TABLE 6.1

Supporting Table 1. *Average AMSE* $\langle \|f - \widehat{f}\|_2^2/m \rangle$ *and measured standard deviation in parentheses for several estimators of* f *in the Monte Carlo study described in section* 5.1. *The subscript corresponds to the design matrix used, and the "noise ratio" refers to the ratio of* σ_f^2/σ_{df}^2. *The correlation coefficient* ρ *was set here to* 0.5 *(this parameter is needed to specify* $\boldsymbol{R}$*).*

Noise ratio	f_1	f_2	f_3	f_4	f_5
1/16	0.28(0.10)	0.27(0.10)	0.29(0.10)	1.05(0.36)	0.47(0.18)
1/4	0.21(0.08)	0.21(0.08)	0.22(0.08)	0.36(0.13)	0.47(0.18)
1/2	0.18(0.07)	0.18(0.07)	0.19(0.07)	0.24(0.09)	0.47(0.19)
2	0.22(0.11)	0.22(0.11)	0.23(0.11)	0.29(0.14)	0.86(0.26)
4	0.33(0.19)	0.33(0.19)	0.36(0.20)	0.51(0.26)	1.31(0.46)
16	0.80(0.63)	0.80(0.63)	0.86(0.68)	1.95(0.92)	3.00(2.06)

TABLE 6.2

Supporting Table 2. *Average AMSE* $\langle \|f - \widehat{f}\|_2^2/m \rangle$ *and measured standard deviation in parentheses for several estimators of* f *in the Monte Carlo study described in section* 5.1. *The subscript corresponds to the design matrix used, and the "noise ratio" refers to the ratio of* σ_f^2/σ_{df}^2. *The correlation coefficient* ρ *was set here to* 0 *(this parameter is needed to specify* $\boldsymbol{R}$*).*

Noise ratio	f_1	f_2	f_3	f_4	f_5
1/16	0.36(0.14)	0.36(0.14)	0.39(0.15)	1.06(0.36)	0.47(0.18)
1/4	0.27(0.10)	0.27(0.10)	0.29(0.10)	0.36(0.13)	0.46(0.18)
1/2	0.22(0.09)	0.22(0.09)	0.24(0.09)	0.24(0.09)	0.47(0.18)
2	0.27(0.13)	0.27(0.13)	0.29(0.14)	0.29(0.13)	0.86(0.26)
4	0.41(0.22)	0.41(0.22)	0.43(0.24)	0.51(0.25)	1.30(0.44)
16	0.93(0.69)	0.93(0.69)	0.98(0.72)	1.94(0.91)	2.98(2.03)

6. Conclusions and outlook. We demonstrated how a single-molecule time series can be transformed, via local maximum likelihood-type methods, into scatterplot data approximating pointwise function and derivative information associated with an SDE. The functions needed by an SDE approximating the global dynamics of the time series were obtained using P-spline techniques. The PuDI design matrix was shown to be useful in this context. The PuDI design matrix exploited some of the advantageous properties of the TPF basis; numerical difficulties were overcome with a recent algorithm [20]. The use of GLS along with P-splines was shown to influence the estimated curves, and the difference was shown to be relevant in regards to predicting/simulating physical quantities of interest. For example, the work computation associated with the ion-channel system studied benefited substantially from the GLS implementation. When this procedure was repeated for different time series, it was shown that the global SDE functions estimated from different time series exhibited variation in part due to a latent process; i.e., our data consisted of "subject specific curves." We briefly discussed why this is relevant information to modern biophysics applications [4, 26].

Although we focused on simulation data, the methodology is also applicable to experimental data [6, 8, 15]. Applications making fuller use of pointwise function estimates and derivative proxies calibrated from time series, as the PuDI method was demonstrated to do, show promise as tools that can be used for understanding the rich amount of information contained in recent single-molecule experiments and computer simulations. Other areas where function and derivative scatterplot information is available and a PuDI might be helpful include geosciences [14] and finance [13]. MATLAB scripts illustrating the PuDI method can be found in the Supporting Material which is available online from http://www.caam.rice.edu/tech_reports/2009/