Supplementary Information:
Exploiting network topology for large-scale inference of nonlinear reaction models
Journal of the Royal Society Interface

Nikhil Galagali and Youssef M. Marzouk
Massachusetts Institute of Technology
Cambridge, MA 02139 USA

1 Network analysis for improved sampling efficiency

1.1 Centered Gaussian parameter proposals

In the main paper, we describe the network-unaware approach that does not recognize the effective networks during the between-model moves. We provide further details here. For nested models, as is the case in the reaction network inference problem, a natural choice of the jump function $f$ is to choose the identity function. Thus, when proposing a move from a lower-dimensional model $M_m$ to a higher-dimensional model $M_n$, the rate constants of the newly added reactions are proposed according to $q(u|k_m)$ and the values of the rate constants of reactions common to the two models are kept fixed. Suppose that model $M_m$ has $i$ reactions and that model $M_n$ has $a > i$ reactions, with the first $i$ reactions common; then an identity $f$ mapping $(k_m,1:i,u)$ to $(k_n,1:i,k_n,i+1:a)$ is simply

\[ k_{n,1:i} = k_{m,1:i}, \quad k_{n,i+1:a} = u \]  

and the acceptance probability is given by

\[ \alpha(k_m,k_n) = \min \left\{ 1, \frac{p(M_n,k_n|\mathcal{D})q(M_m|M_n)}{p(M_m,k_m|\mathcal{D})q(M_n|M_m)q(u|k_m)} \right\}. \]  

The reverse move in this case is deterministic. Let the proposal $q(u|k_m)$ be given by

\[ q(u|k_m) = \mathcal{N}(u; \mu, \Sigma). \]
Gaussian approximation to the conditional posterior distribution. In the framework of Brooks et al., the above construction is equivalent to the centered second-order conditions \[ ? \]. In the scheme described above, the mean vector \( \mu \) is set to the conditional maximum:

\[
\mu = \arg \max_{k_{n,i+1:a}} p(k_{n,i+1:a}|k_{m,1:i}, M_n, \mathcal{D}).
\] (4)

A proposal centered at the posterior conditional maximum satisfies the first order condition:

\[
\nabla_u \log A(M_m, k_m \rightarrow M_n, k_n)|_{\mu} = \nabla \log \mathcal{L}(\mathcal{D}; k_{m,1:i}, u) + \log p(k_{m,1:i}, u) \\
- \log \mathcal{N}(u; \mu, \Sigma)|_{u=\mu} = 0.
\] (5)

Here, \( \mathcal{L}(\mathcal{D}; k_{m,1:i}, u) \) refers to the likelihood function of the models and \( p(k_{m,1:i}, u) \) to the prior probability distribution of the rate constants. Further, setting the second-derivative of the acceptance ratio at the conditional maximum to be 0, we obtain the second order condition as:

\[
\nabla^2_u \log A(M_m, k_m \rightarrow M_n, k_n)|_{\mu} = \nabla^2 \log \mathcal{L}(\mathcal{D}; k_{m,1:i}, u) + \log p(k_{m,1:i}, u) \\
- \log \mathcal{N}(u; \mu, \Sigma)|_{u=\mu} = 0.
\] (6)

Taking \( \mathcal{H} \) to be the Hessian of the conditional posterior density \( p(k_{n,i+1:a}|k_{m,1:i}, M_n, \mathcal{D}) \) at \( \mu \), (6) yields

\[
\mathcal{H} + \Sigma^{-1} = 0 \implies \Sigma = -\mathcal{H}^{-1}.
\] (7)

### 1.2 Network-aware parameter proposals

As discussed in the paper, many reaction networks can have the same effective network. In such a case, if the proposed move is between two networks with the same effective network (i.e., the two networks belong to the same cluster), the parameter proposal adapts to the prior distribution of the newly added reaction. We propose a network-aware approach in which, because we have determined the effective networks, we design parameter proposals that adapt to the difference between the effective networks of the two networks. When the proposed move is between two networks belonging to different clusters, we construct a proposal that approximates the conditional posterior distribution of the rate constants of all reactions not included in the two effective networks. Formally, suppose that the sampler proposes a move from a lower-dimensional model \( M_m \) to a higher-dimensional model \( M_n \). Let the effective networks of the two models \( M_m \) and \( M_n \) be \( M_{me} \) and \( M_{ne} \), respectively. Suppose the proposed move is such that \( M_{ne} \neq M_{me} \), i.e., the effective networks of the current and the proposed networks are different. Further, suppose that model \( M_{me} \) has \( i \) reactions and that model \( M_{ne} \) has \( a > i \) reactions, with
the first $i$ reactions common; then an identity $f$ mapping $(k_{me,1:i}, u)$ to $(k_{ne,1:i}, k_{ne,i+1:a})$ is simply

$$(k_{ne,1:i}, k_{ne,i+1:a}) = (k_{me,1:i}, u),$$

where $u \sim \mathcal{N}(\mu_{me}, \Sigma_{me})$. The proposal mean

$$\mu_{me} = \arg \max_{k_{ne,i+1:a}} p(k_{ne,i+1:a} | k_{me,1:i}, M_{ne}, \mathcal{D})$$

is obtained by solving an $(a - i)$-dimensional optimization problem. The proposal covariance

$$\Sigma_{me} = -\left[ \nabla^2 \log p(k_{ne,i+1:a} | k_{me,1:i}, M_{ne}, \mathcal{D}) \right]^{-1}\bigg|_{k_{ne,i+1:a} = \mu_{me}}$$

is determined numerically using a finite-difference approximation at the proposal mean. The acceptance probability of the proposed move is given by

$$\alpha((M_{me}, k_{me}), (M_{ne}, k_{ne})) = \min \{1, A\},$$

where

$$A = \frac{p(M_{ne}, k_{ne} | \mathcal{D}) q(M_{m} | M_{n})}{p(M_{me}, k_{me} | \mathcal{D}) q(M_{n} | M_{m}) \mathcal{N}(u; \mu_{me}, \Sigma_{me})}.$$  (12)

The reverse move is deterministic and has an acceptance probability $\min\{1, A^{-1}\}$. The idea behind the construction of our network aware proposals is that by solving for the maximum of the joint conditional posterior density of reactions that are in $M_{ne}$ but not in $M_{me}$, and determining the Hessian approximation at that point, we are building a Gaussian approximation of the conditional probability density $p(k_{ne,i+1:a} | k_{me,1:i}, M_{ne}, \mathcal{D})$. In contrast, the standard network-unaware approach would not; in particular, it produces a proposal that is the product of prior densities for $\text{dim}(M_{ne}) - \text{dim}(M_{me}) - 1$ rate constants of the reactions that do not change the effective network and the conditional posterior density of the rate constant of the final reaction that ultimately leads to the change in effective network from $M_{me}$ to $M_{ne}$.

| Method           | Proposal                                                                 |
|------------------|--------------------------------------------------------------------------|
| Network unaware  | $q_{nu}(u | k_{me}) \approx \prod_{j=i+1}^{a-1} p(k_{ne,j}) p(k_{ne,a} | k_{ne,1:a-1}, M_{ne}, \mathcal{D})$ |
| Network aware    | $q_{na}(u | k_{me}) \approx p(k_{ne} | k_{me}, M_{ne}, \mathcal{D})$         |

Table 1: Cluster switching parameter proposals

Mathematically, the two proposals are shown in Table 1. In moves where the effective networks of the current and the proposed network are the same, both the network-unaware and our network-aware approaches use the prior distribution of the newly added reaction as the proposal and have acceptance probabilities of 1. The steps of our network-aware reversible jump MCMC algorithm are given in Algorithm 1.
Algorithm 1 Network-aware reversible jump MCMC

1: **Given:** A set of models $M_m \in \mathcal{M}$ with corresponding parameter vectors $k_m$, posterior densities $p(M_m, k_m | D)$,
2: $\beta \in (0, 1)$: probability of within-model move
3: Initialize chain: $\{(M^0, k_{M^0}), M^0_e \leftarrow \text{effective network of } M^0\}$
4: for $s = 0$ to $N_{\text{iter}}$ do
5: Sample $b \sim \mathcal{U}_{[0,1]}$
6: if $b \leq \beta$ then
7: Metropolis-Hastings within-model move
8: else
9: Sample $M' \sim q(M'|M^s); M'_e \leftarrow \text{effective network of } M'$
10: if $|M'_e| > |M^s_e|$ then
11: $\alpha(M^s, k_{M^s}), (M', k_{M'}) = \min \left\{1, \frac{p(M'_e, k_{M'_e} | D)q(M'|M')}{p(M^s_e, k_{M^s} | D)q(M^s|M^s)} \right\}$
12: if $|M'_e| < |M^s_e|$ then
13: $\alpha(M^s, k_{M^s}), (M', k_{M'}) = \min \left\{1, \frac{p(M'_e, k_{M'_e} | D)q(M'|M')}{p(M^s_e, k_{M^s} | D)q(M^s|M^s)} \right\}$
14: else
15: Sample $k_{M'} \setminus k_{M^s} \sim p(k_{M'} \setminus k_{M^s})$
16: $\alpha((M^s, k_{M^s}), (M', k_{M'})) = 1$
17: end if
18: Sample $p \sim \mathcal{U}_{[0,1]}$
19: if $p < \alpha((M^s, k_{M^s}), (M', k_{M'}))$ then
20: $\{(M^{s+1}, k_{M^{s+1}}^{s+1}), M^{s+1}_e \} = \{(M', k_{M'}), M'_e\}$
21: else
22: $\{(M^{s+1}, k_{M^{s+1}}^{s+1}), M^{s+1}_e \} = \{(M^s, k_{M^s}^s), M^s_e\}$
23: end if
24: end for
1.3 Sensitivity-based network-aware proposals

In the paper, we presented the sensitivity-based network aware algorithm and explained its construction. Here we provide technical details and a pseudocode for the application of the sampler. Suppose, the sampler move is proposed between a lower-dimensional model \(M_m\) to a higher-dimensional model \(M_n\). Let the effective networks of the two models \(M_m\) and \(M_n\) be \(M_{me}\) and \(M_{ne}\), respectively. Suppose the proposed move is such that \(M_{me} \neq M_{ne}\), i.e., the effective networks of the current and the proposed networks are different. The identity mapping \(f\) between \((k_{me}, u_{me})\), the rate constants and proposal parameters of network \(M_{me}\) to \((k_{ne}, u_{ne})\), the rate constants and proposal parameters of network \(M_{ne}\) for sensitivity-based move proposals is given by

\[
\begin{align*}
&k_{ne,1:i} = k_{me,1:i}, \
&k_{ne,i+1:i+c} = u_{me,1:c}, \
&u_{ne,1:c} = k_{me,i+1:i+c}, \text{ and} \
&k_{ne,i+c+1:i+a} = u_{me,c+1:a}.
\end{align*}
\]

(13)

Here, \(\{1 : i\}\) are indices of reactions that are common to the two networks and whose rate constant values are kept fixed during moves between networks \(M_{me}\) and \(M_{ne}\), \(\{i + c + 1 : i + a\}\) are indices of reactions that are in network \(M_{ne}\) but not in \(M_{me}\), and \(\{i + 1 : i + c\}\) are the critical reactions that are present in both networks, but whose values are proposed according to a proposal distribution rather than keeping them fixed during the between-model move. The proposal distributions are again taken to be Gaussian \((u_{me} \sim \mathcal{N}(\mu_{me}, \Sigma_{me})\) and \(u_{ne} \sim \mathcal{N}(\mu_{ne}, \Sigma_{ne})\)). The proposal means

\[
\mu_{me} = \arg \max_{k_{me,i+1:i+a}} p(k_{me,i+1:i+a}|k_{me,1:i}, M_{ne}, D) \tag{14}
\]

and

\[
\mu_{ne} = \arg \max_{k_{me,i+1:i+c}} p(k_{me,i+1:i+c}|k_{me,1:i}, M_{me}, D) \tag{15}
\]

are obtained by solving \(a\)-dimensional and \(c\)-dimensional optimization problems. The proposal covariances

\[
\Sigma_{mc} = -[\nabla^2 \log p(k_{mc,i+1:i+a}|k_{mc,1:i}, M_{nc}, D)]^{-1} |_{\mu_{mc}} \tag{16}
\]

and

\[
\Sigma_{nc} = -[\nabla^2 \log p(k_{mc,i+1:i+c}|k_{mc,1:i}, M_{mc}, D)]^{-1} |_{\mu_{mc}} \tag{17}
\]

are determined numerically using a finite-difference approximation at the proposal means. The acceptance probability of the proposed move is given by

\[
\alpha((M_{me}, k_{mc}); (M_{ne}, k_{nc})) = \min \{1, A\}, \tag{18}
\]

where

\[
A = \frac{p(M_{ne}, k_{nc}|D)q(M_{m}|M_{n})\mathcal{N}(u_{nc}; \mu_{nc}, \Sigma_{nc})}{p(M_{me}, k_{me}|D)q(M_{n}|M_{m})\mathcal{N}(u_{me}; \mu_{me}, \Sigma_{mc})}. \tag{19}
\]
The reverse move is nondeterministic and has an acceptance probability \( \min\{1, A^{-1}\} \).

The pseudocode for sensitivity-based network aware algorithm is given in Algorithm 2.

1.4 Derandomization of conditional expectations

The above Algorithms 1 and 2 lead to gains in sampling efficiency compared to a reversible jump MCMC algorithm that does not use information on network structure in designing between-model moves and parameter proposals. Identifying clusters of models can be further used for additional variance reduction. With the knowledge that all models belonging to the same cluster have identical model evidence, we can compute some expectations analytically and thereby obtain posterior averages of features with lower variances.

1.4.1 General formulation

Let us assume we are performing model inference with \( F \) as one the quantities of interest. Generally, we may be interested in quantities such as the posterior model probabilities, reaction inclusion probabilities of reactions, or pathway probabilities. The Monte Carlo estimate of \( F \) from posterior samples can be written as:

\[
\hat{F} = p(F = 1|D) = \int p(F = 1|C)p(C|D)dC = \int p(F = 1|M)p(M|C)p(C|D)dMdC = \int E_{p(M|C)}[p(F = 1|M)]p(C|D)dC = \frac{1}{N_s} \sum_{i=1}^{N_s} E_{p(M|C^i)}[p(F = 1|M)],
\]

where \( C \) refers to model clusters, \( N_s \) is the number of posterior samples and \( D \) the available data. In the above equation, \( E_{p(M|C^i)}[p(F = 1|M)] \) is the expected value of \( p(F = 1|M) \) conditioned on the generated sample \( C^i \). Knowing the cluster to which each sample belongs and the dependence of the feature on the models included in the cluster, the above expectation can be computed analytically and allows variance reduction. In contrast, in the network-unaware approach, the expectation is computed through Monte Carlo sampling.

1.4.2 Example: model probability estimates

Consider that the feature of interest is the probability of model \( m \). Thus, applying the above formula to the estimation of model probability, we get
Algorithm 2: Sensitivity-based network-aware reversible jump MCMC

1: Given: A set of models $M \in \mathcal{M}$ with corresponding parameter vectors $k_M$, posterior densities $p(M, k_M | \mathcal{D})$.
2: $\beta \in (0,1)$: probability of within-model move
3: Initialize chain: $\{(M^0, k_{M^0}), M^0_e \leftarrow \text{effective network of } M^0\}$
4: for $s = 0$ to $N_{iter}$ do
5:     Sample $b \sim U_{[0,1]}$
6:     if $b \leq \beta$ then
7:         Metropolis-Hastings within-model move
8:     else
9:         Sample $M' \sim q(M'|M^s)$; $M'_e \leftarrow \text{effective network of } M'$
10:        $r_1 \sim \text{Poisson}(1.5)$ and $r_2 \sim \text{Poisson}(1.5)$
11:        $\{1: i\} =$ reactions that are common to $M^s_e$ and $M'_e$ and whose rate constant
12:        values are kept fixed during moves
13:        $\{i + 1: i + c\} =$ reactions common to $M^s_e$ and $M'_e$ and with top $r_1$ and $r_2$
14:        sensitivities of $M^s_e$ and $M'_e$, respectively
15:        $\{i + c + 1: i + a\} =$ reactions only present in $M^s_e$ or $M'_e$
16:     if $|M'_e| > |M^s_e|$ then
17:         $\mu_M = \arg \max p(k^{i+1:i+c}_{M^s_e} | k^{1:i}_{M^s_e}, M^s_e, \mathcal{D})$, $\Sigma_M = -[\nabla^2 \log p(k^{i+1:i+c}_{M^s_e} | k^{1:i}_{M^s_e}, M^s_e, \mathcal{D})]^{-1}$
18:         $\mu_M' = \arg \max p(k^{i+1:i+c}_{M'_e} | k^{1:i}_{M'_e}, M'_e, \mathcal{D})$, $\Sigma_M' = -[\nabla^2 \log p(k^{i+1:i+c}_{M'_e} | k^{1:i}_{M'_e}, M'_e, \mathcal{D})]^{-1}$
19:         Sample $u \sim \mathcal{N}(\mu_M, \Sigma_M)$
20:         Set $(k^{1:i}_{M^s_e}, k^{1:i+1:c}_{M^s_e}, u^{1:c}, k^{i+1:i+c}_{M^s_e}, u^{c+1:a}) = (k^{1:i}_{M'_e}, k^{i+1:i+c}_{M'_e}, u^{1:c}, k^{1:i+1:c}_{M'_e}, u^{c+1:a})$
21:     else if $|M'_e| < |M^s_e|$ then
22:         $\mu_M = \arg \max p(k^{1:i+1:i+c}_{M^s_e} | k^{i}_{M^s_e}, M^s_e, \mathcal{D})$, $\Sigma_M = -[\nabla^2 \log p(k^{1:i+1:i+c}_{M^s_e} | k^{i}_{M^s_e}, M^s_e, \mathcal{D})]^{-1}$
23:         $\mu_M' = \arg \max p(k^{1:i+1:i+c}_{M'^s_e} | k^{i}_{M'^s_e}, M'_e, \mathcal{D})$, $\Sigma_M' = -[\nabla^2 \log p(k^{1:i+1:i+c}_{M'^s_e} | k^{i}_{M'^s_e}, M'_e, \mathcal{D})]^{-1}$
24:         Sample $u \sim \mathcal{N}(\mu_M, \Sigma_M)$
25:         Set $(k^{1:i}_{M^s_e}, k^{1:i+1:i+c}_{M^s_e}, u^{1:c}, k^{i+1:i+c}_{M^s_e}, u^{c+1:a}) = (k^{1:i}_{M'_e}, k^{i+1:i+c}_{M'_e}, u^{1:c}, k^{i+1:i+c}_{M'_e}, u^{c+1:a})$
26:     else
27:         Sample $k_{M'} \backslash k_{M^s} \sim p(k_{M'} | k_{M^s})$
28:     Sample $p \sim U_{[0,1]}$
29:     if $p < \min \left\{ 1 \cdot \frac{p(M'_e, k'_M | \mathcal{D}) q(M'|M^s) \mathcal{N}(u'|\mu_M', \Sigma_M')} {p(M^s_e, k_M | \mathcal{D}) q(M^s | M^s) \mathcal{N}(u|\mu_M, \Sigma_M)} \right\}$ then
30:         $\{(M^{s+1}, k^{s+1}_{M^{s+1}}), M^{s+1}_e\} = \{(M', k_{M'}), M'_e\}$
31:     else
32:         $\{(M^{s+1}, k^{s+1}_{M^{s+1}}), M^{s+1}_e\} = \{(M^s, k_{M^s}), M^s_e\}$
$$\hat{M}_m = p(M_m = 1|D)$$

$$= \int p(M_m = 1|C)p(C|D)dC$$

$$= \int p(M_m = 1|M)p(M|C)p(C|D)dMdC$$

$$= \int E_{p(M|C)}[p(M_m = 1|M)]p(C|D)dC$$

$$= \frac{1}{N_s} \sum_{i=1}^{N_s} E_{p(M|C_i)} [p(M_m = 1|M)]$$

$$= \frac{1}{N_s} \sum_{i=1}^{N_s} p(M_m|C_K)\mathbb{1}_{C_K}(C^i),$$  \hspace{1cm} (20)

where $K : \mathbb{1}_{M_m \in C_K}(M_m) = 1$ and $\mathbb{1}$ is the indicator function. In our network aware schemes, $p(M_m|C_K)$ can be computed analytically. For example, for a cluster $C_K$ with $N_K$ models, taking the prior distribution over models to be uniform, the model probability estimate is

$$\hat{M}_m = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{1}{N_K} \mathbb{1}_{C_K}(C^i)$$ \hspace{1cm} (21)

In contrast, with a standard reversible-jump algorithm, the model probability estimate is

$$\hat{M}_m = \frac{1}{N_s} \sum_{i=1}^{N_s} \mathbb{1}_{M_m}(M^i) \mathbb{1}_{C_K}(C^i)$$ \hspace{1cm} (22)

2 Results

In our paper, we present two example problems and demonstrate the efficiency of our network-aware sampling approaches compared to the network-unaware approach. The observables in our examples are species concentrations and the concentration evolution is modeled using the law of mass action/Michaelis-Menten functionals. The law of mass action gives the rate of a chemical reaction (say $X + Y \rightarrow Z$) as the product of a reaction-specific rate constant $k$ with reactant concentrations $[X]$ and $[Y]$:

$$\text{Rate} = -k[X][Y].$$ \hspace{1cm} (23)

Under some assumptions, the law of mass action produces Michaelis-Menten reaction rate expression
Figure 1: Reaction network of Example 1 and Example 2
\[ Rate = \frac{k[S]}{k_M + [S]}, \]  
\[ Rate = k[E]_0 \frac{[S]}{k_M + [S]}, \]

or when enzyme concentration is taken into account [?]:

where \( k \) denotes the rate constant, \([E]_0\) is the enzyme concentration, \([S]\) the substrate concentration, and \( k_M \) the Michaelis constant.

2.1 Example 1: five-dimensional nonlinear network inference

In the first example, we consider a subset of reactions (15 species and 12 reactions) proposed for a protein-signalling network of the activation of extracellular signal-regulated kinase (ERK) by epidermal growth factor (EGF) [?]. The reaction network used is shown in Figure [1]. The ODE forward model governing the evolution of species concentrations is described in detail in Section [3]. We keep reactions 1, 2, 8, 9, 10, 11, and 12 fixed and thus they are included in all the inferred models. The rate constants of all fixed reactions (shaded pink in Table 2) and Michaelis constants of all reactions are set to their base values (Table 2).

| Reaction | \( \log_{10} k^a \) | \( k^b_M \) | Prior uncertainty |
|----------|-----------------|-----------------|------------------|
| 1 | BEGFR → DEGFR | 0.0 | - | - |
| 2a | EGF + UEGFR → BEGFR | 1.5 | - | - |
| 2b | BEGFR → EGF + UEGFR | 0.0 | - | - |
| 3 | inactiveC3G+BEGFR → activeC3G+BEGFR | 0.5 | 3386.3875 | \( \log_{10} \) \( k \) = \( N \) (1.1, 0.2) |
| 4 | activeC3G → inactiveC3G | 2.0 | - | \( \log_{10} \) \( k \) = \( N \) (1.4, 0.2) |
| 5 | inactiveRap1+activeC3G → activeRap1+activeC3G | 2.0 | 3566 | \( \log_{10} \) \( k \) = \( N \) (2.6, 0.2) |
| 6 | BRAf+activeRap1 → BRAfPP+activeRap1 | 0.4 | 17991.179 | \( \log_{10} \) \( k \) = \( N \) (1.0, 0.2) |
| 7 | activeRap1+Gap → inactiveRap1+Gap | 1.0 | 6808.32 | \( \log_{10} \) \( k \) = \( N \) (0.4, 0.2) |
| 8 | BRAf+activeRas → BRAfPP+activeRas | 0.5 | 7631.63 | - |
| 9 | activeRas+Gap → inactiveRas+Gap | 0.0 | 12457.816 | - |
| 10 | inactiveRas+activeSOS → activeRas+activeSOS | 0.5 | 13.73 | - |
| 11 | activeSOS → inactiveSOS | 4.0 | 9834.13 | - |
| 12 | inactiveSOS+BEGFR → activeSOS+BEGFR | 2.5 | 8176.56 | - |

\( a \) logarithm (base rate constant value)  
\( b \) Base value of Michaelis constant (Obtained from Xu et al. [?])

Table 2: Proposed reactions for Example 1

There are five effective networks in Example 1. Figure 2 shows the network diagrams of all effective networks in Example 1. As an example of our network-aware proposal construction, we present the between-model proposal for the move between effective network 1 and effective network 2 in Figure 3. The move from effective network 1 to effective network 2 involves the addition of reactions 4. The identify jump function \( f \) for the between-model move between the two networks is:

\[ k_{2, \{3,5,6\}} = k_{1, \{3,5,6\}}, \quad k_{2,4} = u, \]  
(26)
where the proposal distribution \( q(u|k_{1, (3,5,6)}) \) is \( \mathcal{N}(u; \mu, \Sigma) \), with
\[
\mu = \arg \max_{k_{2,4}} p(k_{2,4}|k_{1, (3,5,6)}, M_2, \mathcal{D})
\]
and
\[
\Sigma = -[\nabla^2 \log p(k_{2,4}|k_{1, (3,5,6)}, M_2, \mathcal{D})]^{-1}|_{k_{2,4}=\mu}
\]

2.2 Example 2: ten-dimensional nonlinear network inference

Our second example is a large scale nonlinear network inference problem with 10 uncertain reactions. Once again, we consider a protein-signalling network consisting of 15 species and 12 potential species interactions (Figure 1). The ODE forward model governing the evolution of species concentrations is described in detail in Section 3. The rate constants of all fixed reactions and Michaelis constants of all reactions are set to their base values (Table 3). Reactions 3–12 are uncertain and the concentration of BRaf is again the observable.

| Reaction | \( \log_{10} k^a \) | \( k_M^b \) | Prior uncertainty |
|----------|-----------------|--------------|------------------|
| 1 | BEGFR → DEGFR | 0.0 | - | - |
| 2a | EGF + UEGFR → BEGFR | 1.5 | - | - |
| 2b | BEGFR → EGF + UEGFR | 0.0 | - | - |
| 3 | inactiveC3G+BEGFR → activeC3G+BEGFR | 0.5 | 3386.3875 | \( \log_{10} k = \mathcal{N}(1.2, 0.1) \) |
| 4 | activeC3G → inactiveC3G | 2.0 | - | \( \log_{10} k = \mathcal{N}(2.0, 0.1) \) |
| 5 | inactiveRap1+activeC3G → activeRap1+activeC3G | 2.0 | 3566 | \( \log_{10} k = \mathcal{N}(2.7, 0.1) \) |
| 6 | BRaf+activeRap1 → BRafPP+activeRap1 | 0.4 | 17991.179 | \( \log_{10} k = \mathcal{N}(1.1, 0.1) \) |
| 7 | activeRap1+Gap → inactiveRap1+Gap | 1.0 | 6808.32 | \( \log_{10} k = \mathcal{N}(1.0, 0.01) \) |
| 8 | BRaf+activeRas → BRafPP+activeRas | 0.5 | 7631.63 | \( \log_{10} k = \mathcal{N}(0.5, 0.1) \) |
| 9 | activeRas+Gap → inactiveRas+Gap | 0.0 | 12457.816 | \( \log_{10} k = \mathcal{N}(0.0, 0.01) \) |
| 10 | inactiveRas+activeSOS → activeRas+activeSOS | 0.5 | 13.73 | \( \log_{10} k = \mathcal{N}(0.5, 0.1) \) |
| 11 | activeSOS → inactiveSOS | 4.0 | 9834.13 | \( \log_{10} k = \mathcal{N}(4.0, 0.01) \) |
| 12 | inactiveSOS+BEGFR → activeSOS+BEGFR | 2.5 | 8176.56 | \( \log_{10} k = \mathcal{N}(2.5, 0.1) \) |

\( a \) logarithm (base rate constant value)  
\( b \) Base value of Michaelis constant (Obtained from Xu et al. [?])

Table 3: Proposed reactions for Example 2

There are twenty-four effective networks in Example 2. Figure 4 shows the network diagrams of all effective networks in Example 2. We again consider a move between the effective network 1 and effective network 2 of Figure 3 to demonstrate a prototypical between-model move proposal using our sensitivity-based network-aware proposals. Assume that the critical reaction to which the two networks are most sensitive is Reaction 8 and we only include one highly sensitive reaction for proposal construction. Thus in a move from effective network 1 to effective network 2, the rate constant of reaction 8 would not be kept fixed but instead be proposed according to a distribution. Reaction 4 is the additional reaction in network 2. All other reactions common to the two networks, namely reactions 3, 5, 6, 9, 10, 11, and 12 will have their rate constant values kept fixed during the move. The identity jump function \( f \) for the between-model move between the two effective networks is:
\[ k_{2,\{3,5,6,9,10,11,12\}} = k_{1,\{3,5,6,9,10,11,12\}} \text{, } k_{2,8} = u_{1,8} \text{, } u_{2,8} = k_{1,8} \text{, } k_{2,4} = u_{1,4}, \]

(29)

where the proposal distributions \( q(u_{1,\{4,8\}}|k_{1,\{3,5,6,9,10,11,12\}}) \) and \( q(u_{2,8}|k_{1,\{3,5,6,9,10,11,12\}}) \) are \( \mathcal{N}(u_1; \mu_1, \Sigma_1) \) and \( \mathcal{N}(u_2; \mu_2, \Sigma_2) \), respectively, with proposal means

\[
\mu_1 = \arg \max_{k_{2,\{4,8\}}} p(k_{2,\{4,8\}}|k_{1,\{3,5,6,9,10,11,12\}}, M_2, \mathcal{D})
\]

(30)

and

\[
\mu_2 = \arg \max_{k_{1,8}} p(k_{1,8}|k_{2,\{3,5,6,9,10,11,12\}}, M_1, \mathcal{D})
\]

(31)

and proposal covariances

\[
\Sigma_1 = -[\nabla^2 \log p(k_{2,\{4,8\}}|k_{1,\{3,5,6,9,10,11,12\}}, M_2, \mathcal{D})]^{-1}|_{\mu_1}
\]

(32)

and

\[
\Sigma_2 = -[\nabla^2 \log p(k_{1,8}|k_{2,\{3,5,6,9,10,11,12\}}, M_1, \mathcal{D})]^{-1}|_{\mu_2}.
\]

(33)

3 12-dimensional reaction network

Here we present the details of the set of proposed reactions, the corresponding reaction and species production rate ODE expressions for the 12-reaction network used in Example 1 and 2. The resulting nonlinear system of ordinary differential equations are solved using the multistep BDF integrator available in the SUNDIALS suite [?].

3.1 Reactions

1. \textit{boundEGFR} \rightarrow \textit{degradedEGFR}
2. \textit{EGF} + \textit{unboundEGFR} \leftrightarrow \textit{boundEGFR}
3. \textit{inactiveC3G} + \textit{boundEGFR} \rightarrow \textit{activeC3G} + \textit{boundEGFR}
4. \textit{activeC3G} \rightarrow \textit{inactiveC3G}
5. \textit{inactiveRap1} + \textit{activeC3G} \rightarrow \textit{activeRap1} + \textit{activeC3G}
6. \textit{BRaf} + \textit{activeRap1} \rightarrow \textit{BRafPP} + \textit{activeRap1}
7. \textit{activeRap1} + \textit{Gap} \rightarrow \textit{inactiveRap1} + \textit{Gap}
8. \textit{BRaf} + \textit{activeRas} \rightarrow \textit{BRafPP} + \textit{activeRas}
9. \textit{activeRas} + \textit{Gap} \rightarrow \textit{inactiveRas} + \textit{Gap}
10. \textit{inactiveRas} + \textit{activeSOS} \rightarrow \textit{activeRas} + \textit{activeSOS}
11. \textit{activeSOS} \rightarrow \textit{inactiveSOS}
12. \textit{inactiveSOS} + \textit{boundEGFR} \rightarrow \textit{activeSOS} + \textit{boundEGFR}
3.2 Reaction rates

1. $k_1[\text{boundEGFR}]$
2. $k_2f[\text{EGF}][\text{unboundEGFR}] - k_2r[\text{boundEGFR}]$
3. $\frac{k_3[\text{boundEGFR}][\text{inactiveC3G}]}{k_3' + [\text{inactiveC3G}]}$
4. $k_4[\text{activeC3G}]$
5. $\frac{k_5[\text{activeC3G}][\text{inactiveRap1}]}{k_5' + [\text{inactiveRap1}]}$
6. $\frac{k_6[\text{activeRap1}][\text{BRaf}]}{k_6' + [\text{BRaf}]}$
7. $\frac{k_7[\text{Gap}][\text{activeRap1}]}{k_7' + [\text{activeRap1}]}$
8. $\frac{k_8[\text{activeRas}][\text{BRaf}]}{k_8' + [\text{BRaf}]}$
9. $\frac{k_9[\text{Gap}][\text{activeRas}]}{k_9 + [\text{activeRas}]}$
10. $\frac{k_{10}[\text{activeSOS}][\text{inactiveRas}]}{k_{10}' + \text{inactiveRas}}$
11. $\frac{k_{11}[\text{activeSOS}]}{k_{11}' + [\text{activeSOS}]}$
12. $\frac{k_{12}[\text{boundEGFR}][\text{inactiveSOS}]}{k_{12}' + [\text{inactiveSOS}]}$

3.3 Species production rates

1. $\dot{[\text{unboundEGFR}]} = -k_2f[\text{EGF}][\text{unboundEGFR}] + k_2r[\text{boundEGFR}]$
2. $\dot{[\text{inactiveSOS}]} = -\frac{k_{12}[\text{boundEGFR}][\text{inactiveSOS}]}{k_{12}' + [\text{inactiveSOS}]} + \frac{k_{11}[\text{activeSOS}]}{k_{11}' + [\text{activeSOS}]}$
3. $\dot{[\text{inactiveRas}]} = -\frac{k_{10}[\text{activeSOS}][\text{inactiveRas}]}{k_{10}' + [\text{inactiveRas}]} + \frac{k_9[\text{Gap}][\text{activeRas}]}{k_9 + [\text{activeRas}]}$
4. $\dot{[\text{inactiveRap1}]} = \frac{k_7[\text{Gap}][\text{activeRap1}]}{k_7' + [\text{activeRap1}]} - \frac{k_5[\text{activeC3G}][\text{inactiveRap1}]}{k_5' + [\text{inactiveRap1}]}$
5. $\dot{[\text{boundEGFR}]} = k_2f[\text{EGF}][\text{unboundEGFR}] - k_2r[\text{boundEGFR}] - k_1[\text{boundEGFR}]$
6. $\dot{[\text{activeSOS}]} = \frac{k_{12}[\text{boundEGFR}][\text{inactiveSOS}]}{k_{12}' + [\text{inactiveSOS}]} - \frac{k_{11}[\text{activeSOS}]}{k_{11}' + [\text{activeSOS}]}$
7. $\dot{[\text{activeRas}]} = \frac{k_{10}[\text{activeSOS}][\text{inactiveRas}]}{k_{10}' + [\text{inactiveRas}]} - \frac{k_9[\text{Gap}][\text{activeRas}]}{k_9 + [\text{activeRas}]}$
8. $[\text{activeRap}_1] = \frac{k_7[\text{Gap}][\text{activeRap}_1]}{k'_7 + [\text{activeRap}_1]} + \frac{k_5[\text{activeC}3\text{G}][\text{inactiveRap}_1]}{k'_5 + [\text{inactiveRap}_1]}$

9. $[\text{EGF}] = -k_2f[\text{EGF}][\text{unboundEGFR}] + k_2r[\text{boundEGFR}]$

10. $[\text{BRafPP}] = \frac{k_6[\text{activeRap}_1][\text{BRaf}]}{k'_6 + [\text{BRaf}]} + \frac{k_8[\text{activeRas}][\text{BRaf}]}{k'_8 + [\text{BRaf}]}$

11. $[\text{BRaf}] = -\frac{k_6[\text{activeRap}_1][\text{BRaf}]}{k'_6 + [\text{BRaf}]} - \frac{k_8[\text{activeRas}][\text{BRaf}]}{k'_8 + [\text{BRaf}]}$

12. $[\text{activeC}3\text{G}] = \frac{k_3[\text{boundEGFR}][\text{inactiveC}3\text{G}]}{k'_3 + [\text{inactiveC}3\text{G}]} - k_4[\text{activeC}3\text{G}]$

13. $[\text{inactiveC}3\text{G}] = -\frac{k_3[\text{boundEGFR}][\text{inactiveC}3\text{G}]}{k'_3 + [\text{inactiveC}3\text{G}]} + k_4[\text{activeC}3\text{G}]$

14. $[\text{degradedEGFR}] = k_1[\text{boundEGFR}]$

15. $[\text{Gap}] = 0$

### 3.4 Initial species concentrations

All simulations using the above reactions are performed with the following initial concentrations:

1. $[\text{unboundEGFR}]_0 = 500$
2. $[\text{inactiveSOS}]_0 = 1200$
3. $[\text{inactiveRas}]_0 = 1200$
4. $[\text{inactiveRap}_1]_0 = 1200$
5. $[\text{boundEGFR}]_0 = 0$
6. $[\text{activeSOS}]_0 = 0$
7. $[\text{activeRas}]_0 = 0$
8. $[\text{activeRap}_1]_0 = 0$
9. $[\text{EGF}]_0 = 1000$
10. $[\text{BRafPP}]_0 = 0$
11. $[\text{BRaf}]_0 = 1500$
12. $[\text{activeC}3\text{G}]_0 = 0$
13. $[\text{inactiveC}3\text{G}]_0 = 1200$
14. $[\text{degradedEGFR}]_0 = 0$
15. $[\text{Gap}]_0 = 2400$
References

[1] S. P. Brooks, P. Giudici, and G. O. Roberts. Efficient construction of reversible jump Markov chain Monte Carlo proposal distributions (with discussion). *Journal of Royal Statistical Society B*, 65:3–39, 2003.

[2] A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward. Sundials: Suite of nonlinear and differential/algebraic equation solvers, 2005.

[3] J. D. Murray. *Mathematical Biology: I. An Introduction*. Springer, 3rd edition, 2002.

[4] T.-R. Xu, V. Vyshemirsky, A. Gormand, A. von Krieger, M. Girolami, G. S. Baillie, D. Ketley, A. J. Dunlop, G. Milligan, M. D. Houslay, and W. Kolch. Inferring signaling pathway topologies from multiple perturbation measurements of specific biochemical species. *Science Signaling*, 3(113):ra20, 2010.
Figure 2: Effective networks in Example 1
Figure 3: A between-model move between two effective networks in Example 1 and Example 2
Figure 4: Effective networks in Example 2
Figure 5: Effective networks in Example 2 (contd.)