Supplementary information to «Dynamical Friction Effects on the Photoisomerization of a Model Protonated Schiff Base in Solution»

João Pedro Malhado,[†] Riccardo Spezia,[‡] and James T. Hynes[†,¶]

Département de Chimie, CNRS UMR 8640 PASTEUR, École Normale Supérieure, 24 rue Lhomond, 75005 Paris, France, Laboratoire Analyse et Modélisation pour la Biologie et l'Environement, CNRS UMR 8587, Université d'Evry Val d'Essonne, Bd. F. Mitterrand, 91025 Evry, France, and Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309-0215,USA

[*]To whom correspondence should be addressed
[†]École Normale Supérieure
[‡]Université d'Evry Val d'Essonne
[¶]University of Colorado at Boulder

Generation of correlated stochastic forces by autoregressive series

This appendix details the algorithm used to generate the environment-induced forces term F_r in the generalized Langevin equation (GLE), eq. (9) of the accompanying article, for an arbitrary friction kernel, based on the work of references [1,2].

The fluctuations of F_r are modeled as a stochastic process with the required statistical properties that F_r has a zero average value and that its tcf is be proportional to the applied friction kernel through eq. (10) of the article. While MD simulations have exposed deviations to the generally assumed gaussian distribution [3,4], the gaussian choice is employed in this study.

In the autoregressive method [1,2,5], a new occurrence of F_r is determined by linear prediction from previous F_r values, i.e., at a given time $t = n\Delta t$, where Δt is the integration time step and n is the time step number, F_r for one trajectory and one coordinate is given by a linear combination of previously occurred values for that trajectory and a gaussian random number ε with zero expectation value (white noise)

$$F_r(n\Delta t) = \sum_{i=1}^{min(p,n)} a_i F_r((n-i)\Delta t) + \varepsilon(n\Delta t). \tag{1}$$

Here a_i are coefficients termed generators which have to be calculated, and p is the model order corresponding to the maximum number of generators to be used. Equations for the coefficients can be obtained by multiplying (1) by $F_r(n\Delta t)$ and taking the ensemble average to obtain

$$\langle F_r(n\Delta t) \cdot F_r(n\Delta t)\rangle = \sum_{i=1}^{min(p,n)} \langle a_i F_r((n-i)\Delta t) \cdot F_r(n\Delta t)\rangle + \langle \varepsilon(n\Delta t) \cdot F_r(n\Delta t)\rangle . \tag{2}$$

Since the generators a_i are the same for all the ensemble and are constants for given values of p and n, they can be factored out of the correlation under the sum. Further the random number ε is

uncorrelated with the F_r value and the last term in (2) is thus zero, so that eq. (2 can be written

$$\langle F_r(n\Delta t)\cdot F_r(n\Delta t)\rangle = \sum_{i=1}^{min(p,n)} a_i \langle F_r((n-i)\Delta t)\cdot F_r(n\Delta t)\rangle . \tag{3}$$

It is now noted that the stochastic force tcf is related to the friction kernel by eq. (10) of the article, which in this work is taken as a specific analytic form and can thus be calculated at any time value. By taking a number of equations of the form of (3) equal to the smallest value between n and p, a system of linear equations is obtained from which all the generators can be obtained. This system of equations, called Yule-Walker equations[5], can be written in matrix form

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{min(p,n)-1} \\ \gamma_{-1} & \gamma_0 & \cdots & \gamma_{min(p,n)-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{-min(p,n)+1} & \gamma_{-min(p,n)+2} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{min(p,n)} \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{min(p,n)} \end{bmatrix}, \tag{4}$$

where $\gamma_i = \langle F_r(0)\cdot F_r(i\Delta t)\rangle$ and since tcfs are even with respect to time, $\gamma_{-i} = \gamma_i$. The matrix in (4) is thus symmetric and has a Toeplitz structure[6]. There are very efficient algorithms for solving linear systems with such a structure[6]. The greatest concern in solving the Yule-Walker equations has however not been computational efficiency, but numerical stability[7]. In fact for the cases studied in this thesis the matrix in eq. (4) is ill-conditioned, the system of equations is numerically unstable and very sensitive to rounding errors. This means that small variations in the matrix values have very significant effects on the values of the generators, and different sets of values will be obtained in using different numerical methods. In order to minimize this problem, the systems of equations (4), and the full procedure of determining F_r values, was solved using an arbitrary precision numerical implementation[8] using 60 digits of precision. It is however noted that differing sets of generators obtained with different numerical methods provide a solution within a reasonable numerical accuracy to eq. (4), and since the generators bear no physical significance, merely providing an instrument to obtain F_r values from eq. (1), their exact determination is not

crucial for the method.

In order to calculate the stochastic force value from eq. (1), apart from the generators values, it is required to determine the standard deviation σ_ε associated with the gaussian random number ε. This can be done by introducing the expression for $F_r(n\Delta t)$ into the last term in eq. (2) which, by noting that $\varepsilon(n\Delta t)$ is an independent random number uncorrelated to any value of F_r, yields

$$\langle F_r(n\Delta t)\cdot F_r(n\Delta t)\rangle = \sum_{i=1}^{min(p,n)} a_i \langle F_r((n-i)\Delta t)\cdot F_r(n\Delta t)\rangle + \langle \varepsilon(n\Delta t)\cdot \varepsilon(n\Delta t)\rangle . \tag{5}$$

The last term in eq. (5 is the second central moment (variance) of the distribution of the distribution of $\varepsilon(n\Delta t)$ which is the quantity sought. By rearranging eq. (5), one has

$${\sigma_\varepsilon}^2(n\Delta t) = \gamma_0 - \sum_{i=1}^{min(p,n)} a_i\gamma_i, \tag{6}$$

where, as before, $\gamma_i = \langle F_r(0)\cdot F_r(i\Delta t)\rangle$. The terms in the product under the sum have the same index, in contrast to eq. (1).

The model order p in eq. (1) also needs to be determined. It is expected that the larger the value of p is, the more accurate will be the tcf of the stochastic forces with respect to the target function. However, the maximum value of p is limited by the stability of the autoregressive method described by eq. (1) itself. In fact, it was found that for sufficiently high values of p the values of F_r start to diverge, reaching absolute values many orders of magnitude larger than the target variance $\langle F_r(0)^2\rangle$. Further, the maximum model order before divergence increases with the time step, and for the friction kernels used is a relatively small number ($p < 10$). In the present study, for simulations involving friction, the largest time step deemed reasonable was used, of value 0.48fs.

For such low model orders, with generator values calculated from eq. (4), the quality of the stochastic forces tfcs obtained from this method is poor, especially for long times. In order to improve the quality of the generators, the procedure suggested in [1] is followed, by which for a time step $n \leq p$ the generators are determined from eq. (4), but for the following time steps the system

of linear eq. (4) is replaced by a least squares problem of the form

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{min(p,n)-1} \\ \gamma_{-1} & \gamma_0 & \cdots & \gamma_{min(p,n)-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{-p+1} & \gamma_{-p+2} & \cdots & \gamma_0 \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{-min(m,n)+1} & \gamma_{-min(m,n)+2} & \cdots & \gamma_{p-min(m,n)} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_p \\ \vdots \\ \gamma_{min(m,n)} \end{bmatrix}, \tag{7}$$

where m is a fixed number greater than p. From eq. (7) generators (p in number) are obtained which are a best fit to reproduce the F_r tcf up to a time equal to $m\Delta t$. For the simulations presented here the values $p = 8$ and $m = 80$ were used for all friction kernels.

Equation (7) determines approximate values for the generators a_i and in this sense eq. (6) is no longer exact for times greater than $p\Delta t$. σ_ε is a rapid decaying function of time for values $t \leq p\Delta t$ but exponentially increases for values $t > p\Delta t$. This problem does not affect the shape of F_r tcf, but does affect its normalization. In this way, after all the values of F_r for the length of the simulation are determined, they have to be rescaled so as to agree with the required value of $\langle F_r(0)^2 \rangle$.

Given the behaviour of σ_ε with time, it is expected that the aforementioned issues with normalization of the F_r tcf will be relevant after an induction period but less significant for initial times. In order to obtain an homogeneous magnitude of F_r (to be rescaled) a cut-off to the initial values is applied which should be proportional to the time scale of the target tcf and to the model order p. For gaussian F_r tcf values obtained from eq. (1) for time values $t < p\Delta t \times \sigma$ (where σ here is the standard deviation of the gaussian) are rejected. For tcf of the form eq. (11) of the article the $t < p\Delta t/\alpha$ cut-off is used.

Energy evolution on the PSB model including dissipation

Figure 1 presents the evolution of the kinetic energy of each coordinate of the PSB model in solution, with time dependent dissipative effects included.

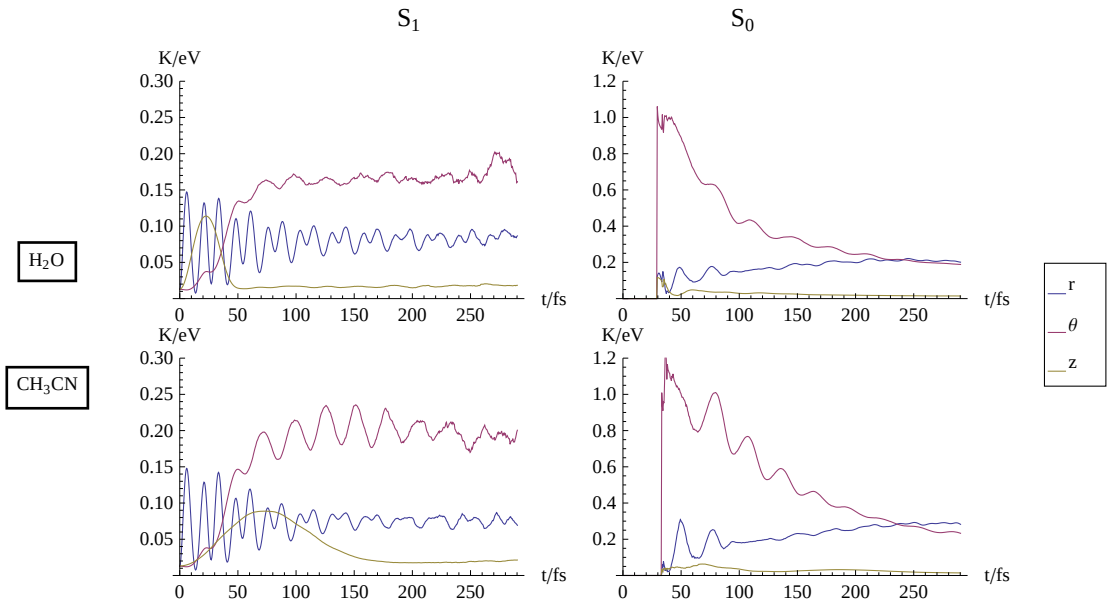

Figure 1: Excited (S_1) and ground (S_0) state evolution of the kinetic energy on each of the three model coordinates, for water and acetonitrile solvents. Motion on each coordinate is governed by a generalized Langevin equation with friction kernels defined in Sec. 3 of the article. (Note the difference in scale for excited and ground state).

The main difference in Figure 1 between the two solvent cases is the clear time scale difference of the solvent coordinate z motion, faster in the case of water. In the excited state the solvent starts out of equilibrium, the kinetic energy of the solvent coordinate rises as the solvent is set into motion, then lowers as the solvent reaches the equilibrium position (compare with Figure 5 of the article). A less apparent difference between the two solvent cases, is that oscillatory features of the torsional coordinate θ are better preserved for CH_3CN than for H_2O.

For both solvents, the kinetic energy in the bond length alternation (BLA) coordinate r, starts with a clear oscillatory pattern which evolves to a stationary value as dephasing of the vibrations occurs. Still in the excited state, the kinetic energy in θ rises as the trajectories go down the potential energy well (see Figure 13 of the article) and then reaches a stationary value. This value

corresponds to the average kinetic energy of the isomerization torsional motion on the excited state before the conical intersection (CI) is reached.

In the ground state the kinetic energy of θ starts at a very high value (higher than the plateau value in the excited state) and then decreases rapidly. Two factors contribute for this behaviour. First, is the fact that the ground state is first populated by the molecules which reach the CI at the twisted geometry first (see Figure 13) of the article, which corresponds to the fraction of molecules with the higher kinetic energy on this coordinate, and as slower molecules reach the ground state the average kinetic energy decreases. Second, is the friction dissipative effect included in the dynamics. In the ground state, after some oscillations, a steady increase of the kinetic energy of r is observed. This is due to the energy transfer from the other degrees of freedom of the model, namely torsion (see the shape of the free energy surfaces, Figure 2 of the article). Friction is not effective in removing energy from r in the time scale studied.

References

(1) Smith, D. E.; Harris, C. B. *J. Chem. Phys.* **1990**, *92*, 1304–1311.

(2) Nilsson, L. G.; Padró, J. A. *Mol. Phys.* **1990**, *71*, 355–367.

(3) Sesé, G.; Guàrdia, E.; Padró, J. A. *J. Stat. Phys.* **1990**, *60*, 501–518.

(4) Tuckerman, M.; Berne, B. J. *J. Chem. Phys.* **1993**, *98*, 7301–7318.

(5) *Nonlinear methods of spectral analysis*, 2nd ed.; Haykin, S., Ed.; Spinger, 1983; Vol. 34.

(6) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes*; Cambridge University Press, 1999.

(7) de Hoon, M. J. L.; van der Hagen, T. H. J. J.; Schoonewelle, H.; van Dam, H. *Ann. Nuc. Ener.* **1996**, *23*, 1219–1228.

(8) Wolfram Research Inc., *Mathematica*, 2007.