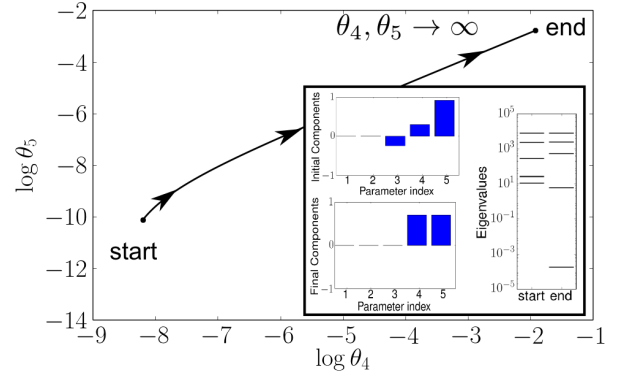


FIG. 3: **Geodesic limit for a model of the isomerization of the  $\alpha$ -pinene.** As two reaction rates become infinite, the concentration of  $y_3$  and  $y_5$  approach an equilibrium ratio determined by the new parameter  $\phi_4 = \theta_4/\theta_5$ .

limit gives the algebraic relation

$$y_5 = \phi_4 y_3. \quad (15)$$

Numerical results confirm that this relation is approached in the geodesic boundary limit. The parameters  $\theta_4$  and  $\theta_5$  can then be removed from the equation for  $\frac{d}{dt}y_3$  by adding to it the equation for  $\frac{d}{dt}y_5$ . The result is

$$\frac{d}{dt}y_3 + \frac{d}{dt}y_5 = \phi_2 y_1 - \phi_3 y_3. \quad (16)$$

The term  $\frac{d}{dt}y_5$  can then be removed using Eq. (15), giving

$$\frac{d}{dt}y_3 = \frac{\phi_2}{1 + \phi_4} y_1 - \frac{\phi_3}{1 + \phi_4} y_3, \quad (17)$$

so that the reduced model now takes the form

$$\frac{d}{dt}y_1 = -(\phi_1 + \phi_2)y_1 \quad (18)$$

$$\frac{d}{dt}y_2 = \phi_1 y_1 \quad (19)$$

$$\frac{d}{dt}y_3 = \frac{\phi_2}{1 + \phi_4} y_1 - \frac{\phi_3}{1 + \phi_4} y_3 \quad (20)$$

$$\frac{d}{dt}y_4 = \phi_3 y_3 \quad (21)$$

$$y_5 = \phi_4 y_3. \quad (22)$$

We here point out a possible source of confusion. In simplifying the equation for  $\frac{d}{dt}y_3$ , it may be tempting to misinterpret the relation in Eq. (15) to imply that that  $\theta_5 y_5 - \theta_4 y_3 = 0$ . This is incorrect, however since the relation  $y_5 - (\theta_4/\theta_5)y_3 = 0$  only emerges in the limit that  $\theta_4, \theta_5 \rightarrow \infty$ . Multiplying this relation by  $\theta_5$  effectively multiplies zero by infinity which is indeterminate and not necessarily zero. The limit must be evaluated carefully by considering the original model. In particular, the combination  $\theta_5 y_5 - \theta_4 y_3 = -\frac{d}{dt}y_5$  must be true for all  $\theta_4$  and  $\theta_5$ , a fact we exploited above.

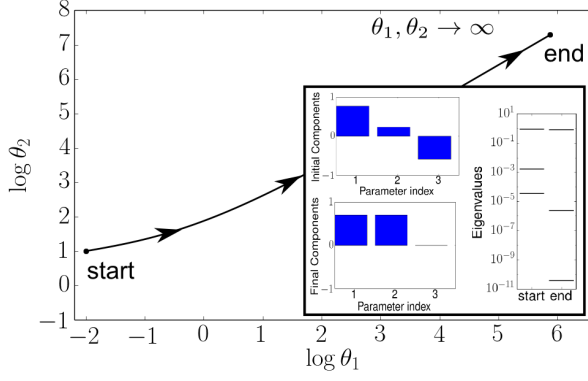


FIG. 4: **Geodesic limit for a model of an enzyme-substrate reaction.** As the parameters  $\theta_1$  and  $\theta_2$  become infinite, the concentration of enzyme, substrate, and intermediate complex approaches an equilibrium. Applying this equilibrium approximation results in the Michaelis-Menten rate law.

We now consider three models not taken from the MINPACK-2 test suite. First, we consider the time course kinetics of an enzyme-substrate reaction. This model appears similar to the  $\alpha$ -pinene model above. However, this example is modeled as a set of coupled nonlinear differential equations derived from mass-action kinetics. We include it here to illustrate how the manifold boundary approximation seamlessly handles the nonlinearity (and increased complexity) of this model. The model describes the chemical process by which an enzyme, E, and a substrate, S, bind (reversibly with rates  $\theta_1$  and  $\theta_2$ ) to form an intermediate complex, ES, which then disassociates (with rate  $\theta_3$ ) to leave an enzyme and a product, P. The equations for the concentrations (denoted by square brackets) of the four chemical species are then given by

$$\frac{d}{dt}[E] = -\theta_1[E][S] + (\theta_2 + \theta_3)[ES] \quad (23)$$

$$\frac{d}{dt}[S] = -\theta_1[E][S] + \theta_2[ES] \quad (24)$$

$$\frac{d}{dt}[ES] = \theta_1[E][S] - (\theta_2 + \theta_3)[ES] \quad (25)$$

$$\frac{d}{dt}[P] = \theta_3[ES]. \quad (26)$$

We take the three reactions rates as parameters. Note that two conservation laws,  $[E] + [ES] = E_0$  and  $[S] + [ES] + [P] = S_0$  where  $E_0$  and  $S_0$  are the total enzyme and substrate concentrations respectively, mean there are only two *independent* equations in this model. We will exploit this fact in the analysis that follows.

A geodesic for this model is given in Figure 4. The limit here, similar to that in Figure 3 involves two parameters becoming infinite with a finite ratio, i.e.  $\phi_1 = \theta_2/\theta_1$  is finite and the new parameter. We also take  $\phi_2 = \theta_3$ .

Evaluating this limit proceeds similarly to that in the

isomerization of the  $\alpha$ -pinene. Dividing the equation for  $\frac{d}{dt}[ES]$  by  $\theta_1$  and taking the limit gives the algebraic relation  $[ES] = [E][S]/\phi_1$ . Combining this relation with the enzyme conservation law gives the relation

$$E = \frac{E_0}{\phi_1 + [S]}. \quad (27)$$

The final model may be written simply as

$$\frac{d}{dt}([S] + [ES]) = -\frac{d}{dt}[P] = -\frac{\phi_3 E_0 S}{\phi_1 + S}, \quad (28)$$

where we have used the substrate concentration law. This equation is the celebrated Michaelis-Menten rate law derived from an equilibrium approximation. We therefore identify the combination  $\phi_1 = \theta_2/\theta_1$  as the saturation constant for the enzyme reaction.

A very similar form to that in Eq. (28) can also be derived from a quasi-steady state approximation. We note that the quasi-steady-state derivation can also be found as a manifold boundary if we also take the total concentrations  $E_0$  and  $S_0$  as parameters. A similar functional form emerges in the limit  $\theta_1, \theta_2$ , and  $\theta_3 \rightarrow \infty$  while  $E_0 \rightarrow 0$ .

Next, we consider a statistical mechanics model describing a three state system. The model describes the likelihood of finding the system in a state as a function of temperature assuming it obeys standard Boltzmann statistics. The functional form is given by

$$y(\beta, \theta) = \frac{e^{-\beta\theta_1 + \theta_2}}{1 + e^{-\beta\theta_1 + \theta_2} + e^{-\beta\theta_3 + \theta_4}}, \quad (29)$$

where  $\beta = 1/k_B T$  and the parameters  $\theta$  represent the energies and entropies of the states measured with respect to the ground state. A geodesic for this model is given in Figure 5 which reveals the limit that  $\theta_3, \theta_4 \rightarrow \infty$  with  $\phi_3 = \theta_4/\theta_3$  remaining finite and becoming the new parameter combination.

To evaluate this limit we consider the terms of the model in Eq. (29). Rewriting the last exponential term in the denominator as  $\exp(-\beta\theta_3 + \theta_4) = \exp(\theta_3(\phi_3 - \beta))$ , we see that in the limit that  $\theta_3 \rightarrow \infty$ , this term will evaluate to either infinity or zero, depending on the sign  $\phi_3 - \beta$ . Eq. (29) can therefore be simplified as

$$y(\beta, \phi) = \frac{e^{-\beta\phi_1 + \phi_2}}{1 + e^{-\beta\phi_1 + \phi_2}} \Theta(\beta - \phi_3), \quad (30)$$

where  $\Theta(x)$  is the heaviside function which is zero for  $x < 0$  and one for  $x > 0$ .

Finally, we consider a model that leads to a combination of a thermodynamic and continuum limit. Consider a large number of particles  $N$  distributed randomly at discrete sites on a one dimensional lattice with spacing  $a$ . Let the total length of the chain be  $L$  so that the number of lattice sites is  $L/a$ . We seek to the model

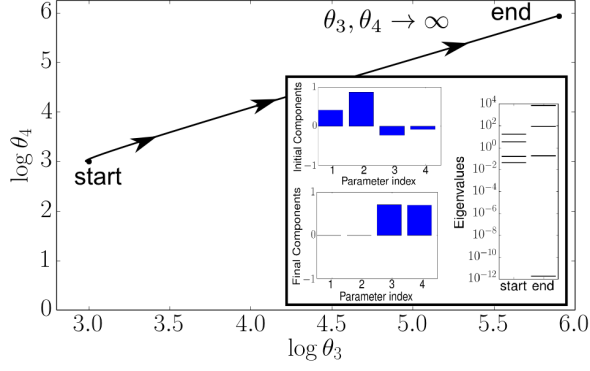


FIG. 5: **Geodesic limit of a multi-state statistical mechanics model.** At the manifold boundary, two infinite parameters are combined into a simple ratio. The simplified model becomes discontinuous in temperature and can be expressed using the heaviside function.

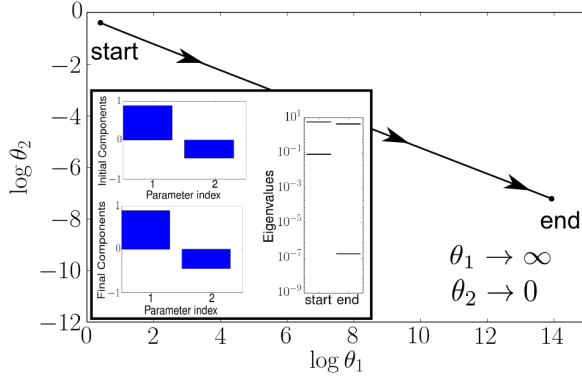


FIG. 6: **Geodesic limit of a Thermodynamic/Continuum Limit.** A model for the number density of particles on a lattice. The particle number can become infinite while the lattice spacing becomes zero in a coordinated way to maintain a finite density. This approximation is both a thermodynamic and a continuum limit.

the probability of observing a particular density of particles at each of the sites. For simplicity, we here assume that all sites are equally likely, so that the distribution of number of particles at a given site is binomial with average value  $Na/L$  and variance  $N(a/L)(1-a/L)$ . The probability of observing a number density  $\rho$  at a single site is approximately gaussian with mean  $Na^2/L^2$  and variance  $N(a/L)^2(1-a/L)$ . The geodesic for this model is given in Figure 6. It corresponds to the case in which  $\theta_1 = N \rightarrow \infty$  and  $\theta_2 = a \rightarrow 0$  such that the combination  $Na^2$  is finite. This is therefore both a thermodynamic limit (infinite particle number) and a continuum limit (lattice spacing becomes zeros) while holding the density finite.

The examples considered in this section are all rela-

tively simple, but serve to illustrate the basic procedure that can be applied to more complicated systems (such as the EGFR signaling model given in the main text in the next section).

### Reduction of a model of EGFR signaling

In this section we provide additional details about the reduction of the protein signaling network described in the main text.

The original model consists of 48 unknown parameters (primarily Michaelis-Menten and rate constants) and fifteen independent differential equations. The model simulates the response of extracellular stimulation of either the Epidermal Growth Factor (EGF) hormone or the Neuronal Growth Factor (NGF) hormone. Under the former, the ERK protein is initially activated and later suppressed, while when the latter is present, the activation persists. The model was proposed to explain these differing behaviors. By inspecting the figure in the main text, one can see three parallel pathways through which information about the stimulation may pass. The left most pathway ( $\text{EGFR} \rightarrow \text{PI3K} \rightarrow \text{Akt/PKB}$ ) was introduced in the model to explain the late time suppression of ERK under EGF stimulation. The subsequent analysis predicted that this pathway was not responsible for the observed behavior, which was confirmed experimentally[4, 5].

To reduce this model using the method described in the text, we choose to preserve the model predictions of several species under 8 distinct experimental conditions. These conditions correspond to either EGF or NGF stimulation for the unperturbed model as well as a perturbed model in which each of the three main pathways has been severed. These conditions were each simulated for two hours after the initial stimulation, following the original analysis, at which point the model has essentially reached a steady state. This choice of predictions corresponds to the same experimental conditions for which the model was originally calibrated and for which it was used to make predictions. Our reduced model, therefore, by construction is able to fit the same data and make the same predictions as the original.

We choose the numerical noise to have variance given by  $\sigma_i^2 = (0.1y_i)^2 + (0.1)^2$  where  $y_i$  is the simulated behavior of the model. This choice corresponds roughly to a ten percent error bar unless the fraction of protein phosphorylation is very small. If less than ten percent of a protein is activated, this choice allows the model to approximate the dynamics of the given protein as being practically zero. This choice is convenient for removing unnecessary dynamical variables from the model.

The differential equations for the original model are:

$$\frac{d}{dt} [\text{EGF}] = k_{ru}\text{EGF} \cdot [\text{bEGFR}] - k_{rb}\text{EGF} \cdot [\text{EGF}] \cdot [\text{fGFR}] \quad (31)$$

$$\frac{d}{dt} [\text{NGF}] = k_{ru}\text{NGF} \cdot [\text{bNGFR}] - k_{rb}\text{NGF} \cdot [\text{NGF}] \cdot [\text{fNGFR}] \quad (32)$$

$$\frac{d}{dt} [\text{fEGFR}] = k_{ru}\text{EGF} \cdot [\text{bEGFR}] - k_{rb}\text{EGF} \cdot [\text{EGF}] \cdot [\text{fEGFR}] \quad (33)$$

$$\frac{d}{dt} [\text{bEGFR}] = k_{rb}\text{EGF} \cdot [\text{EGF}] \cdot [\text{fEGFR}] - k_{ru}\text{EGF} \cdot [\text{bEGFR}] \quad (34)$$

$$\frac{d}{dt} [\text{fNGFR}] = k_{ru}\text{NGF} \cdot [\text{bNGFR}] - k_{rb}\text{NGF} \cdot [\text{NGF}] \cdot [\text{fNGFR}] \quad (35)$$

$$\frac{d}{dt} [\text{bNGFR}] = k_{rb}\text{NGF} \cdot [\text{NGF}] \cdot [\text{fNGFR}] - k_{ru}\text{NGF} \cdot [\text{bNGFR}] \quad (36)$$

$$\begin{aligned} \frac{d}{dt} [\text{SosA}] = & \frac{k_{\text{EGF}} \cdot [\text{bEGFR}] \cdot [\text{SosI}]}{([\text{SosI}] + K_{m\text{EGF}})} \\ & + \frac{k_{\text{NGF}} \cdot [\text{bNGFR}] \cdot [\text{SosI}]}{([\text{SosI}] + K_{m\text{NGF}})} \\ & - \frac{k_{d\text{Sos}} \cdot [\text{P90RskA}] \cdot [\text{SosA}]}{([\text{SosA}] + K_{md\text{Sos}})} \end{aligned} \quad (37)$$

$$\frac{d}{dt} [\text{P90RskA}] = \frac{k_{p\text{P90Rsk}} \cdot [\text{ErkA}] \cdot [\text{P90RskI}]}{([\text{P90RskI}] + K_{mp\text{P90Rsk}})} \quad (38)$$

$$\begin{aligned} \frac{d}{dt} [\text{RasA}] = & \frac{k_{\text{Sos}} \cdot [\text{SosA}] \cdot [\text{RasI}]}{([\text{RasI}] + K_{m\text{Sos}})} \\ & - \frac{k_{\text{RasGap}} \cdot [\text{RasGapA}] \cdot [\text{RasA}]}{([\text{RasA}] + K_{m\text{RasGap}})} \end{aligned} \quad (39)$$

$$\begin{aligned} \frac{d}{dt} [\text{Raf1A}] = & \frac{k_{\text{RasToRaf1}} \cdot [\text{RasA}] \cdot [\text{Raf1I}]}{([\text{Raf1I}] + K_{m\text{RasToRaf1}})} \\ & - \frac{k_{d\text{Raf1}} \cdot [\text{Raf1PPtase}] \cdot [\text{Raf1A}]}{([\text{Raf1A}] + K_{md\text{Raf1}})} \\ & - \frac{k_{d\text{Raf1ByAkt}} \cdot [\text{AktA}] \cdot [\text{Raf1A}]}{([\text{Raf1A}] + K_{m\text{Raf1ByAkt}})} \end{aligned} \quad (40)$$

$$\begin{aligned} \frac{d}{dt} [\text{BRafA}] = & \frac{k_{\text{Rap1ToBRaf}} \cdot [\text{Rap1A}] \cdot [\text{BRafI}]}{([\text{BRafI}] + K_{m\text{Rap1ToBRaf}})} \\ & - \frac{k_{d\text{BRaf}} \cdot [\text{Raf1PPtase}] \cdot [\text{BRafA}]}{([\text{BRafA}] + K_{md\text{BRaf}})} \end{aligned} \quad (41)$$

$$\begin{aligned} \frac{d}{dt} [\text{MekA}] = & \frac{k_{p\text{Raf1}} \cdot [\text{Raf1A}] \cdot [\text{MekI}]}{([\text{MekI}] + K_{mp\text{Raf1}})} \\ & + \frac{k_{p\text{BRaf}} \cdot [\text{BRafA}] \cdot [\text{MekI}]}{([\text{MekI}] + K_{mp\text{BRaf}})} \\ & - \frac{k_{d\text{Mek}} \cdot [\text{PP2AA}] \cdot [\text{MekA}]}{([\text{MekA}] + K_{md\text{Mek}})} \end{aligned} \quad (42)$$

$$\begin{aligned} \frac{d}{dt} [\text{ErkA}] = & \frac{k_{p\text{MekCytoplasmic}} \cdot [\text{MekA}] \cdot [\text{ErkI}]}{([\text{ErkI}] + K_{mp\text{MekCytoplasmic}})} \\ & - \frac{k_{d\text{Erk}} \cdot [\text{PP2AA}] \cdot [\text{ErkA}]}{([\text{ErkA}] + K_{md\text{Erk}})} \end{aligned} \quad (43)$$

$$\begin{aligned} \frac{d}{dt} [\text{PI3KA}] = & \frac{k_{\text{PI3K}} \cdot [\text{bEGFR}] \cdot [\text{PI3KI}]}{([\text{PI3KI}] + K_{m\text{PI3K}})} \\ & + \frac{k_{\text{PI3KRas}} \cdot [\text{RasA}] \cdot [\text{PI3KI}]}{([\text{PI3KI}] + K_{m\text{PI3KRas}})} \end{aligned} \quad (44)$$

$$\frac{d}{dt}[\text{AktA}] = \frac{k_{\text{Akt}} \cdot [\text{PI3KA}] \cdot [\text{AktI}]}{([\text{AktI}] + K_{\text{mAkt}})} \quad (45)$$

$$\frac{d}{dt}[\text{C3GA}] = \frac{k_{\text{C3GNGF}} \cdot [\text{bNGFR}] \cdot [\text{C3GI}]}{([\text{C3GI}] + K_{\text{mC3GNGF}})} \quad (46)$$

$$\begin{aligned} \frac{d}{dt}[\text{Rap1A}] &= \frac{k_{\text{C3G}} \cdot [\text{C3GA}] \cdot [\text{Rap1I}]}{([\text{Rap1I}] + K_{\text{mC3G}})} \\ &\quad - \frac{k_{\text{RapGap}} \cdot [\text{RapGapA}] \cdot [\text{Rap1A}]}{([\text{Rap1A}] + K_{\text{mRapGap}})} \end{aligned} \quad (47)$$

Most of the chemical species described by the model are superfluous to the understanding the model function. Experimental data is available for only a handful of species and one is typically only interested in the model predictions for a select few. We therefore reduce the model by preserving the time course for the experi-

mental data available in the original analysis. This corresponds to the activation levels of **Ras**, **Raf1**, **Mek**, **Rap1**, **BRaf**, and **Erk**. For these predictions, we reduced the model to include only six independent differential equations (the first two below are dependent) and twelve unknown parameters:

$$\frac{d}{dt}[\text{NGF}] = -\theta_1[\text{NGF}][\text{fNGFR}] \quad (48)$$

$$\frac{d}{dt}[\text{bNGFR}] = \theta_1[\text{NGF}][\text{fNGFR}] \quad (49)$$

$$[\text{bEGFR}] = \begin{cases} 1 & \text{EGF Present} \\ 0 & \text{Otherwise} \end{cases} \quad (50)$$

$$\frac{d}{dt}[\text{RasA}] = \theta_2[\text{bEGFR}] + \theta_3[\text{bNGFR}] - [\widetilde{\text{P90}}][\text{RasA}] \quad (51)$$

$$\frac{d}{dt}[\widetilde{\text{Raf1A}}] = \theta_4[\text{RasA}] - \theta_5 \frac{[\widetilde{\text{Raf1A}}]}{[\widetilde{\text{Raf1A}}] + \theta_6} \quad (52)$$

$$\frac{d}{dt}[\text{C3GA}] = \theta_7[\text{bNGFR}][\text{C3GI}] \quad (53)$$

$$[\text{Rap1A}] = \theta_8[\text{C3GA}] \quad (54)$$

$$[\text{MekA}] = [\widetilde{\text{Raf1A}}][\text{MekI}] + \theta_9[\text{Rap1A}] \quad (55)$$

$$\frac{d}{dt}[\text{ErkA}] = -\theta_{10}[\text{ErkA}] + \theta_{11}[\text{MekA}][\text{ErkI}] \quad (56)$$

$$\frac{d}{dt}[\widetilde{\text{P90}}] = \theta_{12}[\text{ErkA}]. \quad (57)$$

The regulatory relationships described by these equations are represented schematically in Figure 3 in the main text. The parameters' relations to the originals that are

inherited through the limiting approximations are given by:

$$\theta_1 = k_{\text{rbNGF}} \quad (58)$$

$$\theta_2 = \frac{[\text{SosI}](k_{\text{EGF}})(K_{\text{mRasGap}})(k_{\text{Sos}})}{[\text{RasGapA}](K_{\text{mEGF}})(k_{\text{RasGap}})} \quad (59)$$

$$\theta_3 = \frac{[\text{SosI}](k_{\text{NGF}})(K_{\text{mRasGap}})(k_{\text{Sos}})}{[\text{RasGapA}](K_{\text{mNGF}})(k_{\text{RasGap}})} \quad (60)$$

$$\theta_4 = \frac{(k_{\text{RasToRaf1}})(k_{\text{pRaf1}})(K_{\text{mdMek}})}{[\text{PP2AA}](k_{\text{dMek}})(K_{\text{mpRaf1}})} \quad (61)$$

$$\theta_5 = \frac{[\text{Raf1PPtase}](\text{kdRaf1})(\text{kpRaf1})(\text{KmdMek})}{[\text{PP2AA}](\text{kdMek})(\text{KmpRaf1})} \quad (62)$$

$$\theta_6 = \frac{(\text{KmdRaf1})(\text{kpRaf1})(\text{KmdMek})}{[\text{PP2AA}](\text{kdMek})(\text{KmpRaf1})} \quad (63)$$

$$\theta_7 = \text{kC3GNGF}/\text{KmC3GNGF} \quad (64)$$

$$\theta_8 = \frac{(\text{KmRapGap})(\text{kC3G})}{[\text{RapGapA}](\text{kRapGaP})} \quad (65)$$

$$\theta_9 = \frac{[\text{BRaf1}](\text{kRap1ToBRaf})(\text{KmdBRaf})(\text{kpBRaf})(\text{KmdMek})}{[\text{PP2AA}][\text{Raf1PPtase}](\text{kdBRaf})(\text{KmRap1ToBRaf})(\text{kdMek})} \quad (66)$$

$$\theta_{10} = [\text{PP2AA}](\text{kdErk})/\text{KmdErk} \quad (67)$$

$$\theta_{11} = \text{kpMekCyttoplasmic}/\text{KmpMekCyttoplasmic} \quad (68)$$

$$\theta_{12} = \frac{[\text{P90/RSKI}](\text{kpP90Rsk})(\text{kdSos})}{(\text{KmpP90Rsk})(\text{KmdSos})}. \quad (69)$$

parameter	Reduced Model Value	Uncertainty
$\theta_1$	$2.37 \times 10^{-3}$	$1.1 \times 10^{-3}$
$\theta_2$	$9.34 \times 10^{-2}$	$2.7 \times 10^{-2}$
$\theta_3$	$7.57 \times 10^{-1}$	$3.9 \times 10^{-1}$
$\theta_4$	$9.88 \times 10^{-1}$	$5.4 \times 10^{-1}$
$\theta_5$	$3.40 \times 10^{-1}$	$2.1 \times 10^{-1}$
$\theta_6$	$2.70 \times 10^{-1}$	$1.9 \times 10^{-1}$
$\theta_7$	$1.11 \times 10^0$	$7.8 \times 10^{-1}$
$\theta_8$	$1.30 \times 10^{-1}$	$3.6 \times 10^{-2}$
$\theta_9$	$1.75 \times 10^0$	$5.5 \times 10^{-1}$
$\theta_{10}$	$2.56 \times 10^{-1}$	$1.0 \times 10^{-1}$
$\theta_{11}$	$4.51 \times 10^0$	$1.6 \times 10^0$
$\theta_{12}$	$8.21 \times 10^{-1}$	$3.5 \times 10^{-1}$

TABLE I: Parameter values for the reduced model of 12 parameters. As the model is simplified, the uncertainty in the inferred parameter values becomes much smaller than in the original model. Here, we report the one standard-deviation uncertainty estimate from the inverse Fisher Information.

The parameter values for this reduced model are given in Table I. As the model is reduced, the uncertainty in the inferred parameters becomes much smaller than in the original analysis.

### One-dimensional Ising Model

In the main text we considered a one-dimensional Ising model with Hamiltonian given by

$$H = - \sum_{\mu}^N J_{\mu} s_{\mu} s_{\mu+1} \quad (70)$$

where  $J_{\mu}$  are the parameters that represent the nearest neighbor coupling strengths of the spins  $s_{\mu}$ . Since the spin chain is one dimensional, it can be solved by introducing the random variables  $\xi_{\mu} = s_{\mu} s_{\mu+1}$ . One can then

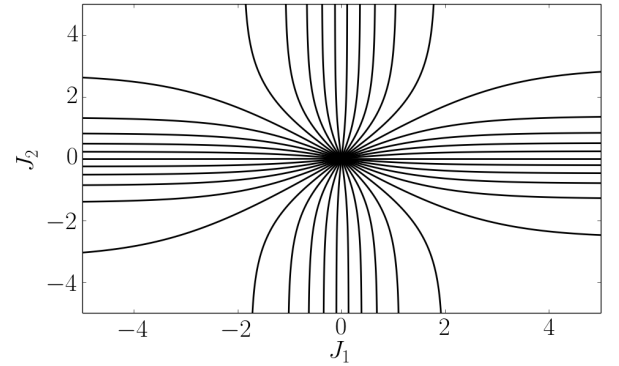


FIG. 7: **Geodesics for the Ising model in Eq (70).** Geodesics in the Ising model with metric tensor given by Eq. (71) for paths originating from the origin with initial directions along radial lines. The geodesics identify four possible limiting cases  $J_{\mu} \rightarrow \pm\infty$ .

show that these random variables are independently distributed with expectation value  $\langle \xi_{\mu} \rangle = \tanh J_{\mu}$ . Note that since  $\xi_{\mu}^2 = 1$ , the distribution of  $\xi_{\mu}$  is completely determined by specifying its first moment which is directly controlled by the parameter  $J_{\mu}$ . For clarity, we consider here the model with three spins and two parameters, but basic the results generalize to more complicated scenarios. The metric for this model is given by

$$g = \begin{pmatrix} \frac{4e^{2J_1}}{(1+e^{2J_1})^2} & 0 \\ 0 & \frac{4e^{2J_2}}{(1+e^{2J_2})^2} \end{pmatrix}. \quad (71)$$

And geodesics in the  $J_1, J_2$  plane are illustrated in Figure 7. These geodesics identify four limiting boundaries:  $J_{\mu} \rightarrow \pm\infty$  for  $\mu = 1, 2$ .

Now consider the coarse-graining scenario described in the main text. If the distribution given by Eq. (70) is marginalized over of the spins  $s_2$ , the distribution is over-parameterized since it has more parameters than in-



dependent moments in the distribution. Consequently, the Fisher Information becomes singular and we can take  $J_\mu \rightarrow \pm\infty$  without losing any predictive power in the model. In the reduced model, the single remaining parameter  $\tilde{J}$  determines the expectation value of  $s_1 s_3 = \xi_1 \xi_2$ . We can therefore express the parameter of the reduced models in terms of the originals:  $\tanh \tilde{J} = \tanh J_1 \tanh J_2$ . From this it follows that  $|\tilde{J}| < |J_1| |J_2|$ . In a larger system with many parameters and spins, after many similar coarse-graining iterations,  $\tilde{J} \rightarrow 0$ , consistent with the standard results obtained from RG for the one-dimensional Ising model.

Now consider an alternate generalization of the one-dimensional Ising model with Hamiltonian given by

$$H = - \sum_{\mu\alpha} J_\alpha s_\mu s_{\mu+\alpha}, \quad (72)$$

where the parameter  $J_1$  represents the nearest neighbor coupling,  $J_2$  the second nearest neighbor coupling, and so forth. Again for clarity, we consider the model with two parameters and four spins. The metric, more complicated than before, is given by,

$$g = \begin{pmatrix} \frac{32e^{4J_1+8J_2}(1+e^{8J_1}+2e^{4J_2}+2e^{8J_1+4J_2}+2e^{4J_1+8J_2})}{(2e^{4J_1}+e^{8J_2}+4e^{4(J_1+J_2)}+e^{8(J_1+J_2)})^2} & \frac{64e^{4J_1+8J_2}(-1+e^{8J_1})(1+e^{4J_2})}{(2e^{4J_1}+e^{8J_2}+4e^{4(J_1+J_2)}+e^{8(J_1+J_2)})^2} \\ \frac{64e^{4J_1+8J_2}(-1+e^{8J_1})(1+e^{4J_2})}{(2e^{4J_1}+e^{8J_2}+4e^{4(J_1+J_2)}+e^{8(J_1+J_2)})^2} & \frac{64e^{4(J_1+J_2)}(2e^{4J_1}+2e^{4J_2}+e^{8J_2}+e^{8(J_1+J_2)}+2e^{8J_1+4J_2})}{(2e^{4J_1}+e^{8J_2}+4e^{4(J_1+J_2)}+e^{8(J_1+J_2)})^2} \end{pmatrix}, \quad (73)$$

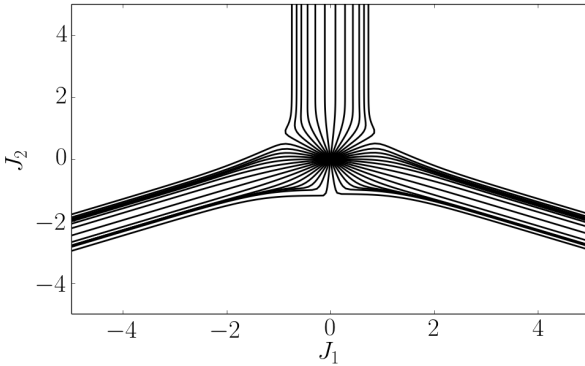


FIG. 8: **Geodesics for the Ising model in Eq (72).** Geodesics in the Ising model with metric tensor given by Eq. (73) for paths originating from the origin with initial directions along radial lines. The geodesics indicate three possible limiting cases in which the parameters become infinite in various combinations.

and geodesics in the  $J_1, J_2$  plane are given in Figure 8. In this case, there are three limiting boundaries. First  $J_1 \rightarrow \infty$  and  $J_2 \rightarrow -\infty$  (toward the lower right in Figure 8). Second,  $J_1 \rightarrow -\infty$  and  $J_2 \rightarrow -\infty$  (toward the lower left in Figure 8). Third,  $J_2 \rightarrow \infty$  while  $J_1$  remains finite.

Here we focus on the first limit  $J_1 \rightarrow \infty$  and  $J_2 \rightarrow -\infty$ . By inspecting the distribution along geodesic paths approaching this boundary, we find that among the 16 distinct spin configurations in the model, all but 6 occur

with zero probability at this boundary. To understand the relationship among these 6 remaining states, it is convenient to rewrite the spins as a Fourier transform:

$$\tilde{s}_\mu = \sum_{\nu=1}^4 s_\nu e^{2\pi i(\nu-1)\mu/4}, \quad (74)$$

for  $\mu = 0, \pm 1, 2$ . After this transformation the Hamiltonian in Eq. (72) becomes

$$H = -\frac{J_1}{4} (\tilde{s}_0^2 - \tilde{s}_2^2) - \frac{J_2}{4} (\tilde{s}_0^2 - 2\tilde{s}_1\tilde{s}_{-1} + \tilde{s}_2^2). \quad (75)$$

In this notation, we find that the 6 remaining states on the boundary are those for which  $\tilde{s}_2 = 0$  (i.e. those for which the highest frequency mode has no contribution to the energy).

The ratio of Boltzmann factors for two of the remaining states is given by

$$P_1/P_2 = e^{4J_1+8J_2}, \quad (76)$$

indicating that although  $J_1 \rightarrow \infty$  and  $J_2 \rightarrow -\infty$ , the combination

$$\tilde{J} = J_1 + 2J_2, \quad (77)$$

must remain finite, which combination we identify as our reduced parameter. Using Eq. (77) to remove  $J_2$  from the Hamiltonian we find

$$H = -\frac{\tilde{J}}{8} (\tilde{s}_0^2 - 2\tilde{s}_1\tilde{s}_{-1} + \tilde{s}_2^2) - J_1 (\tilde{s}_0^2 + 2\tilde{s}_1\tilde{s}_{-1} - 3\tilde{s}_2^2) \quad (78)$$

$$= -\frac{\tilde{J}}{8} (\tilde{s}_0^2 - 2\tilde{s}_1\tilde{s}_{-1}) - J_1 (\tilde{s}_0^2 + 2\tilde{s}_1\tilde{s}_{-1}), \quad (79)$$

where in the second line we have used the fact that  $\tilde{s}_2 = 0$  for the coarse-grained states.

Although Eq. (79) appears to be a two parameter model, it can be checked that  $\tilde{s}_0^2 + 2\tilde{s}_1\tilde{s}_{-1}$  takes the same value for each of the 6 states of interest, so the  $J_1$  term simply adds an irrelevant constant to the Hamiltonian and can be ignored. The final, reduced model is given by

$$H = -\frac{\tilde{J}}{8} (\tilde{s}_0^2 - 2\tilde{s}_1\tilde{s}_{-1}). \quad (80)$$

Thus, this particular boundary corresponds to a reduced model in which the high frequency contributions to the energy have been removed.

The other two boundaries illustrated in Figure 8 can be analyzed in a similar way. The lower left boundary ( $J_1, J_2 \rightarrow -\infty$ ) correspond to a limit in which the

low frequency contributions are removed from the model, i.e.  $\tilde{s}_0 = 0$ . The limit of  $J_2 \rightarrow \infty$  corresponds to the scenario in which the intermediate frequency modes are removed, i.e.  $\tilde{s}_{\pm 1} = 0$ .

- 
- [1] M. K. Transtrum, B. B. Machta, J. P. Sethna: *Physical Review Letters* **104** (2010) 1060201
  - [2] M. K. Transtrum, B. B. Machta, J. P. Sethna: *Physical Review E* **83** (2011) 036701
  - [3] B. Averick, R. Carter, J. Moré (1991)
  - [4] K. Brown, J. Sethna: *Physical Review E* **68** (2003) 21904
  - [5] K. Brown, C. Hill, G. Calero, C. Myers, K. Lee, J. Sethna, R. Cerione: *Physical biology* **1** (2004) 184