A Proof of Lemma 1

Suppose there are two stationary points, *i.e.*, $\phi'_L(x) = \phi'_L(y) = 0$, thus, by continuity of $\phi_L(.)$ in $(t_{s_i}, t_{s_{i+1}})$ there must be a $z \in (x, y)$ such that $\phi''_L(z) = 0$. We show it is a contradiction as

$$\phi''_L(t_s) = \sum_l^L \gamma_l \beta_l^2 e^{\beta_l t_s} > 0 \tag{12}$$

for all $1 \leq l \leq L$.

B Proof of Lemma 2

Assume we would like to find the maximizer of $\phi_L(.)$ in interval (a, b) and consider two points at one-third and two-third of the interval, *i.e.*, $c = a + \frac{b-a}{3}$ and $d = a + 2\frac{b-a}{3}$. It can be easily shown that, if $\phi_L(c) < \phi_L(d)$, then the maximizer will be on interval (c, b) and, if $\phi_L(c) < \phi_L(d)$, then the maximizer must lie on interval (a, d). Therefore, by two evaluations, we can shrink the interval containing the maximizer by a factor of $\frac{2}{3}$. Then, to reach the ϵ-neighborhood of the real maximizer, we need evaluate the function $2 * r$ times, where

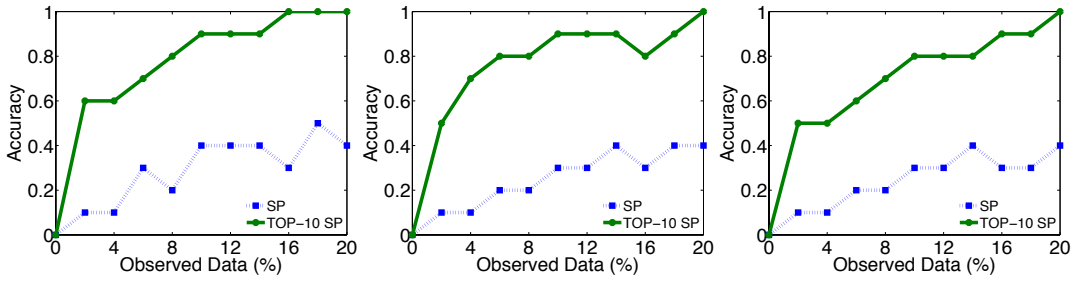
$$(t_{s_{i+1}} - t_{s_i})(2/3)^r < \epsilon. \tag{13}$$

This will prove our claim.

C Additional Experimental Results

In this section, we provide additional experimental results on synthetic data, including an evaluation of the performance of our method against the percentage of observed infections and the number of Montecarlo samples, as well as a scalability analysis.

Performance vs. percentage of observed infections. Intuitively, the greater the number of observed infections, the more accurately our method can infer the true source and its infection time. Figure 6 confirms this intuition by showing the success probability against percentage of observed infections. However, we also find that the greater is the percentage of observed infections, the smaller is the effect of observing additional infections; a diminishing return property.



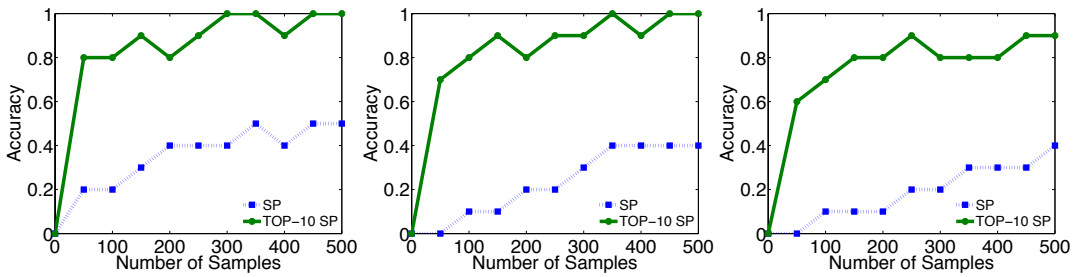
(a) Random (b) Hierarchical (c) Core-periphery

Figure 6: Accuracy vs. % observed infections.

Performance vs. number of Montecarlo samples. Drawing more transmission time samples $\{\tau_{ji}\}_{(j,i)\in\mathcal{E}}$ leads to a better estimate of Eq. 6, and thus a greater accuracy of our method. Figure 7 shows the success probability against number of samples. Importantly, we observe that as long as the number of samples is large enough, the performance of our method quickly flattens and does not depend on the number of samples any more.

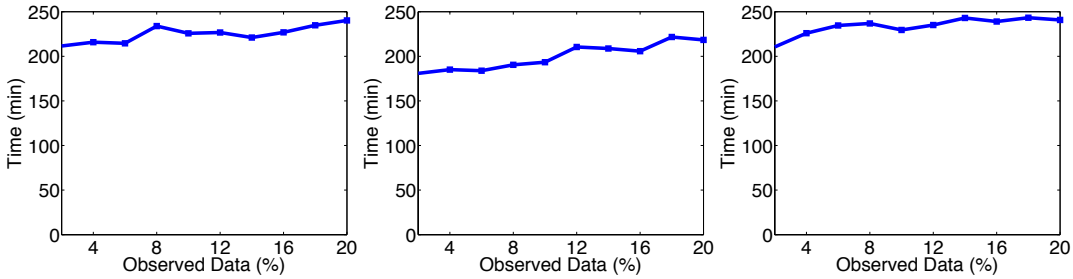
Running time vs. percentage of observed infections. Figure 8 plots the average running time to infer the source of a single cascade against the percentage of observed infections. Perhaps surprisingly, the running time barely increases with the percentage of observed infections.

Running time vs. number of samples. Figure 9 plots the average running time against the number of Montecarlo samples used to approximate the likelihood, Eq. 6.



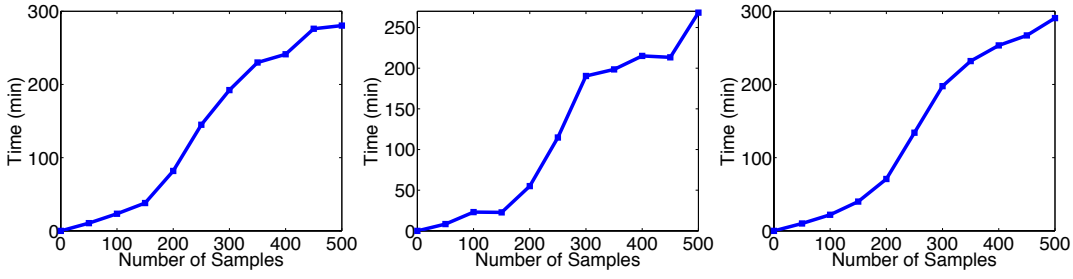
(a) Random (b) Hierarchical (c) Core-Periphery

Figure 7: Accuracy vs. number of samples.



(a) Random (b) Hierarchical (c) Core-periphery

Figure 8: Running time vs. % observed infections.



(a) Random (b) Hierarchical (c) Core-periphery

Figure 9: Running time vs. number of samples.

Toy example. We consider the same 64-node hierarchical Kronecker network as in Section 5.1 and visualize the approximate likelihood given by Eq. 8 against number of observed cascades ($C = 1, \ldots, 8$) for each node in the network using 150 Monte Carlo samples.

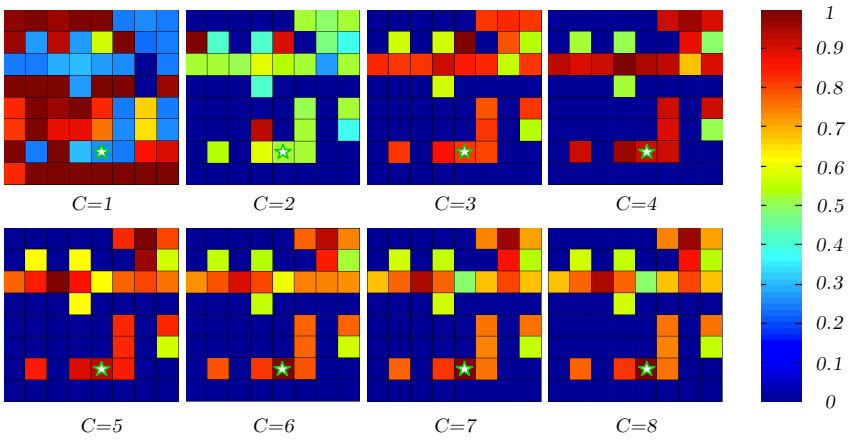
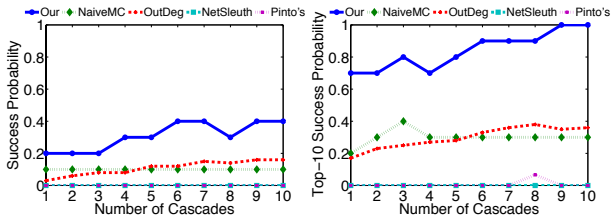


Figure 10: Evolution of the proposed method with respect to the number of cascades.

Accuracy on a hierarchical Kronecker network. We additionally evaluate the accuracy of our method in comparison with the same two state of the art methods and two baselines as in Section 5.1 in a Kronecker hierarchical network. Figure 11 shows the success probability (SP) and top-10 success probability, and mean squared error (MSE) on the estimation of t_s.



(a) SP (b) Top-10 SP

Figure 11: Success Probability (SP), Top-10 Success Probability (Top-10 SP) and Mean-squared error (MSE) on the estimation of t_s for a hierarchical Kronecker network.