

# Estimation of reproductive number and prevalence using genomic and time series data

Alexander E. Zarebski, Louis du Plessis, Kris V. Parag, and Oliver G. Pybus

## Motivation

- Inferring the dynamics of pathogen transmission during an outbreak is an important problem in infectious disease epidemiology.
- Combining datatypes is challenging: methods to analyse time series of cases and sequence data have developed in isolation of each other.
- Existing approaches for estimating the reproduction number and prevalence of infection may struggle with large datasets (Andréoletti *et al* (2022)).

## Methods

- We used an approximation to make the birth-death model from Manceau *et al* (2021) feasible for large datasets. Our algorithm has linear complexity.
- We extended the method to enable the use of a time series of confirmed (but not sequenced) cases. Time-series Integration Method Through Approximation of Moments
- The resulting method has been implemented as a BEAST2 package called TimTam available at [aезarebski.github.io/timtam/](https://aезarebski.github.io/timtam/) with several tutorials.

## Discussion

- TimTam is a computationally tractable way to estimate key quantities while using more of the available data.
- The BEAST2 package makes it possible for people to use TimTam in their existing workflows.
- The approximation scheme could be extended to incorporate additional data types.

## References

Zarebski, A. E. et al 2022 A computationally tractable birth-death model that combines phylogenetic and epidemiological data. *PLoS Comput. Biol.*

Manceau, M. et al., 2021. The probability distribution of the ancestral population size conditioned on the reconstructed phylogenetic tree with occurrence data. *J. Theor. Biol.*

Andréoletti, J. et al 2022 The Occurrence Birth-Death Process for combined-evidence analysis in macroevolution and epidemiology. *Syst. Biol.*

Estimate  $R_0$  and the number of infectious from sequences and a time series of cases with a birth-death model.

Now available as the TimTam BEAST2 package!

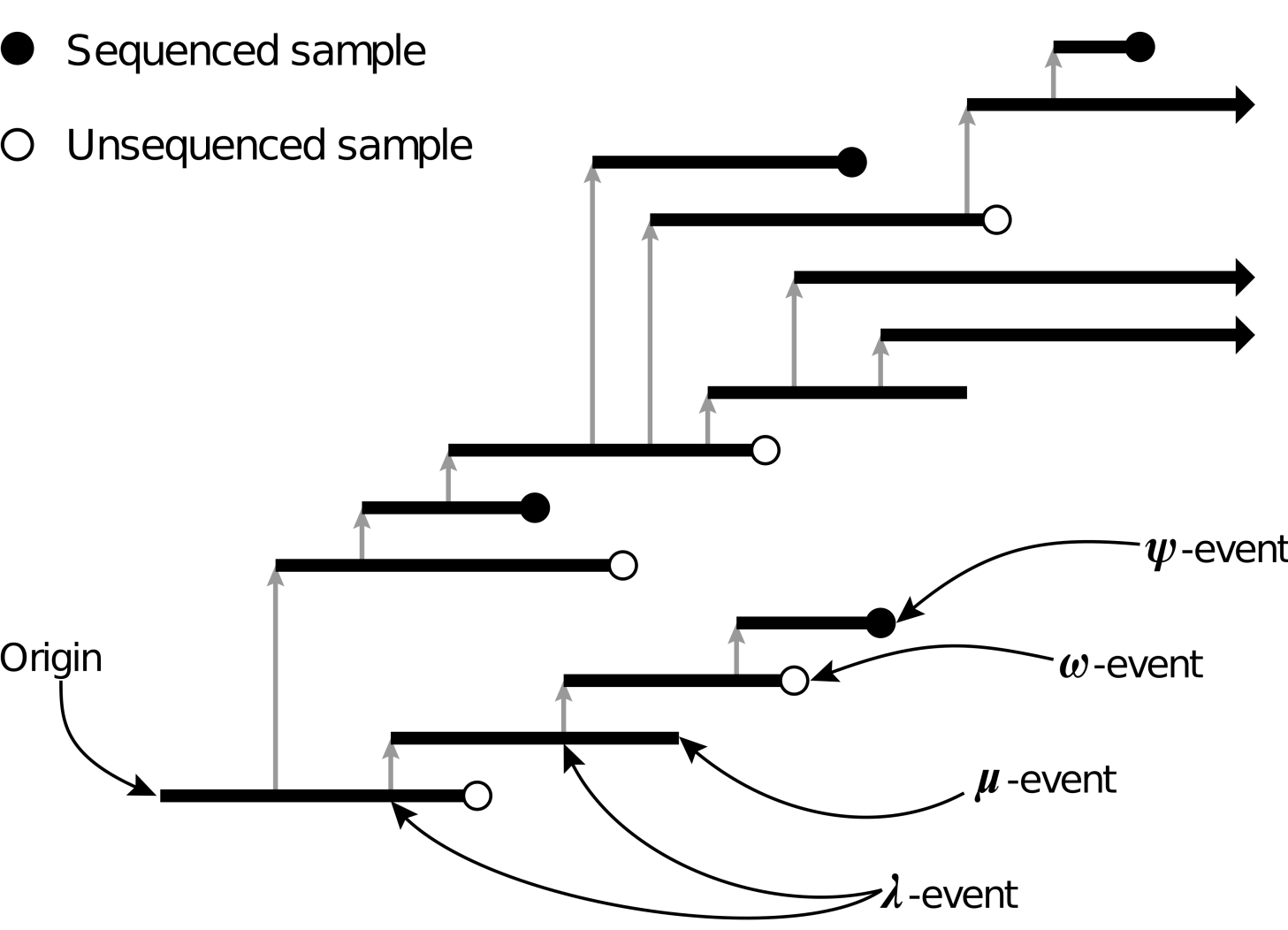
Package



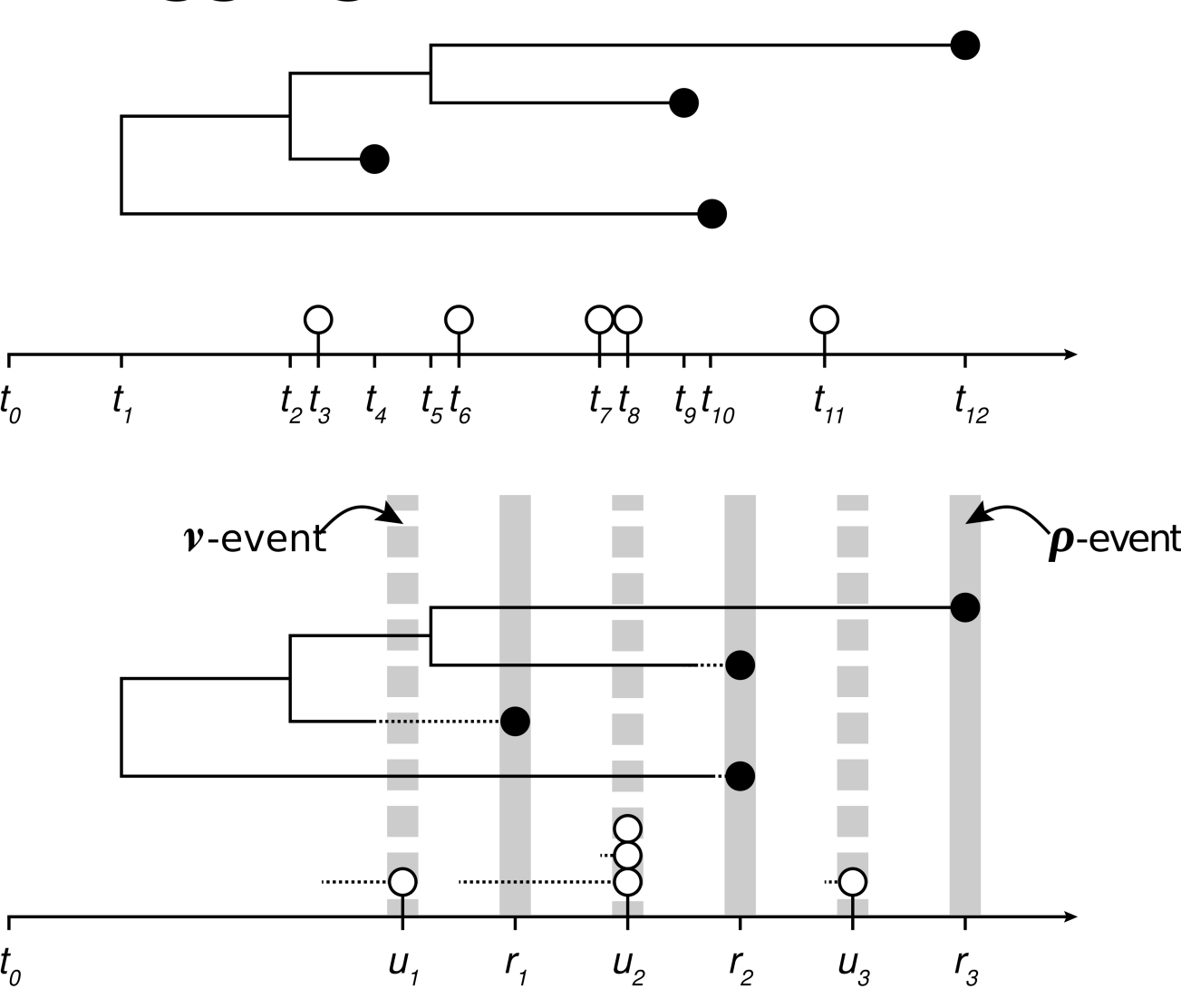
Paper



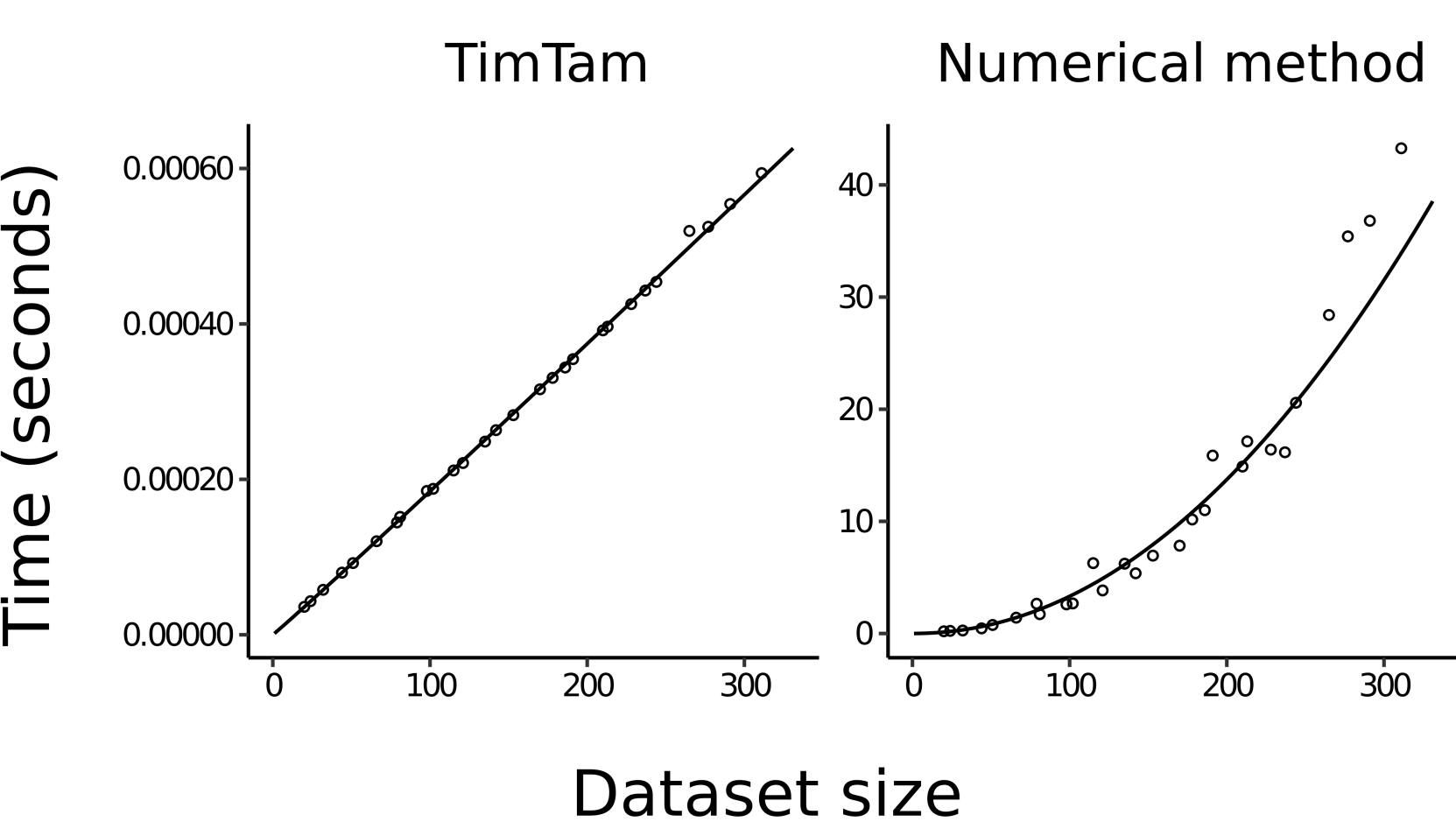
A birth-death process is observed in two ways



Observations in continuous time may be aggregated into a time series



Evaluation of the log-likelihood is linear in the size of the dataset



The  $R_0$  estimates converge to the truth as the datasets get bigger

