Appendices

A. reference tables

**TABLE I: Table of definitions**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$\min \left( t_{\text{final}}, \left( \min_t I_{{f(x(t</td>
</tr>
<tr>
<td>$N_T$</td>
<td>total number of reactions fired in the interval $[0, T]$</td>
</tr>
<tr>
<td>$\delta(k)$</td>
<td>distance from the target, $\mathcal{P}<em>\mathcal{E} - \hat{p}</em>{SPaSE}(k)$</td>
</tr>
<tr>
<td>$\xi$</td>
<td>intermediate events</td>
</tr>
<tr>
<td>$\rho$</td>
<td>fractions used to compute intermediate events Eqs. (10,11)</td>
</tr>
<tr>
<td>$\eta(k)$</td>
<td>$\sum_{i=1}^{N} \left[ I_{{f(x_i(t</td>
</tr>
<tr>
<td>$\hat{p}_{SPaSE}(k)$</td>
<td>$\frac{1}{N} \eta(k)$</td>
</tr>
</tbody>
</table>

**TABLE II: List of input parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Default value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{E}$</td>
<td>NA</td>
<td>target event</td>
</tr>
<tr>
<td>$f(x(t))$</td>
<td>NA</td>
<td>event (distance) function</td>
</tr>
<tr>
<td>$\mathcal{P}_\mathcal{E}$</td>
<td>NA</td>
<td>target probability for $\mathcal{E}$</td>
</tr>
<tr>
<td>$\epsilon_{\mathcal{P}_\mathcal{E}}$</td>
<td>0.05</td>
<td>absolute error tolerance</td>
</tr>
<tr>
<td>$N$</td>
<td>$5 \cdot 10^4$</td>
<td>number of simulations in computing $\hat{p}_{SPaSE}(k)$</td>
</tr>
<tr>
<td>$\mathcal{L}$</td>
<td>10</td>
<td>maximum number of multilevel CE method per $k^{(0)}$</td>
</tr>
<tr>
<td>$\mathcal{I}$</td>
<td>3</td>
<td>maximum number of exponential interpolation per $k^{(0)}$</td>
</tr>
</tbody>
</table>
### TABLE III: Table of definitions for variables used in Algorithms 1-5

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>(M)</td>
<td>number of reactions in the system</td>
</tr>
<tr>
<td>(\gamma_{\text{hist}})</td>
<td>matrix of past intermediate biasing parameter values normalized with respect to (k^{(0)}). (\gamma_{(i,j)}^{\text{hist}}) denotes (i^{th}) past intermediate biasing parameter for (R_j).</td>
</tr>
<tr>
<td>(\eta_{\text{hist}})</td>
<td>vector of past (\eta) values computed. (\eta_{i}^{\text{hist}}) denotes (i^{th}) past (\eta) value.</td>
</tr>
<tr>
<td>(\gamma_{\text{num}})</td>
<td>matrix whose entries are used to accumulate the numerator in Eq. (13) (\gamma_{(i,j)}^{\text{num}}) denotes values of the matrix at row (i) and column (j).</td>
</tr>
<tr>
<td>(\gamma_{\text{denom}})</td>
<td>matrix whose entries are used to accumulate the denominator in Eq. (13) (\gamma_{(i,j)}^{\text{denom}}) denotes values of the matrix at row (i) and column (j).</td>
</tr>
<tr>
<td>(\bar{k})</td>
<td>matrix of candidate reaction rates from executing Steps 1-3 in Exponential Interpolation of Biasing Parameters (\bar{k}_{(l,s)}) denotes (s^{th}) candidate reaction rate for (R_j) in (l^{th}) interpolation.</td>
</tr>
<tr>
<td>(\bar{\gamma})</td>
<td>matrix of candidate biasing parameter rates from fitting exponential interpolant and executing Step 1. (\bar{\gamma}_{(l,s)}) denotes (s^{th}) candidate biasing parameter for (R_j) in (l^{th}) interpolation.</td>
</tr>
</tbody>
</table>

### B. SPARSE PSEUDOCODE

In the below pseudocode, \(\nu_j\) and \(t_f\) represent the state change vector for reaction \(j\) and the simulation end time, respectively. The maximum number of multilevel cross-entropy method for any given set of initial reaction rates is denoted by \(\mathcal{L}\), and exceeding \(\mathcal{L}\) is considered a failure in convergence. Similarly, the maximum number of exponential interpolation is denoted by \(\mathcal{I}\). The computational complexity of Algorithms 2-4 is identical to that of dwSSA.

**Algorithm 1.** SParSE Driver
Input: $k^{(0)}$, $L$, $E$, $N$, $\mathcal{P}_E$, $\epsilon_{P_E}$

\[ l \leftarrow 0 \]

2: $\phi_{\text{type}} \leftarrow \text{Eq. (8)}$

\[ k^{(\text{cur})} \leftarrow k^{(0)}, \quad \gamma^{(l)} \leftarrow \text{matrix of size } 1 \times M \]

4: $\gamma_{\text{hist}}$, $\eta_{\text{hist}} \leftarrow \emptyset$

doInterpolate, $\text{isOverPerturbed} \leftarrow \text{false}$

6: while $l < L$ do

\[ \text{for } i = 1 \text{ to number of rows in } \gamma^{(l)} \text{ do} \]

8: $\gamma^{(l,i)} \leftarrow i^{th} \text{ row of } \gamma^{(l,:)}$, $k^{(\text{int},i)} \leftarrow \text{Eq. (12)}$

\[ \xi, \eta \left( k^{(\text{int},i)} \right) \leftarrow \text{Algorithm 2} \]

10: $\hat{\mathcal{P}_{\text{ParSE}}} \left( k^{(\text{int},i)} \right) = \frac{1}{N} \cdot \eta \left( k^{(\text{int},i)} \right)$, $\delta \left( k^{(\text{int},i)} \right) = \mathcal{P}_E - \hat{\mathcal{P}_{\text{ParSE}}} \left( k^{(\text{int},i)} \right)$

\[ \text{if } |\delta \left( k^{(\text{int},i)} \right) | < \epsilon_{P_E} \text{ then} \]

12: $k^* \leftarrow k^{(\text{int},i)}$

\[ \text{go to step (90)} \]

14: \text{end if}

\[ \gamma_{\text{hist}} \left| \gamma_{\text{hist}}^{(\text{end}+1,j)} = k_j^{(\text{int},i)} / k_j^{(0)} \right., \quad j = \{1, \ldots, M\} \]

16: $\eta_{\text{hist}} | \eta_{\text{hist}}^{\text{end}+1} = \eta \left( k^{(\text{int},i)} \right)$

\text{if isOverPerturbed == false then}

18: \text{if } \text{sgn} \left( \delta \left( k^{(\text{int},i)} \right) \right) \neq \phi_{\text{type}} \text{ then} \{\text{reached over perturbed state}\}

\[ \text{if } l == \text{number of rows in } \gamma^{(l)} \text{ then} \]

20: \text{if } (\forall \eta \in \eta_{\text{hist}}, \eta > \mathcal{P}_E \cdot N) \text{ then}

\[ \text{isOverPerturbed} \leftarrow \text{true} \]

22: \text{else}

break out of while loop in step (6)

24: \text{end if}

else

26: doInterpolate $\leftarrow \text{true}$, continue for loop in step (7)

end if

28: \text{end if}

if doInterpolate == true then

30: break out of while loop in step (6)

else
32: \( k^{(\text{cur})} \leftarrow k^{(\text{int},i)} \)

break out of for loop in step (7)

34: end if

end if

36: if isOverPerturbed == true then

if \( \text{sgn} \left( \delta \left( k^{(\text{int},i)} \right) \right) == \phi_{\text{type}} \) then \{reached under perturbed state\}

38: if \( l == \) number of rows in \( \gamma^{(l,:)} \) then

break out of while loop in step (6)

40: else

doInterpolate \leftarrow \text{true}, \quad \text{continue for loop in step (7)}

42: end if

else

44: if doInterpolate == true then

break out of while loop in step (6)

46: else

\( k^{(\text{cur})} \leftarrow k^{(\text{int},i)}, \quad \xi \leftarrow \text{Algorithm 3} \)

48: break out of for loop in step (7)

end if

50: end if

end if

52: end for

\( \gamma^{\text{num}}, \gamma^{\text{denom}} \leftarrow \text{Algorithm 4 with } k^{(\text{cur})} \)

54: \( l \leftarrow l + 1 \)

if isOverPerturbed == false then

56: \( \gamma^{(l)}|_{(i,j)}^{(l)} = \frac{\gamma^{\text{num}}_{(i,j)}}{\gamma^{\text{denom}}_{(i,j)}} \)

else \{inverse biasing\}

58: \( \gamma^{(l)}|_{(i,j)}^{(l)} = \frac{\gamma^{\text{denom}}_{(i,j)}}{\gamma^{\text{num}}_{(i,j)}} \)

end if

60: end while

retry \leftarrow 1 \{\text{used to count the number of interpolation stage}\}

62: \( k^{(\text{retry},:)} \), ind\( _k \leftarrow \text{Algorithm 5} \)

while retry \leq I do
\[ k^{\text{(cur)}} \leftarrow \bar{k}^{\text{(retry, ind}_k}\text{)} \]
\[ \xi, \eta \left( k^{\text{(cur)}} \right) \leftarrow \text{Algorithm 2} \]
\[ \hat{p}_{\text{SPSE}} \left( k^{\text{(cur)}} \right) = \frac{1}{N} \cdot \eta \left( k^{\text{(cur)}} \right), \quad \delta \left( k^{\text{(cur)}} \right) = p_{E} - \hat{p}_{\text{SPSE}} \left( k^{\text{(cur)}} \right) \]

if \( \text{ind}_k \) was visited twice before then

if weighted average was tried then

Failure in convergence. Exit Algorithm 1.

else

\[ k^{\text{(cur)}} \leftarrow \text{weighted average between } \bar{k}^{\text{(retry, ind}_k}\text{)} \text{ and the previous } k^{\text{(cur)}} \]

else

if \( |\delta \left( k^{\text{(cur)}} \right)| < \epsilon_{p_E} \) then

\[ k^* \leftarrow k^{\text{(cur)}} \]

end if

end if

\[ \gamma^{\text{hist}} |, \eta^{\text{hist}} |_{\text{end+1, j}} = k_j^{\text{end+1}} / k_j^{(0)}, \quad j = \{1, \ldots, M\} \]
\[ \eta^{\text{hist}} |_{\text{end+1}} = \eta \left( k^{\text{(cur)}} \right) \]

if \( \delta \left( k^{\text{(cur)}} \right) < 0 \text{ and } \text{ind}_\gamma > 1 \) then

\[ \text{ind}_k \leftarrow \text{ind}_k - 1 \]

else if \( \delta \left( \text{ind}_k \right) > 0 \text{ and } \text{ind}_\gamma < \text{number of rows in } \bar{k}^{\text{(retry,.)}} \) then

\[ \text{ind}_k \leftarrow \text{ind}_k + 1 \]

else

\[ \text{retry} \leftarrow \text{retry} + 1 \]

end if

end if

end while

90: return \( k^* \)


Input: \( k, E, N, f(x(t)) \)

\[ v_k, I_\xi \leftarrow 0 \quad \{\text{zero vector of size } N\} \]

2: for \( i = 1 \) to \( N \) do
\( t \leftarrow 0, \quad x \leftarrow x_0 \)

4: evaluate all \( a_j(x), \ j = \{1, \ldots, M\} \); calculate \( a_0(x) \)

\[
v_k[i] \leftarrow f(x_0)
\]

6: while \( t \leq t_f \) do

if \( f(x) == \mathcal{E} \) then

8: \( I_\xi[i] \leftarrow 1 \)

break out of the while loop

10: end if

generate two unit-interval uniform random numbers \( r_1 \) and \( r_2 \)

12: \( \tau \leftarrow a_0^{-1}(x) \ln(1/r_1) \)

\( j' \leftarrow \) smallest integer satisfying \( \sum_{j=1}^{j'} a_j(x) \geq r_2a_0(x) \)

14: \( t \leftarrow t + \tau, \quad x \leftarrow x + \nu_{j'} \)

update all \( a_j(x), \ j = \{1, \ldots, M\} \); recalculate \( a_0(x) \)

16: if \( \phi_{\text{type}} == 1 \) then

18: \( v_k[i] \leftarrow \max(v_k[i], f(x)) \)

else

18: \( v_k[i] \leftarrow \min(v_k[i], f(x)) \)

20: end if

end while

22: end for

\[
\eta(k) \leftarrow \sum_{i=1}^{N} I_\xi[i], \quad \rho(\delta) \leftarrow \text{Eq. (10)}
\]

24: if \( \phi_{\text{type}} == 1 \) then

Sort \( v_k \) in descending order

26: else

Sort \( v_k \) in ascending order

28: end if

\( \xi \leftarrow \) unique elements of \( v_k \) at indices \( \lceil \rho(\delta) \times N \rceil \)

30: return \( \xi \) and \( \eta(k) \)


Input: \( k, \ N \)

\[
v_k \leftarrow \overrightarrow{0} \quad \{\text{zero vector of size} \ N\}\]
2: **for** $i = 1$ to $N$ **do**
   
   $t \leftarrow 0, \quad x \leftarrow x_0$

4: evaluate all $a_j(x), j = \{1, \cdots, M\}$; calculate $a_0(x)$
   
   $v_k[i] \leftarrow f(x_0)$

6: **while** $t \leq t_f$ **do**

   generate two unit-interval uniform random numbers $r_1$ and $r_2$

8: $\tau \leftarrow a_0^{-1}(x) \ln(1/r_1)$
   
   $j' \leftarrow$ smallest integer satisfying $\sum_{j=1}^{j'} a_j(x) \geq r_2a_0(x)$

10: $t \leftarrow t + \tau, \quad x \leftarrow x + \nu_{j'}$

   update all $a_j(x), j = (1, \cdots, M)$; recalculate $a_0(x)$

12: **if** $\phi_{\text{type}} == 1$ **then**

   $v_k[i] \leftarrow \max(v_k[i], f(x))$

14: **else**

   $v_k[i] \leftarrow \min(v_k[i], f(x))$

16: **end if**

**end while**

18: **end for**

   $\rho(\delta) \leftarrow \text{Eq. (11)}$

20: **if** $\phi_{\text{type}} == 1$ **then**

   Sort $v_k$ in descending order

22: **else**

   Sort $v_k$ in ascending order

24: **end if**

   $\xi \leftarrow$ unique elements of $v_k$ at indices $[\rho(\delta) \times N]$

26: **return** $\xi$

**Algorithm 4. Biasing Parameter Computation.**

**Input:** $k, \xi, N$

1: $q \leftarrow$ length of $\xi$

2: $\gamma_{\text{num}}, \gamma_{\text{denom}} \leftarrow 0$ matrix of size $q \times M$

   **for** $i = 1$ to $N$ **do**

4: $\gamma_{\text{numi}}, \gamma_{\text{denomi}} \leftarrow 0$ matrix of size $q \times M$
\textbf{Algorithm 5.} Exponential Interpolation of Biasing Parameters.

\textbf{Input:} $\mathcal{P}_E$, $\epsilon_{\mathcal{P}_E}$, $k^{(0)}$, $\gamma^{\text{hist}}$, $\eta^{\text{hist}}$, $N$

1: $\eta^{\text{interp}} \leftarrow \frac{(-2 - 1 - 0.5 0.5 1 2) \cdot \epsilon_{\mathcal{P}_E} + \mathcal{P}_E \times N}{\mathcal{P}_E \cdot N}$

2: [Optional] keep at most 5 elements in $\eta^{\text{hist}}$ closest to $\mathcal{P}_E \cdot N$ and the corresponding $\gamma^{\text{hist}}$ with at least one element of $\eta^{\text{hist}} < \mathcal{P}_E \cdot N$ and one element $> \mathcal{P}_E \cdot N$

3: for all $j \in 1, \cdots, M$ do

4: solve for $a$ and $b$ in $y = ax + b$, where $x \leftarrow \gamma^{\text{hist}}_{(i,j)}$ and $y \leftarrow \log \left(\eta^{\text{hist}}\right)$
5: \( \tilde{\eta}_{j}^{(s)} \leftarrow (\log(\eta_{\text{interp}}^{s}) - b)/a, \ s = \{1, \cdots, 7\} \)

6: apply Steps 2-3 in Exponential Interpolation of Biasing Parameters and compute \( \tilde{k}_{j}^{(s)} \)

7: end for

8: remove duplicate sets in \( \tilde{k}_{j}^{(s)} \) if applicable

9: \( \text{ind}_{\tilde{k}} \leftarrow \text{index where} \ \eta_{\text{hist}}^{\text{ind}_{\tilde{k}}} == P_{\varepsilon} \cdot N \)

10: return \( \tilde{k}, \ \text{ind}_{\tilde{k}} \)

C. FAILURES IN CONVERGENCE

Results and Discussion section contains a total of 330 test cases of SParSE among three example systems provided: birth-death process, reversible isomerization process, and SIRS disease transmission model. Out of the 330 cases, we observed three failures in convergence: two in birth-death process and one in reversible isomerization process. For birth-death process with \( P_{\varepsilon} = 0.60 \) and \( \epsilon_{P_{\varepsilon}} = 0.01 \), two samples, \( k_{3}^{(0)} = [1.606 \ 0.0140] \) and \( k_{27}^{(0)} = [1.684 \ 0.0148] \) (subscript representing the index of initial reaction rates), failed to converge after three rounds of exponential interpolations. One failure in reversible isomerization is caused by \( k_{27}^{(0)} = [0.205 \ 0.414] \) with \( P_{\varepsilon} = 0.95 \) and \( \epsilon_{P_{\varepsilon}} = 0.005 \). We note that the following discussion involves variables and procedures from Methods section.

For all three cases, the failure to converge was a result of poor agreement between the past simulation data and the corresponding interpolants. Figure C.1 displays the exponential interpolants from the first iteration for \( k_{3}^{(0)} \) and \( k_{27}^{(0)} \) from birth-death process. Plots of successive interpolations and of reversible isomerization exhibit similar qualitative fit and are omitted from display. We see from Figure C.1 that an interpolant cannot pass through the past data without changing its inflection, of which an exponential function is incapable. Consequently all 7 candidate biasing parameters underwent Step 3 to satisfy Assumption 1 and yielded a single set after removing duplicates. At this point, the final set of candidate biasing parameters is the same as \( k^{(e)} \), and the resulting SParSE estimate supplies no additional information. Similar results occur in the next two subsequent interpolations. We note that for both failed cases, taking five past data sets closest to \( P_{\varepsilon} \) does not improve the quality of interpolation. This is because three sets of biasing parameters that underperturb the system (i.e., below the yellow dotted line in Figure C.1) produce estimates that
are farther from $\mathcal{P}_E$ than the initial set produces. Thus, upon completion of the first interpolation, we have two estimates with $k^{(v)}$, which are closer to $\mathcal{P}_E$ than the rest, causing the biasing parameters responsible for the smallest estimate to be removed as the 6th data in the second interpolation. Following the same logic, the five sets of biasing parameters used in the last interpolation are the set corresponding to the largest estimate less than $\mathcal{P}_E$, three sets corresponding to $k^{(v)}$, and the initial set. Because the second inflection point is never removed, an exponential interpolant is unable to compute candidates that do not over-perturb the system. The failure would have been avoided if SParSE was configured to either (1) take 4 closest data sets instead of 5, or (2) set the maximum number of interpolations to 4 instead of 3, both of which would have removed the initial set of biasing parameters in the last iteration and yielded viable biasing parameter candidates. Although an additional round of exponential interpolation or decreasing the number of data sets used in it would have prevented the three failures in, we remind the reader that such changes may induce non-convergence in other samples. Lastly, we observed this type of poor agreement between the past data and its exponential interpolant only on over-perturbing initial rates.
FIG. C.1: Illustration of first exponential interpolation for the two failed initial reaction rates in birth-death process with $P_E = 0.60$ and $\epsilon P_E = 0.01$. Figure 1-A represents $k_3^{(0)} = [1.606 \ 0.0140]$ and Figure 1-B $k_27^{(0)} = [1.684 \ 0.0148]$. Yellow horizontal dotted line is the desired number of successful trajectories, i.e., $N \times P_E = 3 \times 10^4$. Blue and green circles denote the past intermediate biasing parameters for $R_1$ and $R_2$, respectively, normalized with respect to $k^{(0)}$. Blue and green dashed lines are the interpolants constructed from the past intermediate biasing parameters, and the red triangles are candidate biasing parameters computed according to Step 3 in Subsection *Updating intermediate reaction rates*.