APPENDIX: VARIANCE CORRECTION BETWEEN THE CASE OF FIXED PROBABILITY OF OBSERVATIONS AND FIXED NUMBER OF OBSERVATIONS

To quantify the effect on the variance of an acceptance ratio free energy estimate in the case of the restriction from fixed probability of forward and reverse measurements to fixed number of forward and reverse measurements, we first label the number of forward and reverse simulations in a single realization of the measurement, as m_F and m_R respectively.

If m_F and m_R are different from their expectation values n_F and n_R, then the free energy that will zero Eq. (11) of the main text will be off by the difference $kT\log(n_R/n_F) - kT\log(m_F/m_R)$ in the estimate. This is because we are effectively using the wrong correction factor M in from Eq. (11) of the main text for the numbers m_F and m_R. We see from Eq. (11) of the main text that deviations of m from n will change the value of M while $M - \Delta F$ remains fixed, and therefore change the estimate of ΔF.

Writing $p = n/n_{tot}$, and $r = (m-n)/n_{tot}$, where n and m stand for either n_F and n_R or m_F and m_R respectively, we can approximate this difference μ in the limit $r \ll p$ as:

$$\mu = kT\log\frac{(1-p)(p+r)}{p(1-(p-r))} = kT\log\left(1+\frac{r}{p-p^2-pr}\right) \approx \frac{r}{\beta p(1-p)} \tag{1}$$

Knowing the dependence of the estimated free energy on m_F, we must now try to deconvolute the variance over all possibilities with the variance at the expectation value for m_F and m_R given probability p, equivalent to finding the variance at fixed n_F and n_R.

If we have k sets of measurements with means $\mu_1, \mu_2, \ldots, \mu_k$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2$ and the fraction of the total number in each set is f_k, then the variance of the set of all these points can be expressed as:

$$\sum_{i=1}^{k}(f_i\sigma_i^2 + f_i\mu_i^2) - \left(\sum f_i\mu_i\right)^2 \tag{2}$$

In the case of our variance estimate, the sum is from 1 to n_{tot}, and the μ_i's are given by Eq. (1). The f_i's are the binomial weights, $C_m^n p^m(1-p)^{n_{tot}-m}$. By examining Eqs. (1) of the Appendix and (12) of the main text, we see that although deviations of m from n will change the value of ΔF, since $M - \Delta F$ remains constant, the variance of ΔF will not change. The σ_i are therefore all equal; we will call this value simply σ^2. This is the variance at each individual value of m for which we need to solve. The variance for the combined sets, denoted by σ^2_{Fisher}, is evaluated in Eq. (12) of the main text.

We note that the expectation value of m is simply $n_{tot}p$. This implies that the expectation value of r and therefore the expectation value of μ_i is 0, since r is simply $(m-n)/n_{tot}$, and so $\sum f_i\mu_i = 0$. We also note that variance of a binomial variable such as m is $n_{tot}p(1-p)$. Therefore the variance of r is simply $p(1-p)/n_{tot}$. Since the expectation value of r is zero, the average of r^2 is also $p(1-p)/n_{tot}$.

We therefore find that

$$\sum f_i\mu_i^2 = \frac{p(1-p)}{n_{tot}}\frac{1}{\beta^2p^2(1-p)^2} = \frac{1}{\beta^2 n_{tot}p(1-p)} \tag{3}$$

and therefore

$$\sigma^2 \approx \sigma^2_{Fisher} - \frac{1}{\beta^2 n_{tot}p(1-p)} \tag{4}$$

To complete the derivation, we then note that:

$$\frac{1}{n_{tot}p(1-p)} = \frac{1}{n_{tot}\frac{n_F}{n_{tot}}n_R n_{tot}} = \frac{n_F+n_R}{n_F n_R} = \frac{1}{n_F} + \frac{1}{n_R} \tag{5}$$

allowing us to obtain Eq. (13) of the main text.