Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Electronic Supplementary Information for

Improved Reweighting Protocols for Variationally Enhanced Sampling Simulations with Multiple Walkers

Baltzar Stevensson*, and Mattias Edén*

Department of Materials and Environmental Chemistry, Stockholm University, SE-106 91 Stockholm, Sweden.

*Corresponding authors. E-mail: *baltzar.stevensson@mmk.su.se*; *mattias.eden@mmk.su.se*

Contents

Table S1. Computational Efforts for Reweighting.[a]

Simulation time	$\mathcal{T}$ =	90 ps	900 ps	9000 ps
Number of time points	$N_\mathcal{T}$ =	100	1000	10000
M_Σ^t, M_w^t	$N_\mathrm{F} N_\mathcal{T}(N_\mathcal{T}+1)N_W/2$	0.012 s	1.2 s	120 s
$M_\Sigma^\mathcal{T}$, $M_w^\mathcal{T}$	$N_\mathrm{F} N_\mathcal{T}(N_\mathcal{T}+1)N_W N_\mathrm{iter}$	0.23 s	23 s	2300 s
M_Σ^TP, M_Σ^VES	$N_\mathrm{F} N_\mathcal{T} N_s$	0.056 s	0.56 s	5.6 s

[a] Number of floating-point operations (N_op) associated with each reweighting protocol along with the CPU times required for the reweighting results of Fig. **3**a with $N_W = 4$. Here, $N_\mathrm{F} = 12$ is the number of Fourier coefficients [eqn (14)], $N_\mathrm{iter} = 10$ is the number of iterations to solve eqn (22), and $N_s = 1000$ is the total number of collective-variable samples/coordinates. The $M_\Sigma^\mathcal{T}$ and $M_w^\mathcal{T}$ methods are not suitable for "on-the-fly" calculations because they require consecutive re-evaluations of $c(t)$ at all time points $t \leqslant \mathcal{T}$ for each additional sampled time point, whereas the M_Σ^t, M_w^t, M_Σ^TP, and M_Σ^VES methods require only $N_\mathrm{op}/N_\mathcal{T}$ operations. The CPU time periods required to calculate $c(t)$ (rightmost columns) in serial mode were obtained on an Intel® Core™ i7-9700K CPU (3.60 GHz) desktop computer running Linux Ubuntu (version 18.04) and compiled using GNU Fortran (version 7.5.0) with optimization and double precision.

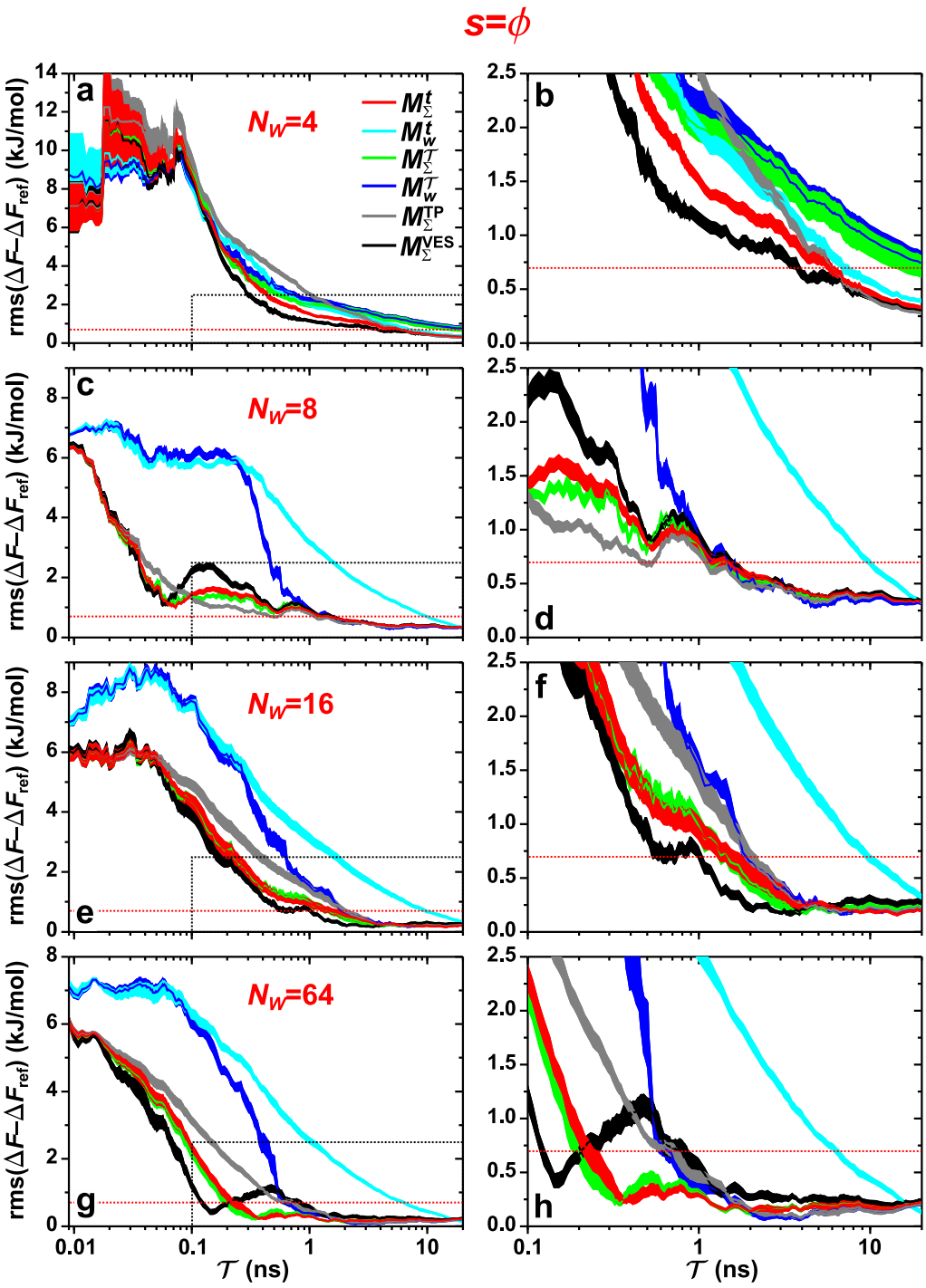

Fig. S1. Convergence curves for $s = \phi$ as in Fig. **3** (left panel) and the zoomed counterparts of Fig. **4** (right panel) but plotting the rms($\Delta F - \Delta F_{\rm ref}$) results along with the $\pm\sigma$ spread among the (a–d) 16, (e, f) 8 and (g, h) 6 independent simulations for each as-indicated reweighting and walker ensemble.

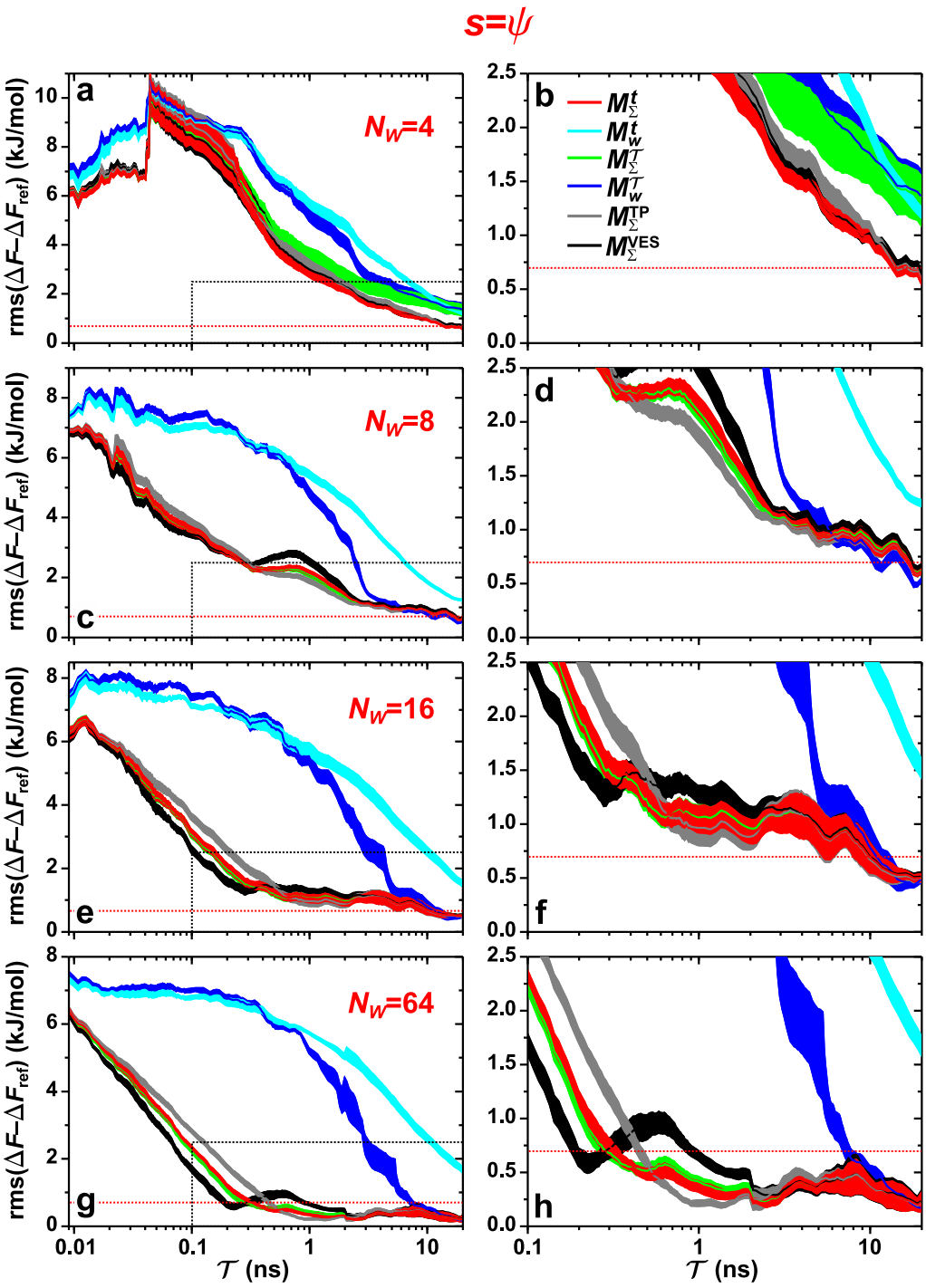

Fig. S2. Convergence curves for $s = \psi$ as in Fig. **3** (left panel) and the zoomed counterparts of Fig. **4** (right panel) but plotting the rms($\Delta F - \Delta F_{\rm ref}$) results along with the $\pm\sigma$ spread among the (a–d) 16, (e, f) 8 and (g, h) 6 independent simulations for each as-indicated reweighting and walker ensemble.

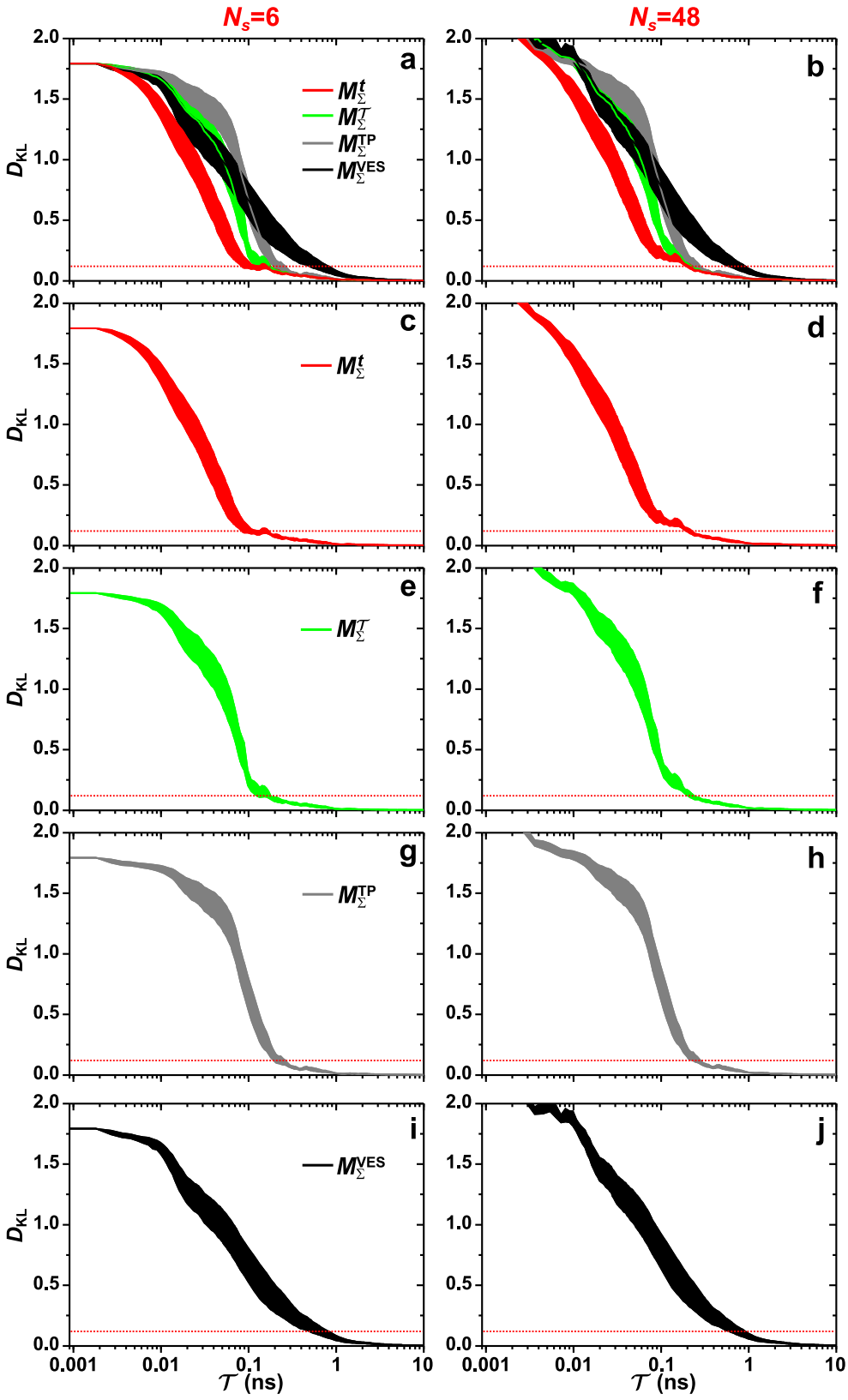

Fig. S3. (a, b) $D_{\mathrm{KL}}(\mathcal{T})$ convergence plots as in Fig. **6**b,c but also including the $\pm\sigma$ data variabilities among 32 independent simulations for each reweighting protocol for $N_s = 6$ (left panel) and $N_s = 48$ (right panel). (c–j) The result from each as-indicated reweighting method plotted separately.

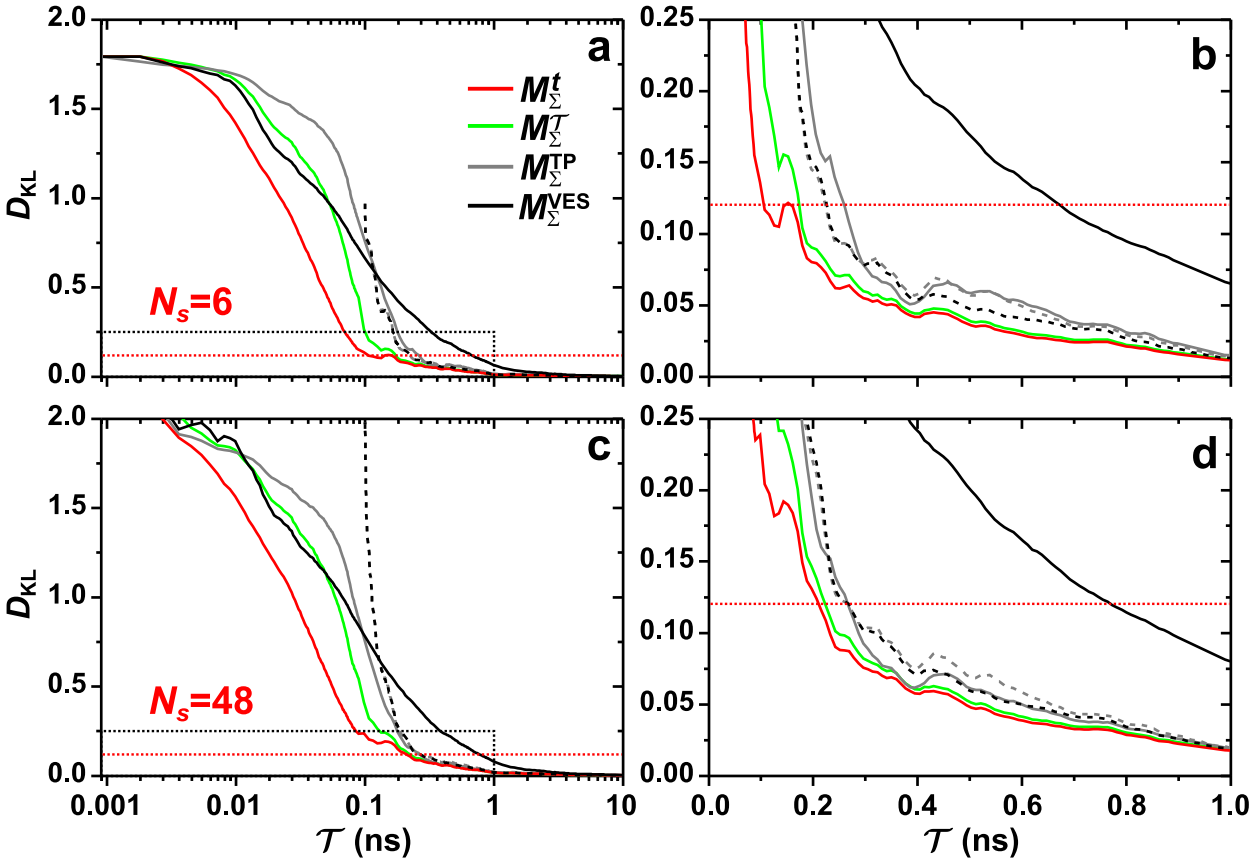

Fig. S4. (a, c) $D_{\mathrm{KL}}(\mathcal{T})$ convergence plots as in Fig. **6**b,c for (a) $N_s = 6$ and (c) $N_s = 48$ bins [eqn. (35)] but additionally including the results after excluding all data with $t < t_{\mathrm{min}} = 100$ ps for the M_Σ^{TP} and M_Σ^{VES} schemes in the reweighting, shown by dashed grey and black curves, respectively. (b, d) Zooms around the near-convergence region, i.e., the black dotted rectangles in (a, c). The value $t_{\mathrm{min}} = 100$ ps ensured optimal reweighting performance of both CV-integration-based M_Σ^{VES} and M_Σ^{TP} methods, after which they essentially indistinguishable convergence curves for simulation periods $\mathcal{T} \lesssim 400$ ps, and requiring $\mathcal{T} = 230$ ps ($N_s = 6$) and $\mathcal{T} = 270$ ps ($N_s = 48$) to reach convergence.