Supplementary Material: A Parameter Estimation Method for Single Molecule Fluorescence Lifetime Data

Dan Sewell,† Hajin Kim,‡ Taekjip Ha,‡ and Ping Ma*,†

Department of Statistics, University of Illinois at Urbana-Champaign, and Department of Physics, University of Illinois at Urbana-Champaign

E-mail: pingma@illinois.edu

*To whom correspondence should be addressed
†Department of Statistics, University of Illinois at Urbana-Champaign
‡Department of Physics, University of Illinois at Urbana-Champaign
Simulation Results by Parameter Configuration

Figures 1 to 5 show, for each of 30 parameter configurations, a two-dimensional histogram of the quantity \( \log(\hat{\tau}^2/\tau^2) \), where \( \tau \) is the true parameter value and \( \hat{\tau} \) is the estimate of \( \tau \); the vertical axes correspond to the estimates obtained from fitting the biexponential distribution directly, and the horizontal axes correspond to the estimates obtained from the gamma conversion method.

Figures 6 to 15 show, for each of the 30 parameter configurations, a two-dimensional histogram of Hellinger Distances vs. Pearson’s \( \chi^2 \) statistics. For each configuration there is a plot corresponding to the estimated curves from fitting the biexponential distribution directly (left) and to the estimated curves from the gamma conversion method (right). As mentioned in the main text, the Hellinger Distance is a metric between two probability distributions. Here the estimated curve is compared with the true curve, and hence a small (large) distance implies that the estimate is very close to (very far from) the true curve. Pearson’s \( \chi^2 \) statistic indicates how well the estimates fit the data. Therefore if an estimate is overfitting the data it will have a low \( \chi^2 \) value but high distance from the truth.
Figure 1: Two-dimensional histograms of log($\hat{\tau}_2 / \tau_2$), where $\hat{\tau}_2$ is the estimate from either using the biexponential pdf directly (vertical axis) or the gamma conversion method (horizontal axis); each plot corresponds to varying true values of $c$ and $\tau_2 = k \tau_1$. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 2: Two-dimensional histograms of $\log(\hat{\tau}_2/\tau_2)$, where $\hat{\tau}_2$ is the estimate from either using the biexponential pdf directly (vertical axis) or the gamma conversion method (horizontal axis); each plot corresponds to varying true values of $c$ and $\tau_2 = k\tau_1$. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 3: Two-dimensional histograms of $\log(\hat{\tau}_2/\tau_2)$, where $\hat{\tau}_2$ is the estimate from either using the biexponential pdf directly (vertical axis) or the gamma conversion method (horizontal axis); each plot corresponds to varying true values of $c$ and $\tau_2 = k\tau_1$. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 4: Two-dimensional histograms of $\log(\hat{\tau}_2 / \tau_2)$, where $\hat{\tau}_2$ is the estimate from either using the biexponential pdf directly (vertical axis) or the gamma conversion method (horizontal axis); each plot corresponds to varying true values of $c$ and $\tau_2 = k \tau_1$. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
$\hat{\tau}_2$ is the estimate from either using the biexponential pdf directly (vertical axis) or the gamma conversion method (horizontal axis); each plot corresponds to varying true values of $c$ and $\tau_2 = k \tau_1$. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 6: Two-dimensional histogram for Hellinger Distance (vertical axis) and Pearson’s $\chi^2$ statistic (horizontal axis) for simulated data, where the plotted values have been aggregated over varying true values of $c$ and $\tau_2$. For each subfigure (a) through (c), fitting the biexponential directly gives the plot on the left, and using gamma conversion method gives the plot on the right. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 7: Two-dimensional histogram for Hellinger Distance (vertical axis) and Pearson’s $\chi^2$ statistic (horizontal axis) for simulated data, where the plotted values have been aggregated over varying true values of $c$ and $\tau^2$. For each subfigure (a) through (c), fitting the biexponential directly gives the plot on the left, and using gamma conversion method gives the plot on the right. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 8: Two-dimensional histogram for Hellinger Distance (vertical axis) and Pearson’s $\chi^2$ statistic (horizontal axis) for simulated data, where the plotted values have been aggregated over varying true values of $c$ and $\tau_2$. For each subfigure (a) through (c), fitting the biexponential directly gives the plot on the left, and using gamma conversion method gives the plot on the right. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 9: Two-dimensional histogram for Hellinger Distance (vertical axis) and Pearson’s $\chi^2$ statistic (horizontal axis) for simulated data, where the plotted values have been aggregated over varying true values of $c$ and $\tau_2$. For each subfigure (a) through (c), fitting the biexponential directly gives the plot on the left, and using gamma conversion method gives the plot on the right. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 10: Two-dimensional histogram for Hellinger Distance (vertical axis) and Pearson’s $\chi^2$ statistic (horizontal axis) for simulated data, where the plotted values have been aggregated over varying true values of $c$ and $\tau_2$. For each subfigure (a) through (c), fitting the biexponential directly gives the plot on the left, and using gamma conversion method gives the plot on the right. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 11: Two-dimensional histogram for Hellinger Distance (vertical axis) and Pearson’s $\chi^2$ statistic (horizontal axis) for simulated data, where the plotted values have been aggregated over varying true values of $c$ and $\tau_2$. For each subfigure (a) through (c), fitting the biexponential directly gives the plot on the left, and using gamma conversion method gives the plot on the right. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 12: Two-dimensional histogram for Hellinger Distance (vertical axis) and Pearson’s $\chi^2$ statistic (horizontal axis) for simulated data, where the plotted values have been aggregated over varying true values of $c$ and $\tau_2$. For each subfigure (a) through (c), fitting the biexponential directly gives the plot on the left, and using gamma conversion method gives the plot on the right. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 13: Two-dimensional histogram for Hellinger Distance (vertical axis) and Pearson’s $\chi^2$ statistic (horizontal axis) for simulated data, where the plotted values have been aggregated over varying true values of $c$ and $\tau_2$. For each subfigure (a) through (c), fitting the biexponential directly gives the plot on the left, and using gamma conversion method gives the plot on the right. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 14: Two-dimensional histogram for Hellinger Distance (vertical axis) and Pearson’s $\chi^2$ statistic (horizontal axis) for simulated data, where the plotted values have been aggregated over varying true values of $c$ and $\tau_2$. For each subfigure (a) through (c), fitting the biexponential directly gives the plot on the left, and using gamma conversion method gives the plot on the right. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Figure 15: Two-dimensional histogram for Hellinger Distance (vertical axis) and Pearson’s $\chi^2$ statistic (horizontal axis) for simulated data, where the plotted values have been aggregated over varying true values of $c$ and $\tau_2$. For each subfigure (a) through (c), fitting the biexponential directly gives the plot on the left, and using gamma conversion method gives the plot on the right. Intensity is graded from blue (lowest) to yellow (highest), white indicating no counts.
Where τ₁ is not known

Here we consider the case where τ₁ is not assumed to be known. The same simulation data sets and real data sets as described in the paper were analyzed. We wish to compare how the estimations compared to the truth; to avoid identifiability issues, we present the estimated decay curves with the true curve superimposed. Figures 16 to 25 give these plots corresponding to fitting the biexponential directly and to the gamma conversion method for each of the thirty simulation configurations. Figure 26 gives this plot for the real data analysis. Besides getting a visualization on how well the estimation methods match the truth, we specifically wish to investigate the tendency to overfit the data. To this end, we again plot the Hellinger Distances, which gives a good idea as to how closely the estimated curves match with the truth, vs. Pearson’s χ² statistics, which show how well the estimations fit the individual small data sets. Figure 27 gives the two-dimensional histogram of Hellinger Distances vs. χ² statistics aggregating over all 30 simulation configurations; Figure 28 gives the same for the real data sets. Figures 29 through 38 give these two-dimensional histograms separated by simulation configuration.
Figure 16. Estimated decay curves for simulated data where $\kappa$ is unknown, with the true curve superimposed. For each subplot (a) through (e), fitting the biexponential directly gives the plot on the left, and using gamma conversion method gives the plot on the right.

(c) $c = 0.6, \gamma = 0.900$

(q) $c = 0.6, \gamma = 0.800$

(a) $c = 0.6, \gamma = 0.500$
Figure 17: Estimated decay curves for simulated data where $\gamma$ is unknown, with the true curve superimposed. For each subplot, (a) through (c), the biexponential directly gives the plot.

$\gamma = 0.6, c = 1.0$

$\gamma = 0.6, c = 0.996$

$\gamma = 0.6, c = 0.950$