# Appendix C

# Quotients of means or random variables

In the gauge P representation, estimates of the expectation value of an observable $\widehat{O}$ are made by evaluating the expression (3.14) for a finite ensemble of $\mathcal{S}$ samples. This has the form of a quotient

$$\langle\widehat{O}\rangle \approx \overline{O} = \frac{\langle f\rangle_{\text{stoch}}}{\langle \text{Re}\{\Omega\}\rangle_{\text{stoch}}} = \frac{\langle f_N\rangle_{\text{stoch}}}{\langle f_D\rangle_{\text{stoch}}}. \tag{C.1}$$

Error estimates are made by subensemble averaging as explained in Section 3.3.2. While for dynamics calculations, one is assured of $\langle \text{Re}\{\Omega\}\rangle_{\text{stoch}} = 1$ in the limit $\mathcal{S} \to \infty$ at all times, so that the denominator can be ignored, thermodynamics calculations require the evaluation of both denominator $\langle f_D\rangle_{\text{stoch}}$ and numerator in (C.1). In particular, the error estimate (3.18) using subensembles requires some care.

The problem is that if there are not enough samples $s$ in (say) the $j$th subensemble, then the $j$th denominator average may have the wrong sign or be close to zero. The result is that this $j$th subensemble average $\overline{O}^{(j)}$ takes on very large or non-realistic values, which are nowhere near $\langle\widehat{O}\rangle$. Since their absolute values can be very large if the denominator was close to zero, these outlier $\overline{O}^{(j)}$ strongly influence the final error estimate. An example of such excessive error estimates is shown in Figure C.1, and typically goes in hand with spiking in the uncertainty $\Delta\overline{O}$.

Roughly, one expects that if there are $\mathcal{S}_E \approx 1000$ subensembles, then the far-

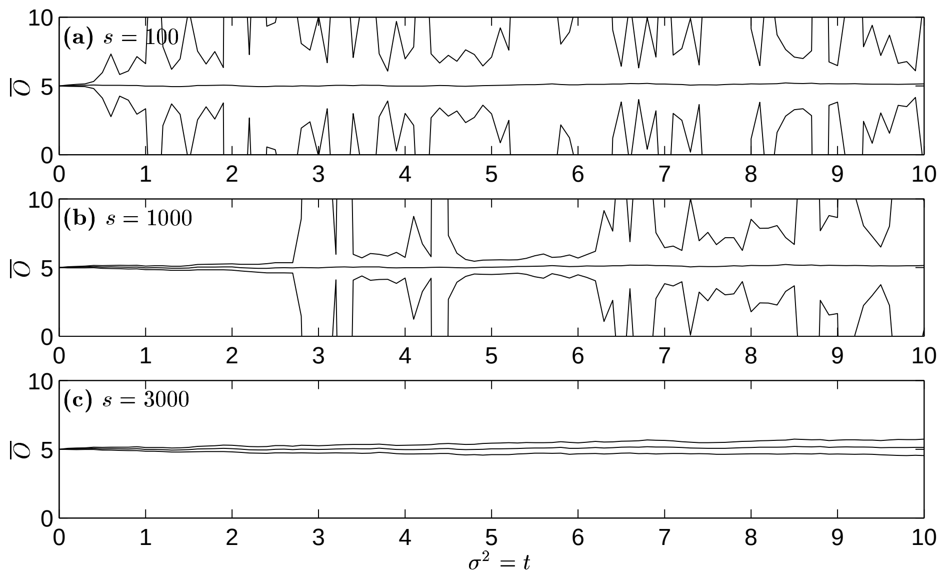


Figure C.1: **Error estimates with different subensemble sizes** $s$. SOLID lines show the means and error estimates obtained using (3.18) for the quantity $\overline{O} = 1/\left(\langle v\rangle_{\rm stoch} + 0.2\right)$, where $v(t)$ obeys the Brownian motion stochastic equation $dv = dW(t)$ ($dW(t)$ is a Wiener increment). The standard deviation of $v$ is $\sigma = \sqrt{t}$. In each simulation there were $\mathcal{S} = 10^5$ trajectories in all. Expression (C.2) suggests $s \gtrsim 2250$ is required at $t = 10$.

thermost outlier will be at about $3\sigma$ from the mean. Thus, excessive uncertainty is likely to occur once $3\sqrt{\text{var}\left[\overline{f}_D^{(j)}\right]} \gtrsim \langle f_D\rangle_{\rm stoch}$, where $\overline{f}_D^{(j)}$ is the denominator average for the $j$th subensemble. For subensembles with $s$ elements, this gives the rough limit

$$s \gtrsim 9\frac{\text{var}\left[f_D\right]}{\langle f_D\rangle^2_{\rm stoch}}. \tag{C.2}$$

Often the variation in the weight $\Omega = e^{z_0}$ is mainly due to a close-to-gaussian distributed $\text{Re}\{z_0\}$. In this case, using (7.39) and (7.40), the requirement (C.2) becomes

$$s \gtrsim 9\left(e^{\text{var}[\text{Re}\{z_0\}]} - 1\right). \tag{C.3}$$