In this document, we provide theoretical details necessary to understand the Bayesian analysis approach introduced in the Main text. We start by deriving a general formula for the posterior density associated with our Bayesian analysis approach. We then discuss the Wegscheider conditions and their implications on the reaction rate constants of a closed biochemical reaction system. Subsequently, we derive appropriate prior probability density functions for the log-equilibrium constants as well as for the log-rate constants, and discuss a practical method for determining the hyperparameters associated with these priors. Finally, we present mathematical arguments to support our belief that the posterior mode should be a more preferable estimator of the kinetic parameters of a biochemical reaction system than the posterior mean.
Bayesian analysis

To develop a Bayesian analysis approach for estimating the kinetic parameters (rate constants) of a biochemical reaction system, we take the log-rate constants $\kappa$ and the error variance $\sigma^2$ to be random variables. Note however that the probability density function of $\kappa$ we consider in this paper depends on the log-equilibrium constants $z = \{z_m, m \in M\}$, where $z_m := \ln(k_{2m-1}/k_{2m})$, which we treat as hyperparameters (details to follow). Since we do not know the exact values of these hyperparameters, we take them to be random variables as well. By following this approach, we can write the posterior joint density of $\kappa$, $z$, and $\sigma^2$, given data $y$, as

$$p(\kappa, z, \sigma^2 | y) = p(\sigma^2 | \kappa, y)p(\kappa | z, y)p(z | y).$$  

(S-1.1)

Since our main objective is to estimate $\kappa$, we are interested in evaluating the posterior density $p(\kappa | y)$. To do so, we must integrate the log-equilibrium constants $z$ and the error variance $\sigma^2$ out of the posterior joint density $p(\kappa, z, \sigma^2 | y)$. In this case,

$$p(\kappa | y) = \int \int p(\kappa, z, \sigma^2 | y)d\sigma^2 dz = \int p(\kappa | z, y)p(z | y)dz,$$

(S-1.2)

by virtue of (S-1.1). On the other hand,

$$p(\kappa | z, y) = \frac{p(y | \kappa)p(\kappa | z)p(z)}{p(z | y)p(y)}.$$  

(S-1.3)

As a consequence of (S-1.2) and (S-1.3), we have that

$$p(\kappa | y) = \frac{p(y | \kappa)}{p(y)} \int p(\kappa | z)p(z)dz \propto p(y | \kappa) \int p(\kappa | z)p(z)dz,$$

(S-1.4)

where $p \propto q$ denotes that $p$ is proportional to $q$, whereas,

$$p(y | \kappa) = \int p(y | \kappa, \sigma^2)p(\sigma^2 | \kappa)d\sigma^2.$$  

(S-1.5)

The term $p(y | \kappa)$ is simply the average of the likelihood $p(y | \kappa, \sigma^2)$ over the conditional prior density $p(\sigma^2 | \kappa)$ of $\sigma^2$ given the values of the log-rate constants $\kappa$. We refer to $p(y | \kappa)$ as the “effective” likelihood. Moreover, we refer to $\int p(\kappa | z)p(z)dz$ as the “effective” prior, since this term contains the prior information about the log-rate constants $\kappa$ after the prior information about the log-equilibrium constants is integrated out of the problem.

\(^1\)It is clear that $z$ can be directly calculated from $\kappa$. Consequently, conditioning on both $\kappa$ and $z$ is equivalent to conditioning only on $\kappa$. Therefore, $p(\sigma^2 | \kappa, z, y) = p(\sigma^2 | \kappa, y)$ and $p(y | \kappa, z) = p(y | \kappa)$. 

2
Wegscheider conditions

As we mentioned in the Main text, the thermodynamic constraints given by Equation (10) are equivalent to the Wegscheider conditions given by Equation (11). To show this fact, let us first assume that Equation (10) is satisfied. Then, for every \( r = \{r_m, m \in M\} \in \text{null}(S) \),

\[
\ln \prod_{m \in M} \left( \frac{k_{2m-1}}{k_{2m}} \right)^{r_m} = \sum_{m \in M} r_m \ln \frac{k_{2m-1}}{k_{2m}} = \sum_{m \in M} r_m \ln \frac{\prod_{n \in N} \left[ \frac{\nu'_n}{\nu_n} \right]^{r_m}}{\prod_{n \in N} \left[ \frac{\nu'_n}{\nu_n} \right]^{s_{nm}}} = \sum_{n \in N} \left( \sum_{m \in M} s_{nm} r_m \right) \ln \frac{\prod_{n \in N} \left[ \frac{\nu'_n}{\nu_n} \right]}{\prod_{n \in N} \left[ \frac{\nu'_n}{\nu_n} \right]} = 0,
\]

where \( \{x_n^{(p)}, n \in N\} \) are the stationary concentrations in the system when the initial concentration of the \( p \)th molecular species is perturbed. This shows Equation (11).

On the other hand, if Equation (11) is satisfied, then

\[
\sum_{m \in M} r_m \ln \frac{k_{2m-1}}{k_{2m}} \prod_{n \in N} \left[ \frac{\nu'_n}{\nu_n} \right]^{r_m} = \sum_{m \in M} \ln \left( \frac{k_{2m-1}}{k_{2m}} \right)^{r_m} - \sum_{m \in M} r_m \ln \prod_{n \in N} \left[ \frac{\nu'_n}{\nu_n} \right]^{s_{nm}} = \sum_{m \in M} \ln \left( \frac{k_{2m-1}}{k_{2m}} \right)^{r_m} - \sum_{n \in N} \left( \sum_{m \in M} s_{nm} r_m \right) \ln x_n^{(p)} = \sum_{m \in M} \ln \left( \frac{k_{2m-1}}{k_{2m}} \right)^{r_m} = 0,
\]

for every \( r \in \text{null}(S) \). We can take the \( m \)th element \( r_m \) of \( r \) to be

\[
r_m = k_{2m-1} \prod_{n \in N} \left[ \frac{\nu'_n}{\nu_n} \right]^{r_m} - k_{2m} \prod_{n \in N} \left[ \frac{\nu'_n}{\nu_n} \right]^{s_{nm}},
\]

since, from Equations (2) and (4) in the Main text, we have \( S r = 0 \), which implies that \( r \in \text{null}(S) \). Note now that

\[
\sum_{m \in M} \left( k_{2m-1} \prod_{n \in N} \left[ \frac{\nu'_n}{\nu_n} \right]^{r_m} - k_{2m} \prod_{n \in N} \left[ \frac{\nu'_n}{\nu_n} \right]^{s_{nm}} \right) \ln \frac{k_{2m-1}}{k_{2m}} \prod_{n \in N} \left[ \frac{\nu'_n}{\nu_n} \right]^{s_{nm}} = 0,
\]

3
by virtue of (S-1.6). However, each term in the previous sum is nonnegative. Hence, for the sum to be equal to zero, each term must be zero, which implies the steady-state thermodynamic constrains given by Equation (10) in the Main text.

We will now show that the Wegscheider conditions are satisfied for all \( \mathbf{r} \in \text{null}(\mathbf{S}) \) so long as they are satisfied for \( M_2 = M - M_1 \) basis vectors \( \{\mathbf{r}(j), j = 1, 2, \ldots, M_1\} \) of the null space of \( \mathbf{S} \), where \( M_1 = \text{rank}(\mathbf{S}) \). Note that, for any \( \mathbf{r} \in \text{null}(\mathbf{S}) \),

\[
\mathbf{r} = \sum_{j=1}^{M_2} a_j \mathbf{r}(j),
\]

for some scalar coefficients \( a_j, j = 1, 2, \ldots, M_2 \). As a consequence, and from Equations (8) and (11) in the Main text, we have that

\[
\ln \prod_{m \in \mathcal{M}} \left( \frac{k_{2m-1}}{k_{2m}} \right)^{r_m} = \sum_{m \in \mathcal{M}} r_m \ln \frac{k_{2m-1}}{k_{2m}}
\]

\[
= \sum_{m \in \mathcal{M}} r_m z_m
\]

\[
= \sum_{m \in \mathcal{M}} \left[ \sum_{j=1}^{M_2} a_j r_m(j) \right] z_m
\]

\[
= \sum_{j=1}^{M_2} a_j \sum_{m \in \mathcal{M}} z_m r_m(j)
\]

\[
= \sum_{j=1}^{M_2} a_j \sum_{m \in \mathcal{M}} r_m(j) \ln \frac{k_{2m-1}}{k_{2m}}
\]

\[
= \sum_{j=1}^{M_2} a_j \ln \prod_{m \in \mathcal{M}} \left( \frac{k_{2m-1}}{k_{2m}} \right)^{r_m(j)}
\]

\[
= 0,
\]

for every \( \mathbf{r} = \{r_m, m \in \mathcal{M}\} \in \text{null}(\mathbf{S}) \), since the Wegscheider conditions are assumed to be satisfied by the basis vectors of the null space of \( \mathbf{S} \). This shows that the Wegscheider conditions are also satisfied for every \( \mathbf{r} \in \text{null}(\mathbf{S}) \).

Let us now rearrange the columns and rows of the stoichiometry matrix \( \mathbf{S} \) (by appropriately relabeling the molecular species and reactions) so that the first \( M_1 \) columns are linearly independent, whereas, the
remaining $M_2$ columns are linearly dependent on the first columns. In this case, we can write the stoichiometry matrix $S$ in the following block matrix form:

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix},$$

where $S_{11}$ is an invertible $M_1 \times M_1$ matrix, whereas, $S_{12}, S_{21}$, and $S_{22}$ are $M_1 \times M_2$, $(N - M_1) \times M_1$, and $(N - M_1) \times M_2$ matrices, respectively. It is a well-known fact (e.g., see [1]) that the general solution of $Sr = 0$ is given by $r' = -S_{11}^{-1}S_{12}r''$, for an arbitrary $r''$, where $r', r''$ are $M_1 \times 1$ and $M_2 \times 1$ vectors, respectively. Let $S^{-1}_{11}S_{12}$ be the element of the $m$th row and the $m'$th column of matrix $S_{11}^{-1}S_{12}$.

This implies that the columns of matrix

$$B = \begin{bmatrix} -S_{11}^{-1}S_{12} \\ I_{M_2} \end{bmatrix},$$

where $I_{M_2}$ is the $M_2 \times M_2$ identity matrix, form a basis for the null space of $S$. As a consequence of this result and the fact that the Wegscheider conditions are satisfied so long they are satisfied by a basis vector of null($S$), we can conclude that the Wegscheider conditions, given by Equation (11) in the Main text, are equivalent to the following conditions [2]:

$$\kappa_{2m'} = \kappa_{2m'-1} + \sum_{m \in M_1} [S_{11}^{-1}S_{12}]_{m,m'}(\kappa_{2m} - \kappa_{2m-1}), \quad \text{for every } m' \in M_2, \quad (S-1.7)$$

where $M_1 = \{1, 2, \ldots, M_1\}$, $M_2 = \{M_1 + 1, M_1 + 2, \ldots, M\}$, and $[S_{11}^{-1}S_{12}]_{m,m'}$ is the element of the $m$th row and the $m'$th column of matrix $S_{11}^{-1}S_{12}$.

Equation (S-1.7) allows us to specify arbitrary values for the forward and reverse log-rate constants \{\kappa_{2m-1}, m \in M_1\}, \{\kappa_{2m}, m \in M_1\}, and calculate the reverse log-rate constants \{\kappa_{2m}, m \in M_2\} so that the Wegscheider conditions are satisfied. If we denote by $\kappa_f$ the $(M + M_1)$ “free” log-rate constants \{\kappa_{2m-1}, m \in M, \kappa_{2m}, m \in M_1\} and by $\kappa_d$ the $M_2$ “dependent” log-rate constants \{\kappa_{2m}, m \in M_2\}, then we can write (S-1.7) in the following compact form:

$$\kappa_d = \mathbb{W}\kappa_f, \quad (S-1.8)$$

where $\mathbb{W}$ is the $(M - M_1) \times (M + M_1)$ matrix implementing the right hand side of (S-1.7), given by

$$\mathbb{W} = \begin{bmatrix} -(S_{11}^{-1}S_{12})^T \\ I_{M_2} \\ (S_{11}^{-1}S_{12})^T \end{bmatrix}. \quad (S-1.9)$$

Thus, for any arbitrary $\kappa_f$, we can construct a thermodynamically feasible set of kinetic parameters by determining the dependent parameters $\kappa_d$ according to (S-1.8), and by setting $\kappa = \{\kappa_f, \kappa_d\}$. 

5
**Prior density of log-equilibrium constants**

From Equations (8) and (10) in the Main text, we have that

\[ z_m = \sum_{n \in \mathcal{N}} s_{nm} \ln x_n^{(p)}, \quad \text{for all } p \in \mathcal{P}, \]

where \( s_{nm} = \nu_{nm}' - \nu_{nm} \) is the net stoichiometry coefficient and \( \{ x_n^{(p)}, n \in \mathcal{N} \} \) are the stationary concentrations when the initial concentration of the \( p \)-th molecular species is perturbed. Therefore,

\[ z_m = \frac{1}{P + 1} \sum_{p \in \mathcal{P}} \sum_{n \in \mathcal{N}} s_{nm} \ln x_n^{(p)}, \quad \text{for all } m \in \mathcal{M}. \]  

(S-1.10)

On the other hand, Equation (5) in the Main text gives

\[ \sum_{p \in \mathcal{P}} \sum_{n \in \mathcal{N}} s_{nm} \ln x_n^{(p)}(t_q) = \sum_{p \in \mathcal{P}} \sum_{n \in \mathcal{N}} s_{nm} y_n^{(p)}(t_q) - \sum_{p \in \mathcal{P}} \sum_{n \in \mathcal{N}} s_{nm} \eta_n^{(p)}(t_q). \]

(S-1.11)

If we assume that the biochemical reaction system and all its perturbed versions are sufficiently close to steady-state at some time point \( t_{q+1} \), then (S-1.10) and (S-1.11) approximately imply that

\[ z_m = \tilde{y}_m - \tilde{\eta}_m, \]  

(S-1.12)

where

\[ \tilde{y}_m := \frac{1}{P + 1} \sum_{p \in \mathcal{P}} \sum_{n \in \mathcal{N}} s_{nm} y_n^{(p)}(t_{q+1}), \]  

(S-1.13)

and

\[ \tilde{\eta}_m := \frac{1}{P + 1} \sum_{p \in \mathcal{P}} \sum_{n \in \mathcal{N}} s_{nm} \eta_n^{(p)}(t_{q+1}). \]

To proceed, note that

\[ p(z \mid \bar{y}) = p(z \mid \bar{y}, \bar{y}) = \int p(z, \sigma^2 \mid \bar{y}, \bar{y}) d\sigma^2 \]

\[ = \int p(z \mid \sigma^2, \bar{y}) p(\sigma^2 \mid \bar{y}) d\sigma^2 \]

\[ = \int p(z \mid \sigma^2, \bar{y}) p(\sigma^2 \mid \bar{y}) d\sigma^2. \]  

(S-1.14)

This is due to the fact that \( \bar{y} := \{ \bar{y}_m, m \in \mathcal{M} \} \) can be calculated from \( \bar{y} := \{ y_n^{(p)}(t_{q+1}), n \in \mathcal{N}, p \in \mathcal{P} \} \) by means of (S-1.13), in which case \( p(z \mid \bar{y}, \bar{y}) = p(z \mid \bar{y}) \) and \( p(\sigma^2 \mid \bar{y}, \bar{y}) = p(\sigma^2 \mid \bar{y}) \). Moreover, since
\[ z_m = \tilde{y}_m - \tilde{\eta}_m, \] 
\[ \bar{y} \] does not provide further information about \( z \), given \( \sigma^2 \) and \( \bar{y} \), in which case 
\[ p(z \mid \sigma^2, \bar{y}) = p(z \mid \sigma^2, \bar{y}). \]

Equation (S-1.14) implies that, in order to calculate the probability density function \( p(z \mid \bar{y}) \), we must determine the probability density functions \( p(z \mid \sigma^2, \bar{y}) \) and \( p(\sigma^2 \mid \bar{y}) \). Note that, given \( \sigma^2 \) and \( \bar{y} \), the log-equilibrium constants \( z \) follow a multivariate Gaussian distribution with means and covariances 
\[ \text{E}[z_m \mid \sigma^2, \bar{y}] = \tilde{y}_m \quad \text{and} \quad \text{cov}[z_m, z_{m'} \mid \sigma^2, \bar{y}] = \frac{\sigma^2}{P + 1} \sum_{n \in \mathcal{N}} s_{nm}s_{nm'}. \]

This implies that 
\[ p(z \mid \sigma^2, \bar{y}) = \frac{(P + 1)^{M/2}}{(2\pi)^{M/2}\sigma^M\|HH^{-1}\|^{1/2}} \exp \left\{ -\frac{P + 1}{2\sigma^2}(z - \tilde{y})^T HH^{-1}(z - \tilde{y}) \right\}, \quad (S-1.15) \]
where \( HH \) is an \( M \times M \) matrix with elements \( h_{mm'} = \sum_{n \in \mathcal{N}} s_{nm}s_{nm'} \). Note that \( HH = S^T S \), where \( S \) is the \( N \times M \) stoichiometry matrix of the biochemical reaction system with elements \( s_{nm} \).

To determine the probability density function \( p(\sigma^2 \mid \bar{y}) \) of the error variance, note that 
\[ p(\bar{X}, \sigma^2 \mid \bar{y}) = \frac{p(\bar{y} \mid \bar{X}, \sigma^2)p(\bar{X}, \sigma^2)}{p(\bar{y})}, \quad (S-1.16) \]
where \( \bar{x} := \{ \ln x_n^{(p)}, n \in \mathcal{N}, p \in \mathcal{P} \} \) are the stationary concentrations when the initial concentration of the \( p \)-th molecular species is perturbed. Moreover, 
\[ p(\bar{X}, \sigma^2 \mid \bar{y}) = p(\bar{X} \mid \sigma^2, \bar{y})p(\sigma^2 \mid \bar{y}). \quad (S-1.17) \]
As a consequence of (S-1.16) and (S-1.17), we have 
\[ p(\sigma^2 \mid \bar{y}) = \frac{p(\bar{X}, \sigma^2 \mid \bar{y})}{p(\bar{X} \mid \sigma^2, \bar{y})} \]
\[ = \frac{p(\bar{y} \mid \bar{X}, \sigma^2)p(\bar{X}, \sigma^2)}{p(\bar{X} \mid \sigma^2, \bar{y})p(\bar{y})} \]
\[ = \frac{p(\bar{y} \mid \bar{X}, \sigma^2)p(\bar{X})p(\sigma^2)}{p(\bar{X} \mid \sigma^2, \bar{y})p(\bar{y})} \]
\[ \propto \frac{p(\bar{y} \mid \bar{X}, \sigma^2)}{p(\bar{X} \mid \sigma^2, \bar{y})}, \]
where we use the fact that \( \bar{x} \) and \( \sigma^2 \) are statistically independent (since \( \bar{x} \) is determined from \( \kappa \) and we have assumed in the Main text that \( \kappa \) and \( \sigma^2 \) are statistically independent). It is now not difficult to see that
\[ p(\tilde{y} \mid \bar{x}, \sigma^2) = p(\bar{x} \mid \sigma^2, \tilde{y}), \] due to the statistical independence and Gaussianity of the error terms \( \eta_n^{(p)} \) in Equation (5) of the Main text. Therefore,
\[
p(\sigma^2 \mid \tilde{y}) = p(\sigma^2) = \frac{b^a}{\Gamma(a)} (\sigma^2)^{-(a+1)} e^{-b/\sigma^2}, \tag{S-1.18}
\]
which does not depend on \( \tilde{y} \), by virtue of Equation (9) in the Main text.

As a consequence of (S-1.14) and (S-1.18), we now have that
\[
p(z \mid \tilde{y}) = p(z \mid \bar{y}, \tilde{y}) = p(z \mid \bar{y}) \tag{S-1.19}
\]
Finally, from (S-1.14), (S-1.15), (S-1.18), and (S-1.19), and after some straightforward, albeit tedious, algebraic manipulations, we can show that
\[
p(z \mid e, \tilde{y}) \propto \int \frac{1}{\sigma^{2(a+1)+M}} \exp \left\{ -\frac{P+1}{2\sigma^2} (z - \tilde{y})^T \mathbb{H}^{-1} (z - \tilde{y}) - \frac{b}{\sigma^2} \right\} d\sigma^2 
\approx \left[ \frac{2b}{P+1} + (z - \tilde{y})^T \mathbb{H}^{-1} (z - \tilde{y}) \right]^{-(M/2+a)} \tag{S-1.20}
\]
A problem associated with the previous formulation is that it may not be possible to evaluate the prior density \( p(z \mid \tilde{y}) \) of the log-equilibrium constants, given by (S-1.20), since the matrix \( \mathbb{H} \) may not be invertible. Indeed, if \( r \) is a (nonzero) vector in the null space of the \( N \times M \) stoichiometry matrix \( S \) (i.e., if \( S \mathbf{r} = 0 \)), then \( r^T \mathbb{H} \mathbf{r} = r^T S^T S \mathbf{r} = 0 \), which shows that matrix \( \mathbb{H} \) is positive semi-definite and, hence, not necessarily invertible.

To address the previous problem, we will follow a well-known technique known as *decorrelation* or *whitening* that allows us to transform the dependent random variables \( z \) into the statistically independent zero-mean random variables \( z_0 \) and obtain a form of matrix \( \mathbb{H} \) that is always invertible. Subsequently, this will allow us to derive a form for the probability density function \( p(z \mid \tilde{y}) \) that we can always evaluate.

Let us consider an \( (M - M_0) \)-dimensional zero-mean Gaussian random vector \( z_0 \) with identity covariance matrix, where \( M_0 \) is the number of zero eigenvalues of \( \mathbb{H} \). Our first objective is to determine an \( M \times (M - M_0) \) matrix \( \mathbb{H}_0 \), such that
\[
z = \mathbb{H}_0 z_0 + \tilde{y} \tag{S-1.21}
\]
Note that \( \mathbb{E}[z \mid \sigma^2, \tilde{y}] = \tilde{y} \), as expected from (S-1.12), whereas, \( \mathbb{H} = [(P + 1)/\sigma^2] \mathbb{H}_0 \mathbb{H}_0^T \). We can use singular value decomposition (SVD) [3] to decompose matrix \( \mathbb{H} = S^T S \) into \( \mathbb{H} = \mathbb{U} \mathbb{D} \mathbb{U}^T \), form the
\[(M - M_0) \times (M - M_0)\) diagonal matrix \(D_0\) by removing the last \(M_0\) zero singular values from \(D\) and the 
\(M \times (M - M_0)\) submatrix \(U_0\) of \(U\) by removing the last \(M_0\) columns of \(U\). Then, \(H_0 = U_0D_0U_0^T/\sqrt{P + I}\). Note now that \(U_0^TU_0 = I\), where \(I\) denotes the identity matrix. As a 
consequence, given vectors \(zzz\) and \(\tilde{y}\), (S-1.21) has a unique solution, given by 
\[z_0 = \frac{\sqrt{P + I}}{\sigma} D_0^{-1/2} U_0^T (zzz - \tilde{y}).\] 
(S-1.22)

This formula transforms the dependent random variables \(zzz\) into the statistically independent zero-mean 
random variables \(z_0\).

Note that
\[p(z | \sigma^2, \tilde{y}) = \int p(z, z_0 | \sigma^2, \tilde{y}) \, dz_0\]
\[= \int p(z | z_0, \sigma^2, \tilde{y}) p(z_0) \, dz_0\]
\[= \frac{1}{(2\pi)^{(M-M_0)/2}} \int \delta(z - H_0z_0 - \tilde{y}) \exp \left\{ -\frac{1}{2} z_0^T z_0 \right\} \, dz_0\]
\[= \frac{1}{(2\pi)^{(M-M_0)/2}} \delta(H_0z_0 - \tilde{y}) \exp \left\{ -\frac{P + 1}{2\sigma^2} (z - \tilde{y})^T U_0 D_0^{-1} U_0^T (z - \tilde{y}) \right\}\]
by virtue of (S-1.21) and (S-1.22), where \(\delta(\cdot)\) is the Dirac delta function. This result shows that 
\[p(z | \sigma^2, \tilde{y}) \propto \exp \left\{ -\frac{P + 1}{2\sigma^2} (z - \tilde{y})^T U_0 D_0^{-1} U_0^T (z - \tilde{y}) \right\},\]
which leads to [compare with (S-1.20)]:
\[p(z | \tilde{y}) \propto \frac{1}{\alpha^{a+1}} \exp \left\{ -\frac{P + 1}{2\sigma^2} (z - \tilde{y})^T U_0 D_0^{-1} U_0^T (z - \tilde{y}) - \frac{b}{\sigma^2} \right\} \, d\sigma^2\]
\[\propto \left[ \frac{2b}{P + 1} + (z - \tilde{y})^T U_0 D_0^{-1} U_0^T (z - \tilde{y}) \right]^{-a}.\]

Note that we can always evaluate this form of \(p(z | \tilde{y})\), since the diagonal matrix \(D_0\) is invertible.
If we now replace $p(z)$ by $p(z \mid \tilde{y})$ in Equation (6) of the Main text, we have

$$
\int p(\kappa \mid y) d\kappa = \frac{1}{p(y)} \int \int p(y \mid \kappa, z)p(\kappa \mid z)p(z \mid \tilde{y}) d\kappa dz
$$

$$
= \frac{1}{p(y)} \int \int p(y \mid \kappa, z)p(\kappa \mid z)p(\tilde{y}) dz
$$

$$
= \frac{1}{p(y)} \int p(y \mid z)p(z \mid \tilde{y}) dz
$$

$$
= \frac{1}{p(y)} \int p(y \mid z)p(z \mid \tilde{y}) dz
$$

$$
= 1, \text{ for all } y.
$$

since $p(y \mid \kappa, z) = p(y \mid \kappa)$ (see footnote 1), provided that $\tilde{y}$ is statistically independent of $y$, in which case $p(y \mid z, \tilde{y}) = p(y \mid z)$ and $p(y \mid \tilde{y}) = p(y)$. Clearly, the independence between $\tilde{y}$ and $y$ is a sufficient condition for the posterior density $p(\kappa \mid y)$ to be proper [i.e., for $p(\kappa \mid y)$ to be finite for every $y$].

**Prior density of log-rate constants**

Let us consider a bimolecular reaction $X_1 + X_2 \rightarrow X_3$. For this reaction to occur, two events must take place: one molecule of $X_1$ must collide with one molecule of $X_2$ and, given that the two molecules have collided, the reaction must take place. Using basic probabilistic arguments and the well-known hard-sphere collision theory [4], it has been shown in [5] that the probability of a randomly selected pair of molecules $X_1$ and $X_2$ at time $t$ to react during an infinitesimal time interval $[t, t + dt)$ is given by $c dt$. Here, $c$ is known as the specific probability rate constant and is given by

$$
c = \frac{\pi(r_1 + r_2)^2}{V} \sqrt{\frac{8k_BT}{\pi\mu^*\gamma}},
$$

(S-1.23)

where $V$ is the volume of the biochemical reaction system, $T$ is the temperature, $k_B$ is the Boltzmann constant ($k_B = 1.3806504 \times 10^{-23}$ J/K), and $\gamma$ is the probability that a randomly selected pair of colliding molecules $X_1$ and $X_2$ will react. This formula is based on the assumption that each molecule $X_n$ is a hard sphere of radius $r_n$ and mass $\mu_n$. In (S-1.23), $\mu^* = \mu_1\mu_2/(\mu_1 + \mu_2)$.

The rate constant $k$ of the previous bimolecular reaction is associated to the specific probability rate constant $c$ by means of $k = AVc$ (for a reaction with different reactants; see [6]), where $A$ is the Avogadro constant ($A = 6.0221415 \times 10^{23}$ mol$^{-1}$). If we assume that a pair $X_1$ and $X_2$ of molecules react only after
collision with impact energy greater than $E$, then $\gamma = \exp(-E/k_B T)$ \[5\]. In this case,

$$k = \alpha e^{-E/k_B T},$$

where

$$\alpha := \frac{A\pi}{8k_B T} \left( \frac{8k_B T}{\pi \mu} \right)^{1/2}.$$

Equation (S-1.24) is the well-known Arrhenius formula for the rate constant of a bimolecular reaction, and holds for other types of reactions (e.g., monomolecular and trimolecular) as well. The coefficient $\alpha$ is usually referred to as the pre-exponential factor, or simply the prefactor, whereas, $E$ is referred to as the activation energy of the reaction. In the following, we use (S-1.24) to derive a probabilistic model for the log-rate constants of a biochemical reaction system that leads to an appropriate prior density $p(\kappa | z)$ for the parameters $\kappa$.

According to (S-1.24), the rate constant of the $m$th forward reaction is given by

$$k_{2m-1} = \alpha_m e^{-E_m/k_B T}.$$

Equation (S-1.25) provides a theoretical expression for the prefactor $\alpha_m$, assuming the ideal situation of both reactant molecules being perfect hard spheres. In reality, however, the situation is much more complex, and we can use (S-1.25) to predict only a portion of the true prefactor value (provided that the masses and radii of the reactant molecules are known). As a consequence, we can decompose the prefactor $\alpha_m$ into a predictable part $\alpha^0_m$, and an unpredictable part $\omega_m$, so that $\alpha_m = \alpha^0_m \omega_m$. This implies that $\ln \alpha_m = \ln \alpha^0_m + g_m$, where $g_m := \ln \omega_m$ is a random additive component. The multiplicative factor $\omega_m$ can also be used to model unpredictable changes in biochemical conditions or changes in the structure of the reactant molecules, which may also affect the probability of particle collision and thus $\alpha_m$.

Therefore, we will be using the following expression for the prefactor of the $m$th forward reaction:

$$\alpha_m = \alpha^0_m e^{g_m}.$$

We will assume that $g_m$ is a zero-mean Gaussian noise component with standard deviation $\lambda_m$. In this case, $\alpha_m$ is a random variable characterized by the log-normal distribution

$$p(\alpha_m) = \frac{1}{\alpha_m \lambda_m \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\lambda_m} \left( \ln \frac{\alpha_m}{\alpha^0_m} \right)^2 \right\} ,$$
with parameters $\ln \alpha_m^0$ and $\lambda_m$. It has been pointed out in [7] that log-normal distributions are very natural for modeling biochemical processes and are a direct consequence of the thermodynamic behavior of biochemical reaction systems.

Unpredictable changes in biochemical conditions can also affect the probability of reaction after collision, or equivalently, the activation energy $E_m$. We may consider $E_m$ to be a random variable that is decomposed into two terms: a deterministic term $E_0^m$ and a random additive term $U_m$, so that

$$E_m = E_0^m + U_m,$$

where $E_0^m, U_m \geq 0$. This is known as the (static) random activation energy (RAE) model [8]. A commonly used probability law for the random energy component $U_m$ is the Maxwell-Boltzmann (exponential) distribution [8]:

$$p(U_m) = \frac{1}{k_B T_m^*} \exp\left\{ -\frac{U_m}{k_B T_m^*} \right\}, \quad U_m \geq 0,$$

at some temperature $T_m^* > T$.

As a consequence of (S-1.26), (S-1.27), and (S-1.28), we have that

$$\kappa_{2m-1}^0 = \kappa_m^0 + g_m - w_m, \quad w_m \geq 0,$$

where

$$\kappa_m^0 := \ln \alpha_m^0 - \frac{E_0^m}{k_B T}$$

and

$$w_m := \frac{U_m}{k_B T}.$$

Moreover, (S-1.29) implies that the probability density function of $w_m$ is given by the following exponential distribution:

$$p(w_m) = \frac{1}{\tau_m} e^{-w_m/\tau_m}, \quad w_m \geq 0, \quad \tau_m > 0,$$

where $\tau_m := T_m^* / T > 1$. Note that the expected value of $w_m$ equals its standard deviation, with $E[w_m] = \text{sd}[w_m] = \tau_m$. In the following, we will assume that, for all $m \in \mathcal{M}$, $w_m$ is statistically independent of $g_m$. This is a reasonable assumption, considering the fact that these two random variables result from two different biophysical mechanisms, namely molecular collision and molecular reaction, which we may consider to be statistically independent.

\footnote{It has been suggested in the literature (e.g., in [8]) that $T_m^*$ must be larger than the system temperature $T$.}
The previous modeling steps lead to the following prior probability density function \( p(\kappa_{2m-1}) \) for the log-rate constant of the \( m \)th forward reaction:

\[
p(\kappa_{2m-1}) = \frac{e^{\lambda_{m}/2\tau_{m}}}{2\tau_{m}} \text{erfc} \left( \frac{1}{\sqrt{2}} \left( \frac{\lambda_{m}}{\tau_{m}} + \frac{\kappa_{2m-1} - \kappa_{0m}}{\lambda_{m}} \right) \right) \exp \left( \frac{\kappa_{2m-1} - \kappa_{0m}}{\tau_{m}} \right), \tag{S-1.31}
\]

where \( \text{erfc}[\cdot] \) is the complementary error function \([9]\). To derive this result, consider a random variable

\[
y = c + g - w, \tag{S-1.32}
\]

where \( c \) is a constant and \( w, g \) are two statistically independent random variables so that

\[
p_w(w) = \frac{1}{\tau} e^{-w/\tau} \quad \text{(exponential)} \quad \text{and} \quad p_g(g) = \frac{1}{\sqrt{2\pi\lambda}} \exp \left( -\frac{g^2}{2\lambda^2} \right) \quad \text{(Gaussian)}.
\]

If we set \( u = w - g \), then \( y = c - u \) and

\[
p_y(y) = p_u(c - y). \tag{S-1.33}
\]

Moreover, since \( w \) and \( g \) are statistically independent, we have \([9]\)

\[
p_u(u) = \int_{-\infty}^{\infty} p_w(x)p_g(x-u)dx
\]

\[
= \int_{0}^{\infty} \frac{1}{\tau} e^{-x/\tau} \frac{1}{\sqrt{2\pi\lambda}} \exp \left( -\frac{(x-u)^2}{2\lambda^2} \right) dx
\]

\[
= \frac{1}{\sqrt{2\pi\lambda\tau}} \int_{0}^{\infty} e^{-[(x+\lambda^2/\tau-u)/2\lambda^2] - u^2/2\lambda^2} dx
\]

\[
= \frac{1}{\sqrt{2\pi\lambda\tau}} \int_{0}^{\infty} e^{-\xi^2/2\lambda^2} \exp \left( \frac{\xi^2}{2\lambda^2} - u^2/2\lambda^2 \right) dx
\]

\[
= \frac{1}{\sqrt{2\pi\lambda\tau}} \exp \left[ \frac{1}{\sqrt{2}} \left( \frac{\lambda}{\tau} - \frac{u}{\lambda} \right) \right] \exp \left( \frac{\kappa_{2m-1} - \kappa_{0m}}{\tau_{m}} \right). \tag{S-1.34}
\]

Finally, by combining (S-1.33) and (S-1.34), we obtain

\[
p_y(y) = \frac{e^{\lambda_{m}/2\tau_{m}}}{2\tau_{m}} \text{erfc} \left( \frac{1}{\sqrt{2}} \left( \frac{\lambda_{m}}{\tau_{m}} + \frac{y - c}{\lambda_{m}} \right) \right) \exp \left( \frac{y-c}{\tau_{m}} \right), \tag{S-1.35}
\]

which provides an analytical expression for the probability density function of \( y \). Equation (S-1.31) is now a direct consequence of (S-1.30), (S-1.32), and (S-1.35).
Let us now focus our attention on the log-rate constant $\kappa_{2m}$ of the $m^{th}$ reverse reaction. Basic thermodynamic arguments imply that

$$k_{2m-1} = k_{2m} \prod_{n \in N} \phi_n^{s_{nm}}, \quad (S-1.36)$$

where $k_{2m-1}$ and $k_{2m}$ are the rate constants of the $m^{th}$ forward and reverse reactions, respectively, $\phi_n$ is the capacity of the $n^{th}$ molecular species, and $s_{nm}$ is the stoichiometry coefficients of the $n^{th}$ molecular species associated with the $m^{th}$ reaction. The capacity is a thermodynamic quantity characteristic to a molecular species and depends on the standard chemical potential of that species (see [10] for details). As a consequence, the log-equilibrium constant $z_m$ of the $m^{th}$ reaction depends only on the stoichiometry of the biochemical reaction system and the capacities of the underlying molecular species, since Equation (8) in the Main text and (S-1.36) imply that

$$z_m = \sum_{n \in N} s_{nm} \ln \phi_n, \quad \text{for all } m \in M.$$ 

Therefore, $z_m$ is a constant characteristic to the $m^{th}$ reaction.

From Equation (8) in the Main text, note that

$$\kappa_{2m} = \kappa_{2m-1} - z_m, \quad \text{for all } m \in M,$$

which implies that $\kappa_{2m}$ and $\kappa_{2m-1}$ are two dependent random variables, given $z_m$. Their joint probability density function satisfies

$$p(\kappa_{2m}, \kappa_{2m-1} \mid z_m) = p(\kappa_{2m} \mid z_m, \kappa_{2m-1}) p(\kappa_{2m-1} \mid z_m) = \delta(\kappa_{2m} - \kappa_{2m-1} + z_m) p(\kappa_{2m-1}), \quad (S-1.37)$$

where $\delta(\cdot)$ is the Dirac delta function. Clearly, $z_m$ is a hyperparameter, since it characterizes the prior joint density $p(\kappa_{2m}, \kappa_{2m-1})$ of the log-rate constants of the $m^{th}$ reaction. We will assume here that, given $z = \{z_1, z_2, \ldots, z_M\}$, the reaction rate constants of different reactions are mutually independent. Then, the prior density of the log-rate constants will be given by

$$p(\kappa \mid z) = \prod_{m \in M} p(\kappa_{2m}, \kappa_{2m-1} \mid z_m) = \prod_{m \in M} \delta(\kappa_{2m} - \kappa_{2m-1} + z_m) p(\kappa_{2m-1}),$$

by virtue of (S-1.37).
A practical method for determining the hyperparameters $\phi, a, b$

In practice, prior information about the rate constants of a biochemical reaction system may be available from which we may be able to estimate their minimum, maximum, and average values. Moreover, some prior information might be available about the error processes $\eta_n^{(p)}$ in Equation (5) of the Main text, which may allow us to estimate the average value and spread of their variance. In this subsection, we show how to use these values to determine the hyperparameters $\phi = \{\kappa_0^m, \tau_m, \lambda_m, m \in M\}$ associated with the prior densities $p(\kappa_{2m-1})$ of the forward log-rate constants and the hyperparameters $a, b$ associated with the prior density $p(\sigma^2)$ of the measurement errors.

From (S-1.30), we have that

$$\text{E}[\kappa_{2m-1}] = \kappa_0^m - \tau_m \quad \text{and} \quad \text{sd}[\kappa_{2m-1}] = \sqrt{\lambda_m^2 + \tau_m^2},$$

(S-1.38)

by virtue of our assumption that $g_m$ and $w_m$ are statistically independent. Clearly, the parameter $\kappa_0^m$ controls the location of $p(\kappa_{2m-1})$, whereas, $\tau_m$ controls both location and scale. Moreover, the parameter $\lambda_m$ controls the scale of $p(\kappa_{2m-1})$, without affecting its location. We illustrate this behavior in Figure S-1.1.

Let $\kappa_{\text{min}}, \kappa_{\text{max}}$ and $\kappa_{\text{avg}}$ be the minimum, maximum, and average values of a forward log-rate constant $\kappa$.

Our objective is to determine the hyperparameters of the prior density $p(\kappa)$, given by (S-1.31), so that $\text{E}[\kappa] = \kappa_{\text{avg}}$ and $p(\kappa) \simeq 0$, for $\kappa \leq \kappa_{\text{min}}, \kappa \geq \kappa_{\text{max}}$.

Since $\text{E}[\kappa] = \kappa_{\text{avg}}$, we must have

$$\kappa_0^m - \tau = \kappa_{\text{avg}},$$

(S-1.39)

by virtue of (S-1.38). Moreover, since we want $p(\kappa) \simeq 0$, for $\kappa \leq \kappa_{\text{min}}$, we can impose the condition

$$e^{(\kappa - \kappa_0^m)/\tau} \leq \tau \epsilon e^{-\lambda^2/2}, \quad \text{for} \quad \kappa \leq \kappa_{\text{min}},$$

for a sufficiently small $\epsilon > 0$, which implies that $p(\kappa) < \epsilon$, for $\kappa \leq \kappa_{\text{min}}$, by virtue of (S-1.31) and the facts that $\tau > 1$ and $\text{erfc}(x) \leq 2$. Since $e^{(\kappa - \kappa_0^m)/\tau}$ is monotonically increasing in $\kappa$, it suffices to set

$$e^{(\kappa_{\text{min}} - \kappa_0^m)/\tau} = \tau \epsilon^*,$$

where

$$\epsilon^* := \epsilon e^{-\lambda^2/2},$$

(S-1.40)

which, together with (S-1.39), implies that

$$\tau \ln \tau + (1 + \ln \epsilon^*)\tau = \kappa_{\text{min}} - \kappa_{\text{avg}}.$$  

(S-1.41)
Figure S-1.1: The prior density \( p(\kappa_{2m-1}) \) of the forward log-rate constant, given by (S-1.31), when: (a) \( \kappa_0^m = -2, \tau_m = 2, \lambda_m = 0.6 \), (b) \( \kappa_0^m = -2, \tau_m = 2, \lambda_m = 3 \), (c) \( \kappa_0^m = -8, \tau_m = 2, \lambda_m = 0.6 \), and (d) \( \kappa_0^m = -2, \tau_m = 6, \lambda_m = 3 \). Note that \( \kappa_0^m \) controls the location of \( p(\kappa_{2m-1}) \), whereas, \( \tau_m \) controls both location and scale. Moreover, the parameter \( \lambda_m \) controls the scale of \( p(\kappa_{2m-1}) \), without affecting its location.

Finally, since we want \( p(\kappa) \simeq 0 \), for \( \kappa \geq \kappa_{\text{max}} \), we can impose the condition

\[
\text{erfc} \left[ \frac{1}{\sqrt{2}} \left( \frac{\lambda}{\tau} + \frac{\kappa - \kappa_0^m}{\lambda} \right) \right] \leq 2\tau e^{-\lambda^2/2} e^{-(\kappa - \kappa_0^m)/\tau}, \quad \text{for } \kappa \geq \kappa_{\text{max}},
\]

which implies that \( p(\kappa) \leq \epsilon \), for \( \kappa \geq \kappa_{\text{max}} \), by virtue of (S-1.31) and the fact that \( \tau > 1 \). Since \( \text{erfc}[\left( \frac{\lambda}{\tau} + \frac{\kappa - \kappa_0^m}{\lambda} \right)/\sqrt{2}] \) is monotonically decreasing in \( \kappa \), it suffices to set

\[
\text{erfc} \left[ \frac{1}{\sqrt{2}} \left( \frac{\lambda}{\tau} + \frac{\kappa_{\text{max}} - \kappa_0^m}{\lambda} \right) \right] = 2\tau e^{-\lambda^2/2} e^{-(\kappa - \kappa_0^m)/\tau},
\]

16
which, together with (S-1.39), implies
\[ \tau + \lambda \left( r \sqrt{2} - \frac{\lambda}{\tau} \right) = \kappa^{\text{max}} - \kappa^{\text{avg}}, \]  
(S-1.42)

where
\[ r := \text{erfc}^{-1} \left[ 2 \tau \epsilon^* e^{-\left( \kappa^{\text{max}} - \kappa^{\text{avg}} - \tau \right) / \tau} \right]. \]

Given \( \kappa^{\text{min}} \), \( \kappa^{\text{max}} \), and \( \kappa^{\text{avg}} \), we may be able to determine the values of the hyperparameters \( \kappa^0 \), \( \tau \), and \( \lambda \) by simultaneously solving (S-1.39), (S-1.41), and (S-1.42). Unfortunately, (S-1.41) and (S-1.42) are nonlinear and they both depend on \( \tau \) and \( \lambda \). Hence, finding a solution to these equations is a rather difficult problem [3]. In the following, we discuss a simple approach for determining the values of the hyperparameters, which works quite well.

Equation (S-1.41) depends on \( \lambda \) only through \( \epsilon^* \) [see (S-1.40)]. We can remove this dependence by setting \( \epsilon^* \) to a sufficient small fixed value, such as 0.001. Then, we are left with a nonlinear equation for \( \tau \), which we can solve by employing an appropriate numerical method, such as Newton-Raphson [3]. Note that we must choose \( \epsilon^* \) so that the resulting value of \( \tau \) is greater than one.

Using the value of \( \tau \) calculated in the previous step, we can calculate the value of \( \lambda \) by solving (S-1.42), in which case
\[ \lambda = \frac{\sqrt{2} r \tau}{2} \pm \frac{1}{2} \sqrt{2(r^2 + 2) \tau^2 - 4 \tau (\kappa^{\text{max}} - \kappa^{\text{avg}})}. \]  
(S-1.43)

If \( \tau < 2(\kappa^{\text{max}} - \kappa^{\text{avg}})/(r^2 + 2) \), we have no real-valued solution for \( \lambda \). This indicates that we cannot find an appropriate prior density that satisfies the required specifications (i.e., \( E[\kappa] = \kappa^{\text{avg}} \) and \( p(\kappa) \approx 0 \), for \( \kappa \leq \kappa^{\text{min}}, \kappa \geq \kappa^{\text{max}} \)). On the other hand, (S-1.43) may produce two different nonnegative values for \( \lambda \). In this case, we can use both values of \( \lambda \) to evaluate the corresponding prior densities \( p(\kappa) \). Then, we can pick the value that leads to a prior density that best satisfies the condition \( p(\kappa^{\text{min}}) = p(\kappa^{\text{max}}) = 0 \).

After calculating \( \lambda \) by (S-1.43), we must use \( \epsilon^* \) and (S-1.40) to determine the value of \( \epsilon \). If the resulting value is sufficiently close to zero, then we can accept the hyperparameter values. Otherwise, we must decrease \( \epsilon^* \) [note from (S-1.40) that \( \epsilon \geq \epsilon^* \)] and repeat the previous procedure. Finally, having computed the values of \( \tau \) and \( \lambda \), we can calculate the value of \( \kappa^0 \) by setting
\[ \kappa^0 = \kappa^{\text{avg}} + \tau. \]  
(S-1.44)

17
As an example, let us take $\kappa_{\text{min}} = -17$, $\kappa_{\text{max}} = -3$, and $\kappa_{\text{avg}} = -7$ (these are values we consider in the numerical example), and set $\epsilon^* = 0.001$. Then, (S-1.41) becomes $\tau \ln \tau - 5.9\tau = -10$, which is satisfied with $\tau = 1.8990$. Subsequently, (S-1.43) results in $\lambda = 0.7409$ or $\lambda \simeq 5.3849$, whereas, (S-1.44) gives $\kappa^0 = -5.1010$. The resulting prior density with $\lambda = 0.7409$ is depicted in Figure S-1.2(a). In this case, $\epsilon = 0.0013$, which is sufficiently close to zero. The resulting prior density with $\lambda = 5.3849$ is depicted in Figure S-1.2(b). Clearly, this prior is not acceptable, since it turns out that $p(\kappa_{\text{min}}) = p(-17) \not\approx 0$ and $p(\kappa_{\text{max}}) = p(-3) \not\approx 0$.

When the average value, avg, and the spread, sd, of the variance $\sigma^2$ of the measurement errors are known, we can uniquely determine the hyperparameters $a$, $b$ associated with the prior error density $p(\sigma^2)$, given by Equation (9) in the Main text. This is due to the fact that $\text{E}[\sigma^2] = b/(a - 1)$ and $\text{var}[\sigma^2] = \{\text{E}[\sigma^2]\}^2/(a - 2) = b^2/[(a - 1)^2(a - 2)]$, for $a > 2$, which imply that

$$a = 2 + \left(\frac{\text{avg}}{\text{sd}}\right)^2,$$

$$b = \text{avg} \cdot (a - 1).$$

For example, if $\text{avg} = \text{sd} = 0.5$, then $a = 3$ and $b = 1$, whereas, if $\text{avg} = 0.2$ and $\text{sd} = 0.1$, then $a = 6$ and $b = 1$. The resulting prior densities are depicted in Figure S-1.3.
Figure S-1.3: The prior error variance density $p(\sigma^2)$, given by Equation (9) in the Main text, with: (a) $a = 3$, $b = 1$, and (b) $a = 6$, $b = 1$. In the first case, $E[\sigma^2] = 0.5$ and $\text{var}[\sigma^2] = 0.25$, whereas, in the second case, $E[\sigma^2] = 2$ and $\text{var}[\sigma^2] = 0.01$.

**Posterior mode vs. posterior mean**

We have mentioned in the Main text that the posterior mode should be a more preferable estimator of the kinetic parameters of a biochemical reaction system than the posterior mean. To see why this is true, suppose that, so long as $|\kappa_{f,m} - \kappa_{f,m}^{\text{true}}| < \epsilon_m$, for $m = 1, 2, ..., M + M_1$, the parameters $\{\kappa_f, \mathbb{W}_f\}$ reproduce the concentration dynamics of the biochemical reaction system faithfully. Note that a small $\epsilon_m$ corresponds to a rate constant that appreciably affects the concentration dynamics, whereas, a large $\epsilon_m$ corresponds to a rate constant whose value has little or no effect on the dynamics. If $c(\kappa_f, \kappa_f^{\text{true}})$ is the cost of estimating the true log-rate constants $\kappa_f^{\text{true}}$ by $\kappa_f$, then we can set

$$c(\kappa_f, \kappa_f^{\text{true}}) = \begin{cases} 0, & \text{if } |\kappa_{f,m} - \kappa_{f,m}^{\text{true}}| < \epsilon_m, \text{ for } m = 1, 2, ..., M + M_1 \\ 1, & \text{otherwise}. \end{cases} \quad (S-1.45)$$

As a consequence, we would like to find the optimal estimator $\hat{\kappa}_f$ of $\kappa_f^{\text{true}}$ by minimizing the mean posterior cost $E[c(\kappa_f, \kappa_f^{\text{true}}) \mid y]$ with respect to $\kappa_f$. Note that (S-1.45) implies that

$$E[c(\kappa_f, \kappa_f^{\text{true}}) \mid y] = 1 - \Pr(\{|\kappa_{f,m} - \kappa_{f,m}^{\text{true}}| < \epsilon_m, \text{ for } m = 1, 2, ..., M + M_1\} \mid y). \quad (S-1.46)$$

Hence, we can find the optimal estimator $\hat{\kappa}_f$ by maximizing the probability

$$\Pr(\{|\kappa_{f,m} - \kappa_{f,m}^{\text{true}}| < \epsilon_m, \text{ for } m = 1, 2, ..., M + M_1\} \mid y) = \int_{\kappa_{f,1} - \epsilon_1}^{\kappa_{f,1} + \epsilon_1} \cdots \int_{\kappa_{f,M+M_1} - \epsilon_{M+M_1}}^{\kappa_{f,M+M_1} + \epsilon_{M+M_1}} p_W(\kappa' \mid y) d\kappa',$n with respect to $\kappa_f$. 

19
Clearly, the optimal solution is the center of a hypercube with edge lengths \(\{2\epsilon_1, 2\epsilon_2, \ldots, 2\epsilon_{M+M_1}\}\) with the highest probability given by (S-1.46). When all \(\epsilon_m\)’s are small (i.e., when all parameters appreciably affect the system dynamics), the hypercube will be small as well. In this case, \(\hat{\kappa}_f\) will approximately be the point in the parameter space with the highest posterior probability and, therefore, \(\hat{\kappa}_f = \hat{\kappa}_f^{\text{mode}}\).

When a parameter does not appreciably affect the system dynamics, the hypercube grows along the corresponding dimension. In this case, the skewness of the posterior density \(p_W(\kappa_f \mid y)\) may draw the optimal estimator away from the posterior mode along the direction of growth. This however is an acceptable loss of optimality, since the parameter does not appreciably affect the concentration dynamics and finding its optimal value is inconsequential.

References
1. Israel A, Greville TNE: Generalized Inverses: Theory and Applications. New York: Springer-Verlag, 2nd edition 2003.
2. Vlad MO, Ross J: Thermodynamically based constraints for rate coefficients of large biochemical networks. WIREs Syst. Biol. Med. 2009, 1:348–358.
3. Press WH, Teukolsky SA, Vetterling WT, Flannery BP: Numerical Recipes: The Art of Scientific Computing. New York: Cambridge University Press, 3rd edition 2007.
4. Berry RS, Rice SA, Ross J: Physical Chemistry. New York: Oxford University Press, 2nd edition 2000.
5. Gillespie DT: A rigorous derivation of the chemical master equation. Physica A 1992, 188:404–425.
6. Gillespie DT: The chemical Langevin equation. J. Chem. Phys. 2000, 113:297–306.
7. Grönholt T, Annila A: Natural distribution. Math. Biosci. 2007, 210:659–667.
8. Vlad MO, Cerofolini G, Oefner P, Ross J: Random activation energy model and disordered kinetics, from static to dynamic disorder. J. Phys. Chem. B 2005, 109:21241–21257.
9. Papoulis A, Pillai SU: Probability, Random Variables and Stochastic Processes. New York: McGraw Hill, 4th edition 2002.
10. Ederer M, Gilles ED: Thermodynamically feasible kinetic models of reaction networks. Biophys. J. 2007, 92:1846–1857.