

# Supplementary Material of “HIV with contact-tracing: a case study in Approximate Bayesian Computation”

Michael G.B. Blum<sup>1</sup>, Viet Chi Tran<sup>2</sup>

<sup>1</sup> CNRS, Laboratoire TIMC-IMAG, Université Joseph Fourier, Grenoble, France

<sup>2</sup> Laboratoire P. Painlevé, Université Lille 1, France

The Supplementary Material contains 4 figures and the dataset consisting of the 29 simulated detection times used in Section 3.4 of the paper.

- Figure 1. Boxplots of the 200 estimated modes, medians, 2.5% and 97.5% quantiles of the posterior distribution for  $\lambda_2$ , the random screening detection rate.
- Figure 2. Boxplots of the 200 estimated modes, medians, 2.5% and 97.5% quantiles of the posterior distribution for  $\lambda_3$ , the contact tracing detection rate.
- Figure 3. Posterior Predictive Distributions (PPD) of  $R_{21.5}^1$ ,  $R_{21.5}^2$ ,  $I_6$ , and the mean sojourn time in the class I. Parameter inference has been performed in the SIR model with a mass action principle for the detection rate (see equation (2.1) of the paper). The smooth rejection ABC algorithm, using the 2 statistics  $R^1$  and  $R^2$  as the summary statistics, was considered.
- Figure 4. Posterior Predictive Distributions (PPD) of  $R_{21.5}^1$ ,  $R_{21.5}^2$ ,  $I_6$ , and the mean sojourn time in the class I. Parameter inference has been performed in the SIR model with a mass action principle for the detection rate (see equation (2.1) of the paper). The NCH-ABC algorithm, using the 51 statistics as the summary statistics, was considered.
- The 29 detection times that have been simulated using  $\lambda_1 = 0.12$ ,  $\lambda_2 = 1$ ,  $S_0 = 30$  and  $I_0 = 1$ .

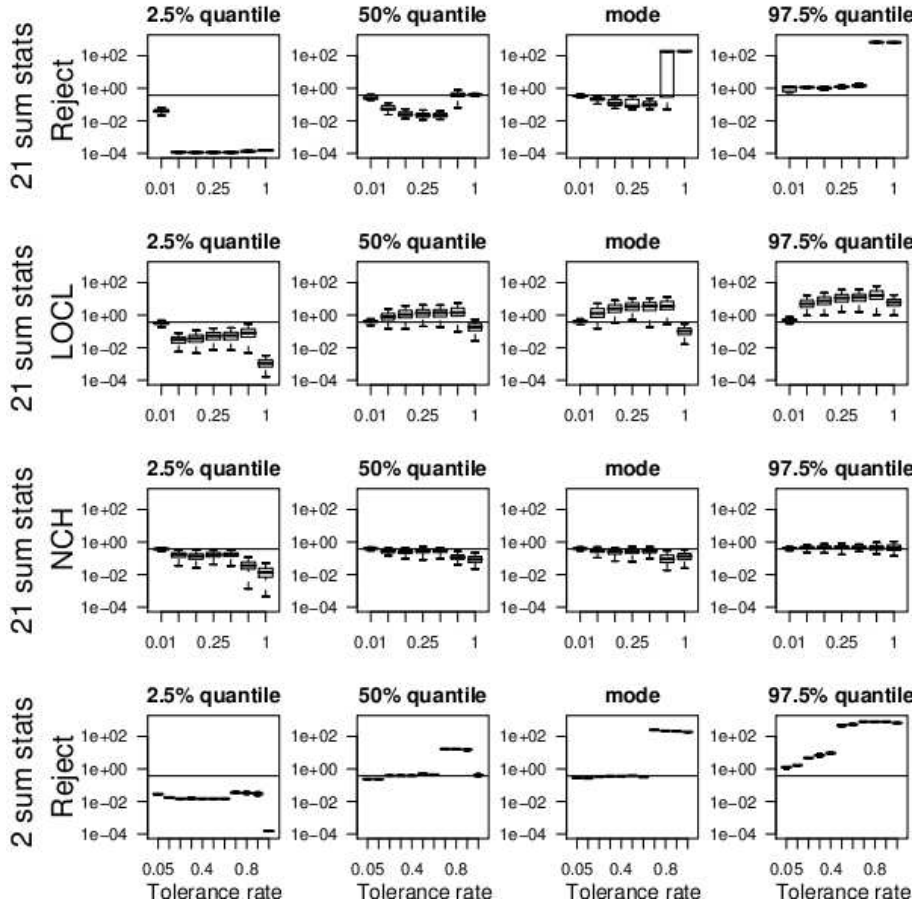


Figure 1: Boxplots of the  $M = 200$  estimated modes and quantiles (2.5%, 50%, and 97.5%) of the approximate posterior distributions of  $\lambda_2$ . For each ABC method and each value of the tolerance rate, 200 posterior distributions are computed for each of the 200 synthetic data sets. The horizontal lines correspond to the true value  $\lambda_2 = 3.75 \times 10^{-1}$  used when simulating the 200 synthetic data sets. The different tolerance rates are 0.01, 0.05, 0.10, 0.25, 0.50, 0.50, 0.75, and 1 for all the ABC methods except the rejection scheme with the two summary statistics. For the latter method, the tolerance rates are 0.007, 0.02, 0.06, 0.13, 0.27, 0.37, 0.45, 0.53, 0.66, 0.80, 1.

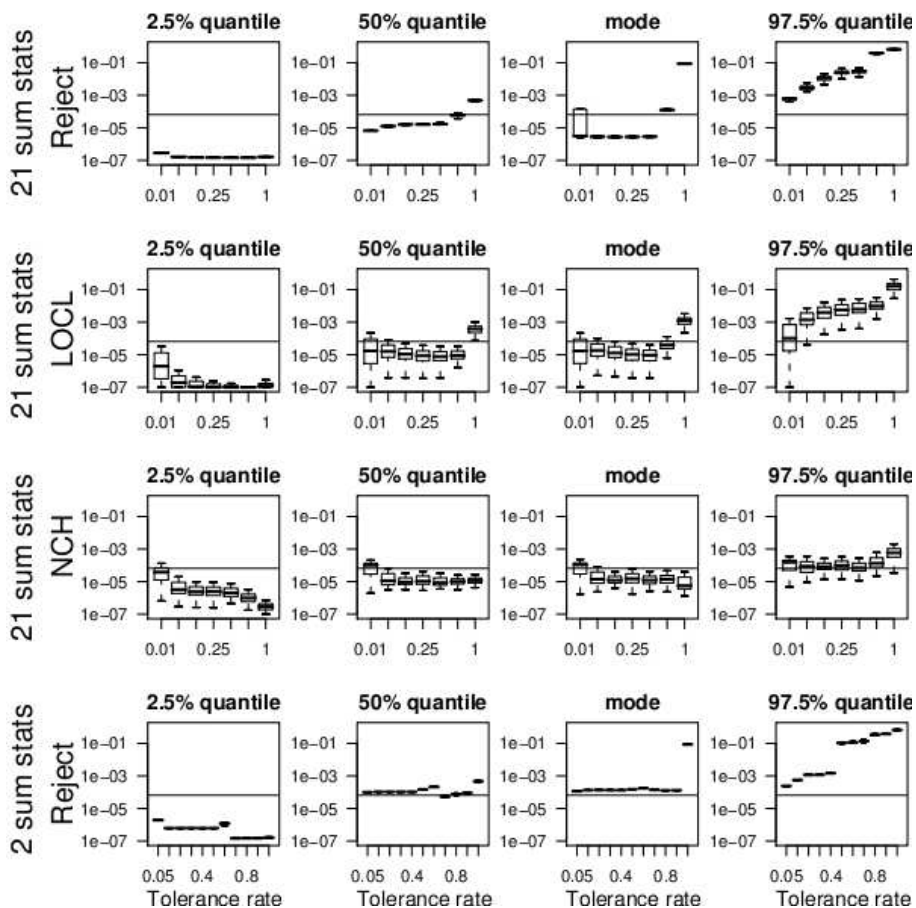


Figure 2: Boxplots of the  $M = 200$  estimated modes and quantiles (2.5%, 50%, and 97.5%) of the approximate posterior distributions of  $\lambda_3$ . For each ABC method and each value of the tolerance rate, 200 posterior distributions are computed for each of the 200 synthetic data sets. The horizontal lines correspond to the true value  $\lambda_3 = 6.55 \times 10^{-5}$  used when simulating the 200 synthetic data sets. The different tolerance rates are 0.01, 0.05, 0.10, 0.25, 0.50, 0.50, 0.75, and 1 for all the ABC methods except the rejection scheme with the two summary statistics. For the latter method, the tolerance rates are 0.007, 0.02, 0.06, 0.13, 0.27, 0.37, 0.45, 0.53, 0.66, 0.80, 1.

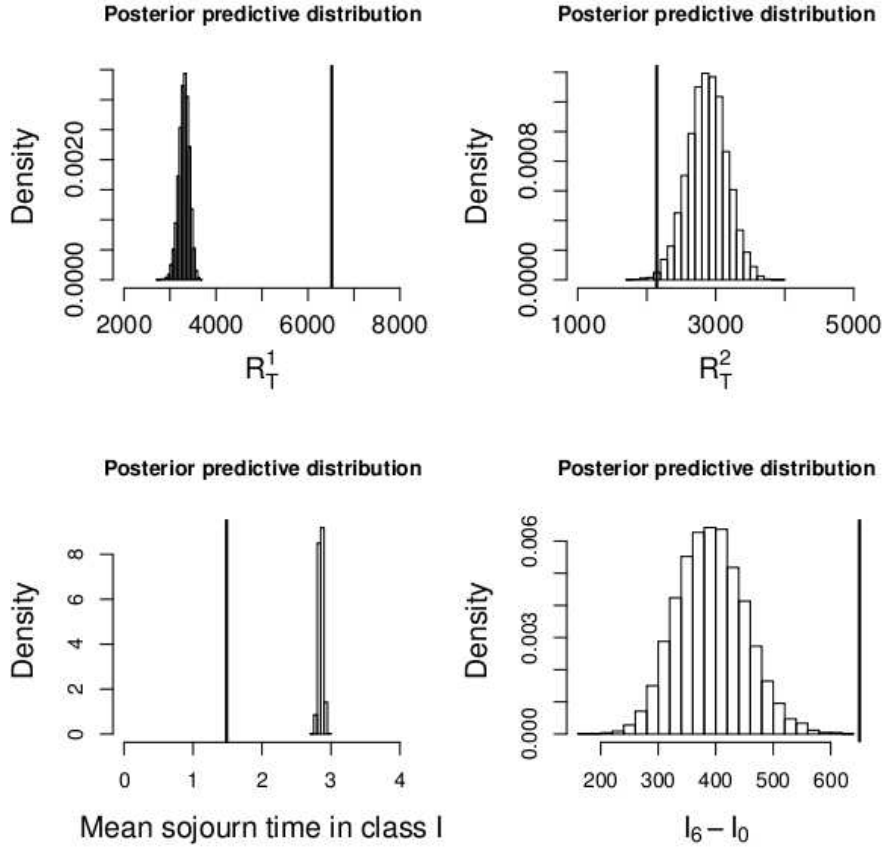


Figure 3: Bayesian posterior predictive distributions of  $R_{21.5}^1$ ,  $R_{21.5}^2$ ,  $I_6$ , and the mean sojourn time in the class I. The vertical lines correspond to the observations. The SIR model corresponds to the model with a mass action principle. The partial posterior samples are obtained with the smooth rejection ABC algorithm by making use of the 2 infinite-dimensional summary statistics  $R^1$  and  $R^2$ . A tolerance rate of  $P_\delta = 1\%$  is considered for each summary statistic.

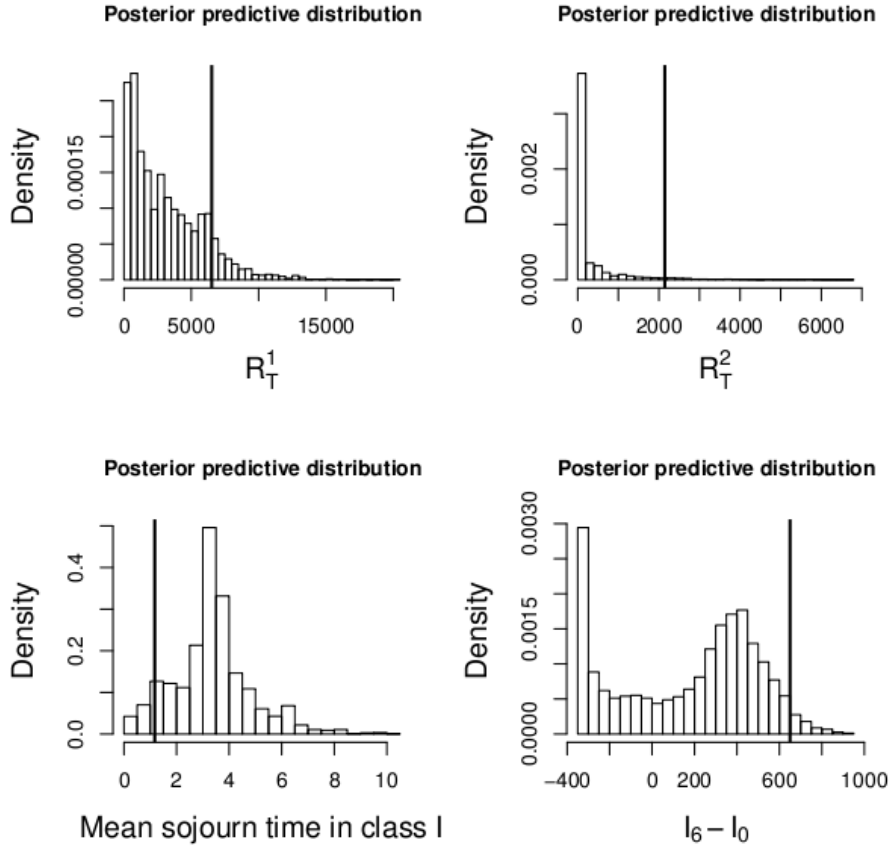


Figure 4: Bayesian posterior predictive distributions of  $R_{21.5}^1$ ,  $R_{21.5}^2$ ,  $I_6$ , and the mean sojourn time in the class I. The SIR model corresponds to the model with a mass action principle. The vertical lines correspond to the observations. The partial posterior samples were obtained with the ABC-NCH method ( $P_\delta = 0.5\%$ ) by making use of the 51 summary statistics.

The 29 simulated detection times are as follows:  $t_1 = 0.00000000$ ,  $t_2 = 0.03585958$ ,  $t_3 = 0.51218642$ ,  $t_4 = 0.79642462$ ,  $t_5 = 0.95457767$ ,  $t_6 = 1.27281694$ ,  $t_7 = 1.34650878$ ,  $t_8 = 1.62610267$ ,  $t_9 = 1.63237045$ ,  $t_{10} = 1.64755829$ ,  $t_{11} = 1.69982485$ ,  $t_{12} = 1.76292975$ ,  $t_{13} = 1.84627091$ ,  $t_{14} = 1.86995518$ ,  $t_{15} = 1.96626950$ ,  $t_{16} = 2.19536465$ ,  $t_{17} = 2.25028658$ ,  $t_{18} = 2.40470011$ ,  $t_{19} = 2.72138201$ ,  $t_{20} = 2.77220509$ ,  $t_{21} = 2.82953372$ ,  $t_{22} = 2.83636998$ ,  $t_{23} = 2.91944671$ ,  $t_{24} = 2.96668428$ ,  $t_{25} = 3.02074108$ ,  $t_{26} = 3.42875757$ ,  $t_{27} = 3.63046765$ ,  $t_{28} = 3.97189122$ , and  $t_{29} = 4.77789949$ .