S1 Text – Supporting text

S1.1 Text

The Gillespie Algorithm

In the section “Predicting epidemic behavior”, we discussed that to use Eq. (11) to project the expected behavior of an epidemic outcome $Z$ using the accrued observation $Y_i$ at time $t_i$, one needs access to a stochastic simulator to sample from the random variable $Z|\nu_i$ for a given set of parameter values $\theta$. We use the Gillespie algorithm [40] for this purpose. To describe the steps of the Gillespie algorithm, we consider an epidemic model with $K$ compartments (e.g. in the model described in §”Design of the Performance Analysis”, $K = 4$). Let $\nu(t) = (\nu^1(t), \nu^2(t), \ldots, \nu^K(t))$ denote the state of the epidemic at time $t$, where $\nu^k(t)$ denotes the number of individuals in compartment $k \in \{1, 2, \ldots, K\}$ at time $t$.

The state of an epidemic (i.e. $\nu(t)$) may change due to the occurrence of specific events, such as the transmission of the disease to a susceptible or the recovery from the disease. We use $Q$ to denote the number of such events. In the model shown in Fig 3, these epidemic events include Infection, Seeking Treatment, and Recovery (and hence $Q = 3$). Let $\Lambda_q(\nu(t), \theta)$ denote the rate at which the event $q$ is occurring when the epidemic is at state $\nu(t)$ and with parameter values $\theta$ that is a transformation of the parameter $\theta$.

Note that the stochastic model usually uses a volume dependent transformation of the parameter $\theta$ from the ODE model. We do not need this transformation as our ODE description is volume independent.

In the model described in §”Design of the Performance Analysis”, the force of infection is $\theta_1 \nu^S(t) \nu^I(t) / N(t)$, the rate of seeking treatment is $\theta_2 \nu^I(t)$, and the rate of recovery is $\theta_3 \nu^T(t)$, where $\nu^S(t)$, $\nu^I(t)$ and $\nu^T(t)$ denote, respectively, the number of individuals in the compartments Susceptible, Infective, and Treatment (see Fig 3). The Gillespie algorithm proceeds as follows.

Assuming that the epidemic is at state $\nu(t)$ at time $t$:

1. Calculate the sum of rates: $\Lambda_0 = \sum_{q=1}^Q \Lambda_q(\nu(t), \theta)$.
2. Sample the time to the next event: $\tau \leftarrow -\frac{1}{\Lambda_0} \ln u_1$, where $u_1$ is a random number drawn from a uniform distribution $U((0,1])$.
3. Determine the event that just realized: Select the event $\tilde{q}$ such that

   \[ \frac{\sum_{q=1}^{\tilde{q}-1} \Lambda_q(\nu(t), \theta)}{\Lambda_0} < u_2 \leq \frac{\sum_{q=\tilde{q}}^Q \Lambda_q(\nu(t), \theta)}{\Lambda_0} \]

   where $u_2$ is another random number drawn from a uniform distribution $U((0,1])$.
4. Update epidemic state given the realized event $\tilde{r}$.
5. Until a desired time condition is satisfied, increment time $t \leftarrow t + \tau$ and move to Step 1.

Figure SI 1. Pseudo-code of Gillespie algorithm for simulating stochastic compartmental models.

To obtain simulation trajectories displayed in Fig 4 A and Figs S1 Fig A to S8 Fig A,
we use the Gillespie algorithm (Fig SI 1) to simulate the SITR compartmental model shown in Fig 1.

These SITR trajectories are used to calculate the number of new diagnoses in week $i$, which is equal to $T_i + R_i - T_{i-1} - R_{i-1}$. If for a simulated trajectory the conditions on attack rate (30% - 50% for the mild scenario, 50%-70% for the severe scenario and 70%-100% for the extreme scenario) and timing of peak (between 10 and 20) are fulfilled, we keep the trajectory for the respective scenarios or we otherwise we reject it.

S1.2 Text

Equations of the SEITR model

The Susceptible - Exposed - Infected - Treatment - Recovered model contains one more compartment than the SITR model, namely Exposed. Disease transmission can be modeled using the following ODE model:

$$
\begin{align*}
\frac{dx_S(t)}{dt} &= -k_1 x_S(t) \frac{x_I(t)}{N(t)}, \\
\frac{dx_E(t)}{dt} &= k_1 x_S(t) \frac{x_I(t)}{N(t)} - k_2 x_E(t), \\
\frac{dx_I(t)}{dt} &= k_2 x_E(t) - \theta_2 x_I(t), \\
\frac{dx_T(t)}{dt} &= \theta_2 x_I(t) - \theta_3 x_T(t), \\
\frac{dx_R(t)}{dt} &= \theta_3 x_T(t)
\end{align*}
$$

where $k_1 = k_2 = 2\theta_1$ and $\theta_1$ is the disease transmission rate, $\theta_2$ is the rate of seeking treatment while infectious, and $\theta_3$ is the rate of recovering.

In our analysis, we assume that at time $t = 0$ one population member becomes infected. The model initial condition can therefore be defined as $x_0 = (x_S(0), x_E(0), x_I(0), x_T(0), x_R(0)) = (N - 1, 1, 0, 0)$. We furthermore assume a constant population size $N(t) = x_S(t) + x_E(t) + x_I(t) + x_T(t) + x_R(t)$ denote the population size at time $t$. 
S1.3 text

Formulas to calculate integrated relative errors (IRE) for different metrics

The exact form of $f_{M_i|\theta}(m|\theta)$ of the IREs in Eq. (16) depends on the metric of interest.

**$R_0$ and duration of infectiousness:** For $R_0$ or the mean duration of infectiousness, we have $f_{M_i|\theta}(m|\theta) = 1$, as $R_0$ and the mean duration of infectiousness only depend on the parameter vector sampled from the posterior $\pi_i$.

**$R_{eff}$:** For $R_{eff}$ as the performance target, $M_i$ is defined as $M_i = R_0 S_i N$. As $S$ is the only state that the target depends on, the function $f_{M_i|\theta}(m|\theta)$ needs to yield the probability for this state. This is accomplished by integrating over the remaining states:

$$f_{M_i|\theta}(m|\theta) = \int_{\nu^{(S)}, \nu^{(T)}, \nu^{(R)}(m)} \Pi_i \left( \left( \nu^{(I)}, \nu^{(T)}, \nu^{(R)} \right) | Y_i; \theta \right)$$

**Infection prevalence:** If the target is the infection prevalence: $M_i = \nu_i^{(S)}$, the function $f_{M_i|\theta}(m|\theta)$ needs to yield the probability for this state and, hence, is calculated by integrating over the remaining states.

$$f_{M_i|\theta}(m|\theta) = \int_{\nu^{(S)}, \nu^{(T)}, \nu^{(R)}(m)} \Pi_i \left( \left( \nu^{(S)}, \nu^{(T)}, \nu^{(R)} \right) | Y_i; \theta \right)$$

As the I.Poi benchmark does not calculate a belief state for time $t_i$, we use the simulation model to carry out simulations with parameters according to $\pi_i(\theta|Y_i)$ and consider these simulations as a sample from the belief state $\Pi_i$.

**All Predictions:** In case of prediction for one week, three weeks (cumulative and specific) or the attack rate, it holds

$$f_{M_i|\theta}(m|\theta) = P(m|Y_i)$$

with $P$ as in Eq. (9).

**Median integrated relative error (mIRE):** Next, we will calculate a median relative error (mIRE) over the 50 trajectories as

$$mIRE(M_i) = \text{median} \left( IRE(M_i^{(1)}), \ldots, IRE(M_i^{(50)}) \right).$$

S1.4 text

**Computational Effort**

As mentioned in the Discussion, the most computationally challenging parts are the ODE system integration for MSS, PF and EnKF and the stochastic simulation for I.Poi. We analyzed the computational time for the severe scenario with a population size of 10000. Carrying out the 1000 stochastic simulations for benchmark I.Poi for the 1000 parameters takes about 100 minutes. The additional time for evaluating the likelihood approximation is rather short with about only 1 second, see table 1 for numbers.

The computational times are shorter before the peak and longer after the peak, when less or more data is taken into account. The computational time will only vary for
the I.Poi. when changing the population size as the other methods do not require stochastic simulations.

We note that the implementation was not carried out in a way to maximize speed, so there is opportunity for additional speed up with more efficient implementation. However, these numbers demonstrate that all the algorithms can be used on a personal computer in real time.

Table 1. Computational time for MSS and benchmarks:

| time to peak     | MSS   | I.Poi. | PF       | EnKF    |
|------------------|-------|--------|----------|---------|
| 8 weeks to peak  | 85.2 ± 32.8 | 100* min + 0.6 ± 0.3 | 7.7 ± 4.1 | 10.6 ± 5.9 |
| 4 weeks to peak  | 125.2 ± 33.2 | 100* min + 1 ± 0.3 | 12.5 ± 4.1 | 17.8 ± 6.0 |
|                   | 157.6 ± 31.6 | 100* min + 1.4 ± 0.3 | 17.4 ± 4.0 | 26.4 ± 6.3 |
| 4 weeks after peak| 196.2 ± 33.8 | 100* min + 1.7 ± 0.3 | 22.6 ± 4.0 | 34.5 ± 6.6 |
| 8 weeks after peak| 228.8 ± 35. | 100* min + 2.1 ± 0.3 | 27.5 ± 4.1 | 41.8 ± 6.3 |

Computational time in seconds for the calibration of one time course in average for the severe scenario in the population with 10000 individuals.
* The 100 min are the computational time to carried out 1000 stochastic simulations for each parameter with the Gillespie algorithm.
S1.5 text

Pseudo code for the benchmark method A (I.Poi)

1. Initialization
   (a) Choose an initial prior probability function \( \pi_0(\theta) \).
   (b) Choose the set \( \Theta \) that includes the values of parameters \( \theta \) for which the approximate likelihood function (10) should be calculated.
   (c) Set \( \ln L^{(I.Poi)}(Y_0; \theta) \leftarrow 0 \) for every \( \theta \in \Theta \).

2. Calibration: For each observation \( y_i \), \( i = 1, \ldots, n \),
   (a) Assume \( P^{(I.Poi)}(\cdot|\theta) \sim \text{Poisson}(\mu_i) \) and use simulated trajectories to estimate the mean \( \mu_i \), for every \( \theta \in \Theta \).
   (b) Update the likelihood function: \( L^{(I.Poi)}(Y_i; \theta) \leftarrow L^{(I.Poi)}(Y_{i-1}; \theta) \times P^{(I.Poi)}(y_i|\theta) \) for every \( \theta \in \Theta \).
   (c) Update the parameter posterior distribution: \( \pi_i(\theta|Y_i) \leftarrow L^{(I.Poi)}(y_i|\theta) \pi_{i-1}(\theta|Y_{i-1}) \) for every \( \theta \in \Theta \).

Figure S1 2. Pseudo code for the benchmark method A (I.Poi) [23].
S1.6 text

Pseudo code for the Particle Filter

1. Initialization
   (a) Choose an initial prior probability function \( \pi_0(\theta) \).
   (b) Choose the set \( \Theta \) that includes the values of parameters \( \theta \) for which the likelihood function (1) should be calculated with the specifications in Eq. (12) and name the specific likelihood approximation \( L^{(PF)} \).
   (c) Set \( \ln L^{(PF)}(Y_0; \theta) \leftarrow 0 \) for every \( \theta \in \Theta \).

2. Calibration: For each observation \( y_i, i = 1, \ldots, n \),
   (a) Calculate the probability \( P^{(PF)}(y_i|Y_{i-1}; \theta) \) for every \( \theta \in \Theta \).
   (b) Update the likelihood function: 
      \[ L^{(PF)}(Y_i; \theta) \leftarrow L^{(PF)}(Y_{i-1}; \theta) \times P^{(PF)}(y_i|Y_{i-1}; \theta) \] 
      for every \( \theta \in \Theta \).
   (c) Update the parameter posterior distribution: 
      \[ \pi_i(\theta|Y_i) \leftarrow L^{(PF)}(y_i|\theta) \pi_{i-1}(\theta, Y_{i-1}) \] 
      for every \( \theta \in \Theta \).

Figure SI 3. Pseudo code for the Particle Filter [27]
S1.7 text

Pseudo code for the Ensemble Kalman filter

1. Initialization
   (a) Choose an initial prior probability function \( \pi_0(\theta) \).
   (b) Choose the set \( \Theta \) that includes the values of parameters \( \theta \) for which the likelihood function (1) should be calculated with the specifications in Eq. (12).
   (c) Choose an initial set of particles \( \Psi \). Its elements \( \psi \in \Psi \) contain values for the parameters and states: \( \psi = (\theta, \nu) \).

2. Calibration: For each observation \( y_i, i = 1, \ldots, n \),
   (a) Propagate each particle by setting \( \psi \leftarrow (\theta, \nu_{prior}) \) with \( \nu_{prior} = x(t_i - t_{i-1}, \nu; \theta) \).
   (b) Calculate the prior for the observations: \( \psi(y) = \nu^{(T)}_{prior} + \nu^{(R)}_{prior} - \nu^{(T)} - \nu^{(R)} \).
   (c) Calculate the variance \( \sigma_{\text{obs},i}^2 \) as in Eq. (13).
   (d) Calculate the prior variance of the observed quantity as \( \sigma_{\text{prior}}^2 = \text{Variance}(\psi(y)) \).
   (e) Calculate the prior co-variance of the observed quantity and each unobserved component of the particle \( \psi \) as \( \sigma_m = \text{co-variance}(\psi(y), \psi^{(m)}) \), where \( m \) denotes the unobserved component.
   (f) Update \( \psi \leftarrow \frac{\sigma_m}{\sigma_{\text{prior}}^2} \delta \) with
      \[
      \delta = \frac{\sigma_{\text{obs}}^2}{\sigma_{\text{obs}}^2 + \sigma_{\text{prior}}^2} \psi^{(y)} + \frac{\sigma_{\text{prior}}^2}{\sigma_{\text{obs}}^2 + \sigma_{\text{prior}}^2} \tilde{y}_i
      \]
      and \( \tilde{y}_i = y_i + \xi_i \), with \( \xi_i \sim N(0, \sigma_{\text{obs},i}^2) \);
   (g) Set \((\theta, \nu) \leftarrow \psi\) for the next iteration.

Figure S1 4. Pseudo code for the Ensemble Kalman filter [27].
Detailed pseudo code for our MSS for SITR

1. Input:
   - ODE solution \(x(t, x_0; \theta)\) for an ODE \(x'\) integrated for time \(t\) with initial value \(x_0\) and parameter \(\theta\) such as in equation (5).
   - Data set \(y_1, \ldots, y_n\)
   - Number of ensemble members \(N_{\text{grid}}\).
   - Total populations size \(N\).

2. Initialization:
   
   FOR \(j = 1, \ldots, N_{\text{grid}}\)
   
   \[
   \theta^{(j)} = \left( R_0^{(j)}, \text{mdi}^{(j)}, \gamma^{(j)} \right) \sim U\left( R_0^{\text{range}} \times \text{mdi}^{\text{range}} \times \gamma^{\text{range}} \right) \\
   \]  
   % Draw a parameter sample from an uninformative prior; mdi - mean duration of infectiousness; 
   % \(\gamma\) - additional parameter for the number of weeks, the epidemics started before the first observation 
   % in our case \(R_0^{\text{range}} = [1, 3]\), \(\text{mdi}^{\text{range}} = [1, 20]\), \(\gamma^{\text{range}} = [0, 5]\).
   
   \[
   S_{\gamma}^{(j)} = N - 1 \\
   I_{\gamma}^{(j)} = 1 \\
   T_{\gamma}^{(j)} = R_{\gamma}^{(j)} = 0 \\
   \hat{\nu}_{\gamma}^{(j)} = \left( S_{\gamma}^{(j)}, I_{\gamma}^{(j)}, T_{\gamma}^{(j)}, R_{\gamma}^{(j)} \right) \\
   L^{}(\gamma^{(j)}) = 1 \\
   y_{\gamma} = y_{\gamma+1} = \ldots = y_0 = 0 \\
   \]  
   % auxiliary observations to model the time before the first case

   END FOR

3. Loop:
   
   FOR \(i = 1, \ldots, n\) % epidemic stared at \(-\gamma\) and data available until \(n\)
   
   FOR \(j = 1, \ldots, N_{\text{grid}}\)
   
   % % % Forward integration
   
   \[
   \hat{\nu}_{\gamma}^{(j)} = x(t_i - t_{i-1}, \hat{\nu}_{\gamma}^{(j)}; \theta^{(j)}); \\
   \]  
   % ODE integration
   
   \[
   \text{Cov}^{(j)} = \sum(t_i - t_{i-1}, \hat{\nu}_{\gamma}^{(j)}; \theta^{(j)}); \\
   \]  
   % LNA co-variance integration
   
   \[
   \text{Cov}^{(j)}_{i(k,k)} = \max(\text{Cov}^{(j)}_{i(k,k)}, \alpha) \text{ for } k \in \{S, I\} \\
   \]  
   % ensure that no negative co-variance, we use \(\alpha = 10^{-6}\)

   % % % calculate state estimate in case of observation noise with standard deviation \(\sigma^2\)
   
   % The state estimate is calculated by maximizing the probability in equation 7.
   
   % LNA \(\Rightarrow p\) normal and \(P(y_i|\nu, \hat{\nu}_{\gamma}^{(j)}; \theta)\) normal as gaussian observation noise
   
   % taking into account that \(\hat{\nu}_{\gamma}^{(j)} + \hat{\nu}_{\gamma}^{(j)} - \nu_{\gamma}^{(j)} + \nu_{\gamma}^{(j)}\) would correspond to a noise free observation
   
   % \(\Rightarrow \max_{\nu_i} (2\pi)^{-\dim(\nu_i)} \left( \text{det}(\text{Cov}^{(j)}_{\nu_i}) \right)^{-\frac{1}{2}} \times \exp \left( -\frac{\left(\nu_i - \hat{\nu}_{\gamma}^{(j)}\right)^T \text{Cov}^{(j)}_{\nu_i}^{-1} \left(\nu_i - \hat{\nu}_{\gamma}^{(j)}\right)}{2} \right) \)

   % \(\Rightarrow \max_{\nu_i} \left( \frac{1}{\sigma^2} \exp \left( -\frac{1}{2} \left(\frac{(\nu_i - \hat{\nu}_{\gamma}^{(j)} + \nu_{\gamma}^{(j)} - \nu_{\gamma}^{(j)})^2}{\sigma^2} \right) \right) \right) \)

   % As the 1st, 2nd and 4th factor do not contain \(\hat{\nu}_{\gamma}^{(j)}\) and max is equal to \(\min(-\log(.)\) , define \(h\) as
\[ h\left( \hat{\nu}^{(j)} \right) = \left( \hat{\nu}^{(j)} - \bar{\nu}^{(j)} \right) \cdot \left( \text{Cov}^{(j)} \right)^{-1} \cdot \left( \hat{\nu}^{(j)} - \bar{\nu}^{(j)} \right) + \frac{\left( \hat{\nu}^{(j)} - \bar{\nu}^{(j)} \right) - y_i}{\sigma^2} \]

and minimize by solving derivative equals 0.

\[
\begin{align*}
\frac{\partial h(\nu^{(j)})}{\partial \nu^{(j)}} &= 0 \\
\frac{\partial h(\nu^{(j)})}{\partial \nu^{(j)}} &= 0 \\
\frac{\partial h(\nu^{(j)})}{\partial \nu^{(j)}} &= 0
\end{align*}
\]

% solve the optimization problem analytically for \( \hat{\nu}^{(S)} \), \( \hat{\nu}^{(I)} \) and \( \hat{\nu}^{(T)} \)

% % % calculate state estimate in case of no observation noise
% The state estimate is calculated by maximizing the probability in equation 7.
% LNA \( \Rightarrow \) p normal
% taking into account that \( y_i = \hat{\nu}_i^{(T)} + \hat{\nu}_i^{(S)} - \hat{\nu}_{i-1}^{(T)} + \hat{\nu}_{i-1}^{(S)} \) (noise free observation), reduces one dimension of the optimization problem

\[
\begin{align*}
h(\hat{\nu}_i^{(j)}) &= \left( g(\hat{\nu}_i^{(j)}) - \nu_i^{(j)} \right) \cdot \left( \text{Cov}^{(j)} \right)^{-1} \cdot \left( g(\hat{\nu}_i^{(j)}) - \nu_i^{(j)} \right) \quad \text{% define function } h
\end{align*}
\]

% function g taking into account the relation of \( y_i \)

\[
\begin{align*}
\frac{\partial h(\nu^{(j)})}{\partial \nu^{(j)}} &= 0 \\
\frac{\partial h(\nu^{(j)})}{\partial \nu^{(j)}} &= 0
\end{align*}
\]

% solve the optimization problem analytically for \( \hat{\nu}_i^{(S)} \) and \( \hat{\nu}_i^{(T)} \)

% % % Calculate state specific bounds in case of observation noise with standard deviation \( \sigma^2 \)
% within the bounds of \( \theta \) and population size

\[
\begin{align*}
\hat{\nu}_i^{(S)} &= \min \left\{ \max \left\{ 0, \hat{\nu}_i^{(S)} \right\}, N \right\} \\
\hat{\nu}_i^{(I)} &= \min \left\{ \max \left\{ 0, \hat{\nu}_i^{(I)} \right\}, N \right\} \\
\hat{\nu}_i^{(T)} &= \min \left\{ \max \left\{ 0, \hat{\nu}_i^{(T)} \right\}, N \right\}
\end{align*}
\]

% % % Calculate state specific bounds in case of no observation noise
% upper bounds: population size, population size - already diagnosed cases - current cases,
% population size - all cases; whichever is lowest

\[
\begin{align*}
\hat{\nu}_i^{(S)} &= \min \left( \hat{\nu}_i^{(S)}, N - 1, N - \hat{\nu}_i^{(T)} - \hat{\nu}_{i-1}^{(S)} - y_i - \alpha, N - \sum_{l=1}^{i} y_l - 1 \right) \\
\hat{\nu}_i^{(I)} &= \min \left( \max \left( 0, \hat{\nu}_i^{(I)} \right), \sum_{l=1}^{i} y_l \right) \\
\hat{\nu}_i^{(T)} &= N - \hat{\nu}_i^{(T)} - \hat{\nu}_{i-1}^{(S)} - \hat{\nu}_{i-1}^{(R)} - y_i
\end{align*}
\]

% % % calculate likelihood
% need to integrate over state space \( \rightarrow \) sampling based

FOR \( l = 1, \ldots, n_r \); \( x_i^{(j)} \sim N\left( \mu_i^{(j)}, \text{Cov}^{(j)} \right) \)
END FOR \( n_r = 10000 \)

newlys_{(i,l)}^{(j)} = \bar{z}_{i}^{(j)} - \mathbf{\hat{\nu}}_{i-1}^{(j)} + \left( N - \bar{z}_{i}^{(S)} - \bar{z}_{i}^{(T)} - \bar{z}_{i}^{(R)} \right) - \hat{\nu}^{(R)}_{i-1}
\]

\[
\text{dist} = N_{\text{truncated}(10^{-10}, N)} \left( \text{mean}\left( \text{newlys}_{(i,l=1,\ldots,n_r)}^{(j)} \right), \text{Variance}\left( \text{newlys}_{(i,l=1,\ldots,n_r)}^{(j)} \right) + \sigma^2 \right)
\]

\[
L(y_i|\theta^{(j)}) = PDF(\text{dist}, y_i)
\]

END FOR

\[
L^{(\text{MSS})}(Y_i|\theta^{(j)}) = L^{(\text{MSS})}(Y_{i-1}|\theta^{(j)}) \times L^{(\text{MSS})}(y_i|\theta^{(j)})
\]
END FOR

4. **Output:** sample of parameters $\theta^{(j)}$ with corresponding likelihood values $L(Y_n|\theta^{(j)})$ and belief states.
1. **Input:**
   - Stochastic simulation model $H(t, x_0; \theta)$ simulating a trajectory until time $t$ with an initial value $x_0$ and a parameter $\theta$.
   - Data set $y_1, \ldots, y_n$
   - Number of ensemble members $N_{grid}$.
   - Total populations size $N$.
   - Number of simulations being carried out for each parameter value $N_{prior,sim}$.

2. **Initialization:**
   FOR $j=1, \ldots, N_{grid}$
   \[
   \theta^{(j)} = \left( R^{(j)}_0, mdi^{(j)} \right) \sim U(R^{(range)}_0 \times mdi^{(range)}) \quad \% \text{mdi - mean duration of infectiousness}
   \]
   \[
   S^{(j)}_0 = N - 1 \\
   I^{(j)}_0 = 1 \\
   T^{(j)}_0 = R^{(j)}_0 = 0 \\
   \nu^{(j)}_0 = \left( S^{(j)}_0, I^{(j)}_0, T^{(j)}_0, R^{(j)}_0 \right)
   \]
   \[
   L(y_0 | \theta^{(j)}) = 1
   \]
   END FOR

3. **Calibration:**
   FOR $i = 1, \ldots, n$
   FOR $j=1, \ldots, N_{grid}$
   FOR $k=1, \ldots, N_{prior,sim}$
   \[
   z^{(j,k)}_{i-1} = H^{(k)}(t_{i-1}, \nu^{(j)}_0; \theta^{(j)}) \quad \% \text{previous state}
   \]
   \[
   z^{(j,k)}_i = H^{(k)}(t_i, \nu^{(j)}_0; \theta^{(j)}) \quad \% \text{current state}
   \]
   \[
   \text{newly}^{(j,k)}_i = z^{(j,k,T)}_i + z^{(j,k,R)}_i - z^{(j,k,T)}_{i-1} - z^{(j,k,R)}_{i-1}
   \]
   END FOR $k$
   \[
   mc = \text{mean} \left( \text{newly}^{(j,1)}, \ldots, \text{newly}^{(j,N_{prior,sim})} \right) \quad \% \text{calculate mean for Poisson distribution}
   \]
   IF $mc > 0$ \% Poisson distribution needs parameter greater 0
   \[
   L(y_i | \theta^{(j)}) = PDF(PoissonDistribution(mean), y_i) \quad \% \text{if so, evaluate}
   \]
   ELSE: IF $y_i = 0$ then 1 ELSE $\tilde{\alpha}$ END IF \% if not, check whether $y_i = 0$, we use $\tilde{\alpha} = 10^{-100}$.
   END IF
   END FOR $j$
   END FOR $i$
   \[
   L^{(1,\text{Poi})}(y_i | \theta^{(j)}) = L^{(1,\text{Poi})}(y_{i-1} | \theta^{(j)}) \times L^{(1,\text{Poi})}(y_i | \theta^{(j)}) \quad \% \text{update likelihood}
   \]
   END FOR

4. **Output:** sample of parameters $\theta$ with corresponding likelihood values $L(Y_i | \theta)$.
**S1.10 Text**

Detailed pseudo code for Particle Filter for SITR

1. **Input:**
   - ODE solution $x(t, x_0; \theta)$ for an ODE $x'$ integrated for time $t$ with initial value $x_0$ and parameter $\theta$ such as in equation (5).
   - Data set $y_1, \ldots, y_n$
   - Number of ensemble members $N_{grid}$.
   - Total populations size $N$.

2. **Initialization:**
   - FOR $j = 1, \ldots, N_{grid}$
     - $\theta^{(j)}_0 = \left( R^{(j)}_{00}, md^{(j)}_{00} \right) \sim U(R_{00}^{(range)} \times md^{(range)})$ % $md$ - mean duration of infectiousness
     - $\%$ in our case $R_{00}^{(range)} = [1, 3], md_{00}^{(range)} = [1, 20], \gamma^{(range)} = [0, 5]$.
     - $S_0^{(j)} \sim U([\alpha N, N])$ % we use $\alpha = 0.9$
     - $I_0^{(j)} \sim U([0, N - S_0^{(j)}])$
     - $T_0^{(j)} \sim U([0, N - S_0^{(j)} - I_0^{(j)}])$
     - $R_0^{(j)} \sim U([0, N - S_0^{(j)} - I_0^{(j)} - T_0^{(j)}])$
     - $\hat{\nu}_0^{(j)} = \left( S_0^{(j)}, I_0^{(j)}, T_0^{(j)}, R_0^{(j)} \right)$
     - $w_0^{(j)} = 1/N_{grid}$
   - END FOR

3. **Loop:**
   - FOR $i = 1, \ldots, n$
     - $W_{i-1} = \sum_{j=1}^{N_{grid}} w_{i-1}^{(j)}$
     - FOR $j = 1, \ldots, N_{grid}$: $\tilde{w}_{i-1}^{(j)} = w_{i-1}^{(j)}/W_i$ END FOR % normalize weights
     - $n_{eff} = 1/\sum_{j=1}^{N_{grid}} (\tilde{w}_{i-1}^{(j)})^2$
     - IF $n_{eff} < N_{grid}/2$ % check whether to re-sample
       - $n_1 = 7$
       - $C_1 = \Pi^{n_1/2}/\Gamma(n_1/2 + 1)$ % with the gamma function $\Gamma$
       - $A_1 = (8/C_1(n_1 + 4) * (2\sqrt{\Pi})^{n_1})^{-1/(n_1+4)}$
       - $h_{opt} = 2 * A_1 * N_{grid}^{1/(n_1+4)}$
       - $x_K = (-1, -0.999, \ldots, 0.999, 1)$
       - $d_K = ((n_1 + 2)/2)/C_1 * (1 - x_K^2)$ % kernel
       - $c_K = d_K / \sum_{j=1}^{\text{length}(d_K)} d_{Kj}$
     - END FOR % cumulative
     - f: function interpolating $(x_K, c_K)$
     - $\hat{\psi}_{i-1}^{(j)}$ drawn from $\{ \left( \theta_{i-1}^{(1)}, \hat{\tau}_{i-1}^{(1)} \right), \ldots, \left( \theta_{i-1}^{(N_{grid})}, \hat{\tau}_{i-1}^{(N_{grid})} \right) \}$ with weights $\{ \tilde{w}_{i-1}^{(1)}, \ldots, \tilde{w}_{i-1}^{(N_{grid})} \}$
     - $\tilde{w}_{i-1}^{(j)} = 1/N_{grid}, j = 1, \ldots, N_{grid}$
\[
\sigma_{i-1} = \text{StdDev} \left( \psi_{i-1}, \ldots, \psi_{i-1}^{(N_{\text{grid}})} \right) \quad \% \text{vector of standard deviations}
\]

IF length \( \left\{ \psi_{i-1}^{(j)} \right\}_{j=1}^{N_{\text{grid}}} \) < max(\( \alpha_1, 0.01N_{\text{grid}} \)) \% we use \( \alpha_1 = 20 \).

\[
\sigma_{i-1} = \text{StdDev} \left( \psi_{i-1}^{(\max(i-2,1,0))}, \ldots, \psi_{i-1}^{(N_{\text{grid}})} \right)
\]

\[
\sigma_{i-1}^{(j)} = \max \left( \sigma_{i-1}^{(j)}, \frac{y_i}{5} \right)
\]

END IF

FOR \( k = 1, \ldots, n : \)

\[
z_k = f(\Xi) \text{ with a } \Xi \sim \text{UniformDistribution} \left( [0, 1]^{N_{\text{grid}}} \right)
\]

END FOR

\[
\psi_{i-1}^{(j)} = \psi_{i-1}^{(j)} + h_{\text{opt}} \times \text{SD} \% \text{with SD a diagonal matrix with entries } \sigma_{i-1} \text{ regularization noise}
\]

FOR \( j = 1, \ldots, N_{\text{grid}} \) \% check boundaries

\[
\hat{\nu}_{i-1}^{(j,m)} = \max \left( 0, \min \left( N, \psi_{i-1}^{(j,m)} \right) \right), m \in \{\text{S, I, T, R} \}
\]

\[
\theta_{i-1}^{(j,R_0)} = \max \left( 1, \min \left( R_0(\text{range}, \text{up}), \psi_{i-1}^{(j,R_0)} \right) \right)
\]

\[
\theta_{i-1}^{(j,mdi)} = \max \left( 1, \min \left( mdi(\text{range}, \text{up}), \psi_{i-1}^{(j,mdi)} \right) \right),
\]

END FOR
ELSE nothing \% do not change \( \theta \) or \( \hat{\nu} \).
END IF

\% % % Forward Propagation

FOR \( j = 1, \ldots, N_{\text{grid}} \)

\[
\hat{\nu}_i^{(j)} = x(t_i - t_{i-1}, \hat{\nu}_{i-1}^{(j)}, \theta_{i-1}^{(j)}), \quad \% \text{forward propagation}
\]

\[
\theta_{i-1}^{(j)} = \theta_{i-1}^{(j)}
\]

END FOR

\% % % Assign weights

FOR \( j = 1, \ldots, N_{\text{grid}} \)

\[
y_{i,\text{prior}}^{(j)} = \hat{\nu}_i^{(j,T)} + \hat{\nu}_i^{(j,R)} - \hat{\nu}_i^{(j,T)} - \hat{\nu}_i^{(j,R)} \% \text{prior observations}
\]

\[
\sigma_{i,\text{obs}}^2 = 10000 + \frac{1}{5} \sum_{j=1}^{i-1} y_j \% \text{observation variance}
\]

\[
w_i^{(j)} = w_{i-1}^{(j)} \times \text{PDF} \left( N \left( y_{i,\text{prior}}, \sigma_{i,\text{obs}}^2 \right), y_i \right)
\]

END FOR

4. Output: posterior sample of parameters \( \theta_n \) and states \( \hat{\nu}_n \).
1. Input:
- ODE solution $x(t, x_0; \theta)$ for an ODE $x'$ integrated until time $t$ with initial value $x_0$ and parameter $\theta$ such as in equation (5).
- Data set $y_1, \ldots, y_n$
- Number of ensemble members $N_{grid}$
- Total populations size $N$.

2. Initialization:
FOR $j=1, \ldots, N_{grid}$

\[ \theta_0^{(j)} = \left( R_0^{(j)}, mdi_0^{(j)} \right) \sim U\left(R_0^{\text{range}} \times mdi^{\text{range}}\right) \]
\% in our case $R_0^{\text{range}} = [1, 3]$, $mdi^{\text{range}} = [1, 20]$, $\gamma^{\text{range}} = [0, 5]$.

\[ S_0^{(j)} \sim U \left( [\bar{N}N, N] \right) \]
\% we use $\bar{N} = 0.9$.

\[ I_0^{(j)} \sim U \left( [0, N - S_0^{(j)}] \right) \]
\[ T_0^{(j)} \sim U \left( [0, N - S_0^{(j)} - I_0^{(j)}] \right) \]
\[ R_0^{(j)} \sim U \left( [0, N - S_0^{(j)} - I_0^{(j)} - T_0^{(j)}] \right) \]
\[ \hat{\nu}_0^{(j)} = \left( S_0^{(j)}, I_0^{(j)}, T_0^{(j)}, R_0^{(j)} \right) \]
\[ \psi_0^{(j)} = \left( \hat{\nu}_0^{(j)}, \theta_0^{(j)}, y_0 = 0 \right) \]
END FOR

3. Loop:
FOR $i=1, \ldots, n$
\% Forward propagation
FOR $j=1, \ldots, N_{grid}$
\[ \nu^{(j)}_{i, \text{prior}} = x(t_i - t_{i-1}, \nu^{(j)}_{i-1, \text{prior}}, \theta^{(j)}) \]
\% propagation
END FOR
FOR $j=1, \ldots, N_{grid}$
\[ \nu^{(j)}_{i, \text{prior}} = \alpha' \left( \nu^{(1)}_{i, \text{prior}} - \text{mean} \left( \nu^{(1)}_{i, \text{prior}}, \ldots, \nu^{(N_{grid})}_{i, \text{prior}} \right) \right) + \nu^{(j)}_{i, \text{prior}} \]
\% Inflation with an inflation parameter $\alpha' = 1.07$
\[ \psi^{(y)}_i = \nu^{(y)}_i + \nu^{(y)}_{i, \text{prior}} - \hat{\nu}^{(y)}_{i-1} - \hat{\nu}^{(y)}_{i-1} \]
\% prior observations
\[ \sigma^2_{i, \text{obs}} = 10000 + \frac{1}{50} \left( \frac{1}{3} \sum_{j=1}^{i-3} y_j \right)^2 \]
\% observation variance
\[ \sigma^2_{i, \text{prior}} = \text{variance} \left( \psi^{(1)}_i, \ldots, \psi^{(N_{grid}, y)}_i \right) \]
\% prior variance
\[ \sigma^{(m)}_i = \text{co-variance} \left( \psi^{(y)}_i, \psi^{(m)}_i \right), m \in \{ R_0, mdi, S, I, T, R \} \]
\% co-variance observation prior
\[ \hat{y}^{(j)}_i = \max \left( 0, y_i + \xi^{(j)}_i \right) \]
\% noisy observation
with $\xi^{(j)}_i \sim \mathcal{N}(0, \sigma^2_{i, \text{obs}})$
END FOR
\% Update
FOR $j=1, \ldots, N_{grid}$
\[ \delta^{(j)}_i = \frac{\sigma^2_{i, \text{obs}}}{\sigma^2_{i, \text{obs}} + \sigma^2_{i, \text{prior}}} \psi^{(y,j)}_i + \frac{\sigma^2_{i, \text{prior}}}{\sigma^2_{i, \text{obs}} + \sigma^2_{i, \text{prior}}} \hat{y}^{(j)}_i. \]
\[ \theta^{(j,m)} = \frac{\theta^{(j)}_{\text{prior}}}{\sigma^{(j)}_{\text{prior}}} \delta^{(j)}_{i} + \hat{\theta}^{(j,m)}_{i}, m \in \{R_0,mdi\} \]

\[ \nu^{(j,m)} = \frac{\nu^{(j)}_{\text{prior}}}{\sigma^{(j)}_{\text{prior}}} \delta^{(j)}_{i} + \hat{\nu}^{(j,m)}_{i-1}, m \in \{S,I,T,R\} \]

END FOR

\% \% Check boundaries

FOR \( j=1,\ldots,N_{\text{grid}} \)

\[ \hat{\nu}^{(j,S)}_{i} = \text{IF} \ \hat{\nu}^{(j,S)}_{i} > N : N - 1 \ \text{ELSE} \ \hat{\nu}^{(j,S)}_{i} \ \text{END IF} \]

\[ \hat{\nu}^{(j,m)}_{i} = \text{IF} \ \hat{\nu}^{(j,m)}_{i} > N : \text{median} \left( \hat{\nu}^{(j=1,\ldots,N_{\text{grid}},m)}_{i} \right) \ \text{ELSE} \ \hat{\nu}^{(j,m)}_{i}, m \in \{I,T,R\} \ \text{END IF} \]

\[ \hat{\theta}^{(j,R_0)}_{i} = \text{IF} \ \hat{\theta}^{(j,R_0)}_{i} > R_0^{(\text{range,up})} : \min \left( R_0^{(\text{range,up})}, \text{median} \left( \hat{\theta}^{(j=1,\ldots,N_{\text{grid}},R_0)}_{i} \right) \right) \ \text{ELSE} \ \hat{\theta}^{(j,R_0)}_{i} \ \text{END IF} \]

\[ \hat{\theta}^{(j,mdi)}_{i} = \text{IF} \hat{\theta}^{(j,mdi)}_{i} > mdi^{(\text{range,up})} : \min \left( mdi^{(\text{range,up})}, \text{median} \left( \theta^{(j=1,\ldots,N_{\text{grid}},mdi)}_{i} \right) \right) \ \text{ELSE} \ \hat{\theta}^{(j,mdi)}_{i} \ \text{END IF} \]

END IF

\[ \hat{\nu}^{(j,m)}_{i} = \text{IF} \ \hat{\nu}^{(j,m)}_{i} < 0 : \max \left( 1, \text{mean} \left( \hat{\nu}^{(j=1,\ldots,N_{\text{grid}},m)}_{i} \right) \right) \ \text{ELSE} \ \hat{\nu}^{(j,m)}_{i}, m \in \{S,I,T,R\} \ \text{END IF} \]

\[ \hat{\theta}^{(j,m)}_{i} = \text{IF} \ \hat{\theta}^{(j,m)}_{i} < 0 : \max \left( 1, \text{mean} \left( \theta^{(j=1,\ldots,N_{\text{grid}},m)}_{i} \right) \right) \ \text{ELSE} \ \hat{\theta}^{(j,m)}_{i}, m \in \{R_0,mdi\} \ \text{END IF} \]

END FOR

END FOR

4. Output: a posteriori sample of parameters \( \theta_n \) and states \( \hat{v}_n \).

References

1. Lipsitch M, Finelli L, Heffernan RT, Leung GM, Redd; for the 2009 H1N1
   Surveillance Group SC. Improving the evidence base for decision making during a
   pandemic: the example of 2009 influenza A/H1N1. Biosecurity and bioterrorism:
   biodefense strategy, practice, and science. 2011;9(2):89–115.

2. Mills CE, Robins JM, Lipsitch M. Transmissibility of 1918 pandemic influenza.
   Nature. 2004;432(7019):904–906.

3. Cauchemez S, Donnelly CA, Reed C, Ghani AC, Fraser C, Kent CK, et al.
   Household transmission of 2009 pandemic influenza A (H1N1) virus in the United
   States. New England Journal of Medicine. 2009;361(27):2619–2627.

4. Cowling BJ, Lau MS, Ho LM, Chau SK, Tsang T, Liu SH, et al. The effective
   reproduction number of pandemic influenza: prospective estimation.
   Epidemiology. 2010;21(6):842.

5. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to
   estimate time-varying reproduction numbers during epidemics. American journal
   of epidemiology. 2013;178(9):1505–1512.

6. Cauchemez S, Epperson S, Biggerstaff M, Swerdlow D, Finelli L, Ferguson NM.
   Using routine surveillance data to estimate the epidemic potential of emerging
   zoonoses: application to the emergence of US swine origin influenza A H3N2v
   virus. PLoS Med. 2013;10(3):e1001399.

7. Cauchemez S, Carrat F, Viboud C, Valleron A, Boelle P. A Bayesian MCMC
   approach to study transmission of influenza: application to household
   longitudinal data. Statistics in medicine. 2004;23(22):3469–3487.

8. Höhle M, Jørgensen E, O’Neill PD. Inference in disease transmission experiments
   by using stochastic epidemic models. Journal of the Royal Statistical Society: Series
   C (Applied Statistics). 2005;54(2):349–366.
9. White LF, Pagano M. A likelihood-based method for real-time estimation of the
serial interval and reproductive number of an epidemic. Stat Med. 2008
Jul;27(16):2999–3016.

10. O’Neill PD. A tutorial introduction to Bayesian inference for stochastic epidemic
models using Markov chain Monte Carlo methods. Mathematical biosciences.
2002;180(1):103–114.

11. Obadia T, Haneef R, Boelle PY. The R0 package: a toolbox to estimate
reproduction numbers for epidemic outbreaks. BMC Med Inform Decis Mak.
2012;12:147.

12. Cauchemez S, Boelle PY, Donnelly CA, Ferguson NM, Thomas G, Leung GM,
et al. Real-time estimates in early detection of SARS. Emerging Infect Dis. 2006
Jan;12(1):110–113.

13. Cauchemez S, Boëlle PY, Thomas G, Valleron AJ. Estimating in real time the
efficacy of measures to control emerging communicable diseases. American
Journal of Epidemiology. 2006;164(6):591–597.

14. Wallinga J, Tenais P. Different epidemic curves for severe acute respiratory
syndrome reveal similar impacts of control measures. American Journal of
Epidemiology. 2004;160(6):509–516.

15. Davoudi B, Miller JC, Meza R, Meyers LA, Earn DJD, Pourbohloul B. Early
real-time estimation of the basic reproduction number of emerging infectious
diseases. Physical Review X. 2012;2.

16. Cauchemez S, Ferguson NM. Methods to infer transmission risk factors in
complex outbreak data. J R Soc Interface. 2012 Mar;9(68):456–469.

17. Saramaki J, Kaski K. Modelling development of epidemics with dynamic
small-world networks. J Theor Biol. 2005 Jun;234(3):413–421.

18. Anderson RM, May RM. Infectious Diseases of Humans: Dynamics and Control.
Oxford University Press; 1992.

19. Daley DJ, Gani JM. Epidemic Modelling: An Introduction. Cambridge; New
York: Cambridge University Press; 1999.

20. Alkema L, Raftery AE, Clark SJ. Probabilistic projections of HIV prevalence
using Bayesian melding. The Annals of Applied Statistics. 2007;p. 229–248.

21. Elderd BD, Dukic VM, Dwyer G. Uncertainty in predictions of disease spread
and public health responses to bioterrorism and emerging diseases. Proceedings
of the National Academy of Sciences. 2006;103(42):15693–15697.

22. Birrell PJ, Ketsetzis G, Gay NJ, Cooper BS, Presanis AM, Harris RJ, et al.
Bayesian modelling to unmask and predict influenza A/H1N1pdm dynamics in
London. Proceedings of the National Academy of Sciences. 2011;108(45):18238–18243.

23. Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, Hedley AJ, et al.
Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of
public health interventions. Science. 2003;300(5627):1961–1966.

24. Choi B, Rempala GA. Inference for discretely observed stochastic kinetic
networks with applications to epidemic modeling. Biostatistics. 2012
Jan;13(1):153–165.
25. Ionides EL, Breto C, King AA. Inference for nonlinear dynamical systems. Proc Natl Acad Sci USA. 2006 Dec;103(49):18438–18443.

26. Shaman J, Karspeck A, Yang W, Tamerius J, Lipsitch M. Real-time influenza forecasts during the 2012-2013 season. Nat Commun. 2013;4:2837.

27. Yang W, Karspeck A, Shaman J. Comparison of Filtering Methods for the Modeling and Retrospective Forecasting of Influenza Epidemics. PLOS Computational Biology. 2014;10:e1003583.

28. Yang W, Cowling BJ, Lau EH, Shaman J. Forecasting Influenza Epidemics in Hong Kong. PLoS Comput Biol. 2015 Jul;11(7):e1004383.

29. Shaman J, Karspeck A. Forecasting seasonal outbreaks of influenza. Proc Natl Acad Sci USA. 2012 Dec;109(50):20425–20430.

30. Ong JB, Chen MI, Cook AR, Lee HC, Lee VJ, Lin RT, et al. Real-time epidemic monitoring and forecasting of H1N1-2009 using influenza-like illness from general practice and family doctor clinics in Singapore. PLoS ONE. 2010;5(4):e10036.

31. Dukic V, Lopes HF, Polson NG. Tracking Epidemics With Google Flu Trends Data and a State-Space SEIR Model. Journal of the American Statistical Association. 2012;107.

32. Chretien JP, George D, Shaman J, Chitale RA, McKenzie FE. Influenza forecasting in human populations: a scoping review. PLoS ONE. 2014;9(4):e94130.

33. Bettencourt LM, Ribeiro RM. Real time Bayesian estimation of the epidemic potential of emerging infectious diseases. PLoS One. 2008;3(5):e2185.

34. Abbey H. An examination of the Reed-Frost theory of epidemics. Human biology. 1952;24(3):201–233.

35. Zimmer C, Sahle S. Deterministic inference for stochastic systems using multiple shooting and a linear noise approximation for the transition probabilities. IET Systems Biology. 2015;9:181 – 192.

36. Zimmer C. Reconstructing the hidden states in time course data of stochastic models. Mathematical Biosciences. 2015;269:117 – 129.

37. Thomas P, Matuschek H, Grima R. Intrinsic noise analyzer: a software package for the exploration of stochastic biochemical kinetics using the system size expansion. PloS one. 2012;7(6):e38518.

38. Van Kampen NG. Stochastic processes in physics and chemistry. vol. 1. Elsevier; 1992.

39. Yaesoubi R, Cohen T. Generalized Markov models of infectious disease spread: A novel framework for developing dynamic health policies. European Journal of Operational Research. 2011;215:679–687.

40. Gillespie DT. A General Method for Numerically Simulating the Stochastic Time Evolution of coupled Chemical Reactions. Journal of Computational Physics. 1976;22 (4):403–434.

41. Influenza prediction challenge. Center for disease control and prevention. 2016:http://www.cdc.gov/flu/news/flu-forecast-website-launched.htm.
42. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, et al. COPASI - a COMplex PAthway SImulator. Bioinformatics. 2006;22 (24):3067–3074.

43. Mathematica, Version 10.4. Wolfram Research, Inc. 2015;Champaign, IL.

44. Pahle J, Challenger JD, Mendes P, McKane AJ. Biochemical fluctuations, optimisation and the linear noise approximation. BMC Systems Biology. 2012;6.

45. Challenger JD, McKane AJ, Pahle J. Multi-compartment linear noise approximation. Journal of Statistical Mechanics: Theory and Experiment. 2012;P11010.

46. Straube R, von Kamp A. LiNA – A Graphical Matlab Tool for Analyzing Intrinsic Noise in Biochemical Reaction Networks. online tutorial. 2013;http://www2.mpi-magdeburg.mpg.de/projects/LiNA/Tutorial_LiNA_v1.pdf.

47. van Kampen NG. Stochastic processes in physics and chemistry. Elsevier; 2007.

48. Grima R. An effective rate equation approach to reaction kinetics in small volumes: Theory and application to biochemical reactions in nonequilibrium steady-state conditions. The Journal of Chemical Physics. 2010;133:035101.

49. Thomas P, Matuschek H, Grima R. Intrinsic Noise Analyzer: A Software Package for the Exploration of Stochastic Biochemical Kinetics Using the System Size Expansion. Plos ONE. 2012;7:e38518.

50. Komorowski M, Finkenstädt B, Harper CV, Rand DA. Bayesian inference of biochemical kinetic parameters using the linear noise approximation. BMC Bioinformatics. 2009;10:343.