2.2 Local Error Decomposition for Trapezoidal Integration

Estimating the uncertainty of a calculated quantity is critically important in many applications.[60] The calculation of ΔG using integration based methods such as TI is no exception. In such calculations, the value of the ensemble average of the derivative of the potential energy with respect to a predefined pathway ($\langle \partial V/\partial \lambda \rangle_\lambda$) is numerically integrated to produce a value of ΔG. Section 2.1 outlined a method based on the KS–statistic (KS_{SE} given by Eq. 8) which can be used to estimate the uncertainty in an ensemble average obtained by MD simulation. In this section we outline a general method for estimating the uncertainty of the Trapezoidal method (section 2.2.3) by considering both uncertainties in the points (section 2.2.1) as well as truncation error due to linear interpolation (section 2.2.2).

The definite integral of a function $f(x)$ represented by a vector of discrete points $\boldsymbol{y}$ can be estimated using the Trapezoidal rule:

$$\int_{x_0}^{x_N} f(x)dx \approx \frac{1}{2}\sum_{i=1}^{N}(y_i + y_{i-1})(x_i - x_{i-1}) \tag{15}$$

The Trapezoidal rule approximates the area under the curve between consecutive points using linear interpolation. Alternative numerical integration algorithms differ in the method of interpolation between points, Simpson's rule for example the interpolation is based on a 2nd order polynomial. All numerical integration algorithms contain an inherent element of uncertainty associated with the method of interpolation. This is discussed in more detail in section 2.2.2. If the discrete points y_i are exact representations of $f(x_i)$, and these points are appropriately spaced in order to reproduce the highest non-zero derivatives, higher-order interpolation methods will yield more accurate results. If, however, $f(x_i)$ is only approximated by y_i, any uncertainty within the points is amplified by the use of higher-order methods. It is for this reason we chose to use the Trapezoidal rule for numerical integration calculations rather than high-order methods such as Simpson's rule. Note that as discussed in section 2.2.3, if truncation errors are allowed to cancel during the summation of local error contributions this effectively incorporates 2nd order (Simpson's method) contributions for functions that contain both concave and convex regions.

2.2.1 Trapezoidal Rule Point Uncertainty Propagation (Known Unknowns)

For the calculation of ΔG using integration based methods we have the case where $f(x)$ is approximated by a series of points y_i which have associated uncertainties σ_{yi}. We therefore would like to analytically calculate how these errors propagate through the Trapezoidal rule and contribute to the uncertainty of the integral. If we assume that the errors corresponding to each of the points ($y_i \pm \sigma_{yi}$) are independent and normally distributed, we can then apply the simplified form of the general Gaussian error propagation formula[61]:

$$f(x \pm \sigma_x, y \pm \sigma_y, ...) = f(x, y, ...) \pm \sqrt{\left(\frac{\partial f}{\partial x}\sigma_x\right)^2 + \left(\frac{\partial f}{\partial y}\sigma_y\right)^2 + ...} \qquad (16)$$

where f is an arbitrary function of variable (x, y, …) with associated independent and normally distributed errors (σ_x, σ_y, ...). The error due to uncertainty in the points of applying the Trapezoidal rule can be obtained by applying Eq. 16 to Eq. 15. Which results in:

$$Error = \frac{1}{2}\sqrt{(x_1 - x_0)^2\sigma_{y_0}^2 + \sum_{i=1}^{N-1}(x_{i+1} - x_{i-1})^2\sigma_{y_i}^2 + (x_N - x_{N-1})^2\sigma_{y_N}^2} \qquad (17)$$

Note that this expression for the total propagation error is a sum of contributions from each y_i point, it is therefore trivial to identify which points contribute most to the overall error.

2.2.2 Trapezoidal Rule Truncation Error (Unknown Unknowns)

When applying the Trapezoidal rule, the errors associated with the points $\boldsymbol{y}$ are a well-defined source of uncertainty. They contain specific information about the value of the function f at a given point x. Another source of error in all numerical integration methods results from implicit assumptions made regarding the form of the function between the evaluated data points i.e. the interpolation method. This source of error, often referred to as truncation error, is not so well-defined and therefore associated with an addition degree of uncertainty. This additional uncertainty is due to the assumption that the discrete points are sufficiently close together to accurately reproduce the highest non-zero derivative of the form of the function being integrated. However, in many applications the reason for using numerical integration is precisely because the form of the underlying function is unknown. Truncation error is therefore much more difficult to quantify accurately than the propagation of uncertainties in the points.

An error analysis of the Trapezoidal rule using a Taylor series expansion—which can be found in most numerical methods text books—results in a truncation error estimate for the interval between two point *a* and *b* given by:

$$Error \approx -\frac{(b-a)^3}{12} f''(\xi) \tag{18}$$

where $f''(\xi)$ is the second derivative over the interval [*a, b*]. The second derivative over this interval can be calculated numerically with either the forward, backward or central difference methods, and thus the 2nd order truncation error for each individual interval can be estimated. There are two primary approximations implicit in this error estimate: (1) that the calculation of the 2nd order term from the discrete points is accurate (i.e. calculating $f''(\xi)$ by numerical differentiation), and (2) that the underlying function can be accurately represented by a 2nd order polynomial on the interval [*a, b*]. Both cases are dependent on whether the discrete points being integrated are sufficiently close together to reproduce the underlying function. Note that in principle the truncation error given by Eq. 18 corresponds to the difference between applying the Trapezoidal and Simpson's rule.

2.2.3 Combined Trapezoidal Error Estimate

The combined estimate of the integration error associated with applying the Trapezoidal rule on a series of points which themselves contain errors is given by the sum of the results from the previous two sections, specifically:

$$\begin{aligned} Error = & \frac{1}{2}\sqrt{(x_1 - x_0)^2\sigma_{y_0}^2 + \sum_{i=1}^{N-1}(x_{i+1} - x_{i-1})^2\sigma_{y_i}^2 + (x_N - x_{N-1})^2\sigma_{y_N}^2} \\ & + \left| \sum_{i=0}^{N-1} -\frac{(x_{i+1} - x_i)^3}{12} f''(\xi) \right| \end{aligned} \tag{19}$$

Since the accuracy of the forward, backward and central difference methods for estimating $f''(\xi)$ over the interval $[x_{i+1}, x_i]$ are the same, the choice of method should have no systematic effect. However, unless the data points are perfectly symmetric they will give slightly different results. In practice, we calculate the sum of truncation errors with both forward and backward difference methods and take the larger of the two.

Note that the truncation error (given by Eq. 18) is not entirely independent of the point uncertainty error (Eq. 17) since the points themselves are used to estimate $f''(\xi)$ over the interval $[x_{i+1}, x_i]$.

Using the technique applied in section 2.2.1, it is straight forward to propagate point uncertainties to obtain an error estimate in $f''(\xi)$. However, since some of the same points are used to estimate the uncertainty in neighbouring values of $f''(\xi)$, correlations between the errors of neighbouring intervals need to be accounted for. This becomes progressively more complex and the practical utility of considering such effects is unclear.

Summing the truncation errors and then taking the absolute value (Eq. 19) allows for cancelation of errors and therefore implicitly accounts for 2nd order contributions in some cases i.e. in cases where the function contain both concave and convex regions. The drawback of this approach however, is that given the uncertainty of the truncation error predications themselves, the errors can inappropriately cancel purely by chance and lead to significant underestimation of the total error. An alternative approach would be to add the absolute values of the individual interval truncation errors—which is the default behaviour for purely concave or convex functions since there can be no cancelation of errors—but this leads to significant over estimation of the error in typical cases. An initial analysis of the performance of Eq. 19 for estimating the error in solvation free enthalpy calculations using TI (details provided in the following section) found that in approximately 8% of cases the error was underestimated. While a failure of 8% may be adequate in many cases, in some applications, for example when parameterising force fields, a more robust estimate of the uncertainty is required.

2.2.4 Additional Uncertainty Heuristic

The addition of a simple heuristic to the integration uncertainty given by Eq. 19 dramatically reduced the underestimate rate without the significant increase in computational cost associated with not allowing for cancelation of errors i.e. summing the absolute value of the truncation error. The heuristic correction consists of adding the maximum interval truncation error found with either the forward or backward difference methods:

$$Error\ Heuristic = \max\{Er_i^{forward}, Er_i^{backward}\}_{i=0}^{N-1} \tag{20}$$

where $Eri^{forward}$ and $Eri^{backward}$ are the interval truncation errors calculated using Eq. 18 where $f''(\xi)$ over the interval $[x_{i+1}, x_i]$ is estimated using the forward and backward difference methods respectively. Thus a more robust Trapezoidal rule error estimate is given by:

$$Error = \frac{1}{2}\sqrt{(x_1 - x_0)^2\sigma_{y_0}^2 + \sum_{i=1}^{N-1}(x_{i+1} - x_{i-1})^2\sigma_{y_i}^2 + (x_N - x_{N-1})^2\sigma_{y_N}^2} + \max\left\{\left|\sum_{i=1}^{N-1} Er_i^{forward}\right|, \left|\sum_{i=1}^{N-1} Er_i^{backward}\right|\right\} + \max\{Er_i^{forward}, Er_i^{backward}\}_{i=0}^{N-1} \quad (21)$$

An analysis on the relative performance of Eq. 19 and Eq. 21 is provided as a part of the following section. Eq. 21 was used for all the calculations of ΔG^{solv} by TI presented in Chapter 6 and Chapter 7.

2.3 Automated Thermodynamic Integration Protocol

The calculation of ΔG by TI involves the numerical integration of the ensemble average of the derivative of the potential with respect to a predefined pathway ($\langle \partial V/\partial \lambda \rangle_\lambda$), described in more detail in section [introduction XX]. The values of $\langle \partial V/\partial \lambda \rangle_\lambda$ are typically obtained by performing a series of fixed length MD simulations at multiple different values of λ which are uniformly distributed between λ=0 and λ=1.[18, 22, 38, 62] In some instances, well converged values of $\langle \partial V/\partial \lambda \rangle_\lambda$ are achieved by running comparatively long simulations with significant portions of the simulation data discarded as equilibration. Using an even distribution of λ-points is also sub-optimal as either too few simulations will be performed at λ-points in regions where $\langle \partial V/\partial \lambda \rangle_\lambda$ has high curvature, or too many in regions where only a few points will suffice i.e. where the change in gradient is small.

A fully automated protocol for the calculation of ΔG using integration based approaches has been developed. The method utilises the ensemble average error analysis outlined in section 2.1 in conjunction with the numerical integration error analysis described in section 2.2. The explicit consideration of the individual sources of error allows the uncertainty in the value of ΔG to be iteratively refined to a specified precision. This is achieved through the extension of equilibrium MD simulations or the addition of λ-points according to the largest sources of uncertainty. The following two sections contain a description of the protocol along with validation and analysis.

2.3.1 Calculation of ΔG^{solv} by TI: Automation by Error Minimisation

Calculations of ΔG^{solv} by TI for small molecules were initialised with 11 MD simulations of the solute in the solvent at equally spaced λ-points between 0 and 1 inclusive ($<\partial V/\partial\lambda>_{\lambda,solvent}$), using the simulation setup described in section 5.1.5. The number of initial λ-points is somewhat arbitrary and is a trade-off between efficiency and robustness as too few λ-points increases the risk of a significant underestimation of the integration truncation error, while too many initial points result in unnecessary simulations being performed. The MD simulations at each λ-point were run initially for 500 ps. The uncertainty and convergence of all ensemble averages were assessed using the scheme described in section 2.1.6; specifically, the KS_{SE} method was used to identify the required equilibration time, estimate the standard error and to measure the confidence with which convergence to within a particular target threshold had been achieved. Initial ensemble averages were converged to within 5 kJ mol^{-1} with simulations extended by 500 ps until a CR_{score} of at least 1 was obtained. A CR_{score} of greater than or equal to 1 corresponds to the case where the standard error values estimated by the KS_{SE} method are below the target error for at least half of the simulation. The final ensemble average error estimates were calculated with any equilibration region removed. These error estimates were used as the point uncertainty values (σ_{yi}) when the combined integration error was determined using Eq. 21.

The calculation of ΔG^{solv} for molecules which contain intermolecular non-bonded interactions requires the calculation of $<\partial V/\partial\lambda>_{\lambda}$ in vacuum ($<\partial V/\partial\lambda>_{\lambda,vac}$) as well as in solvent. Multiple vacuum calculations were performed at the same λ-points as each solvent calculation. These were initialised from evenly spaced solute configurations extracted from the solute in solvent trajectory by simply removing all solvent molecules. 20 vacuum simulations were initially run for 1 ns and then extended by an additional 1 ns until each ensemble average converged to within 5 kJ mol^{-1} with a $CR_{score}>= 1$.

Final values of ΔG^{solv} were calculated by applying the Trapezoidal rule to numerically integrate the difference curve given by $<\partial V/\partial\lambda>_{\lambda,vac}-<\partial V/\partial\lambda>_{\lambda,solvent}$ as per Eq. [introduction XX]. The estimates of the error in ΔG^{solv} were obtained by applying Eq. 21 to the same difference curve. In some cases, the uncertainty obtained by performing the calculation with the initial 11 points was within the target uncertainty. These were predominantly rigid molecules in apolar solvents. For the majority of molecules however the initial error estimate was larger than the target uncertainty and the calculation required additional points and/or the reduction in the uncertainty in ensemble average in order to meet the overall target error.

The overall error was reduced by analysing the individual sources of error in Eq. 21 (i.e. the errors associated with each of the intervals and points). Then points were either added or the simulations extended depending on the largest sources of error. To optimise the use of computational resources each refinement iteration one could simply target the largest source of uncertainty. However, the refinement can be efficiently parallelised by estimating the uncertainty reduction resulting from the addition of a point or the extension of a simulation. This was achieved by calculating the difference between the current and target error, then adding points or extending simulations until the error reduction estimate indicates the new error will be below the target. The conservative nature of the error reduction estimate will ultimately determine the number of iterations required to achieve the target uncertainty. It can be shown analytically for the Trapezoidal rule that the interval truncation error is on average reduced by ¼ of the previous error by the addition of a point at the centre of the interval. For simplicity we also assume this to be the case for extending a simulation and reducing the target error from 5 kJ mol^{-1} to 2 kJ mol^{-1}.

Since the uncertainty in each point of the $\langle \partial V/\partial \lambda \rangle$ difference curve is a combination of the solvent and vacuum contributions, the larger of the two was extended if that point was chosen for error reduction on a given refinement iteration. Vacuum and solvent simulations were extended by 1 ns and 500 ps respectively. The target ensemble average error was also decreased from 5 kJ mol^{-1} to 2 kJ mol^{-1} to ensure that the ensemble error was significantly reduced with the CR_{score} condition remaining at >= 1.

Additional points were always located in the middle of the interval identified for error reduction. The solvent simulations were run first and initialised from the final configuration of either of the two closest λ-points, chosen at random with equal probability. Additional λ-points were run for 500 ps and extended by a further 500 ps until the ensemble average converged to within 5 kJ mol^{-1} with a CR_{score} >= 1. Once the solvent simulations had converged, simulations in vacuum were performed as described above. This protocol was repeated until the overall target error in ΔG^{solv} was obtained.

Figure 2.9 shows 4 example calculations of ΔG^{solv} by TI: the solvation free energy of ethanol, 1-propylguanidine and nonane in SPC water[63] (ΔG^{H2O}), and the solvation free energy of chlorobenzene in the GROMOS 53A6 hexane model[64] (ΔG^{HEX}). The values for $\langle \partial V/\partial \lambda \rangle$ in vacuum and solvent are shown in orange circles and black squares respectively, the difference curve is shown in blue-triangles.

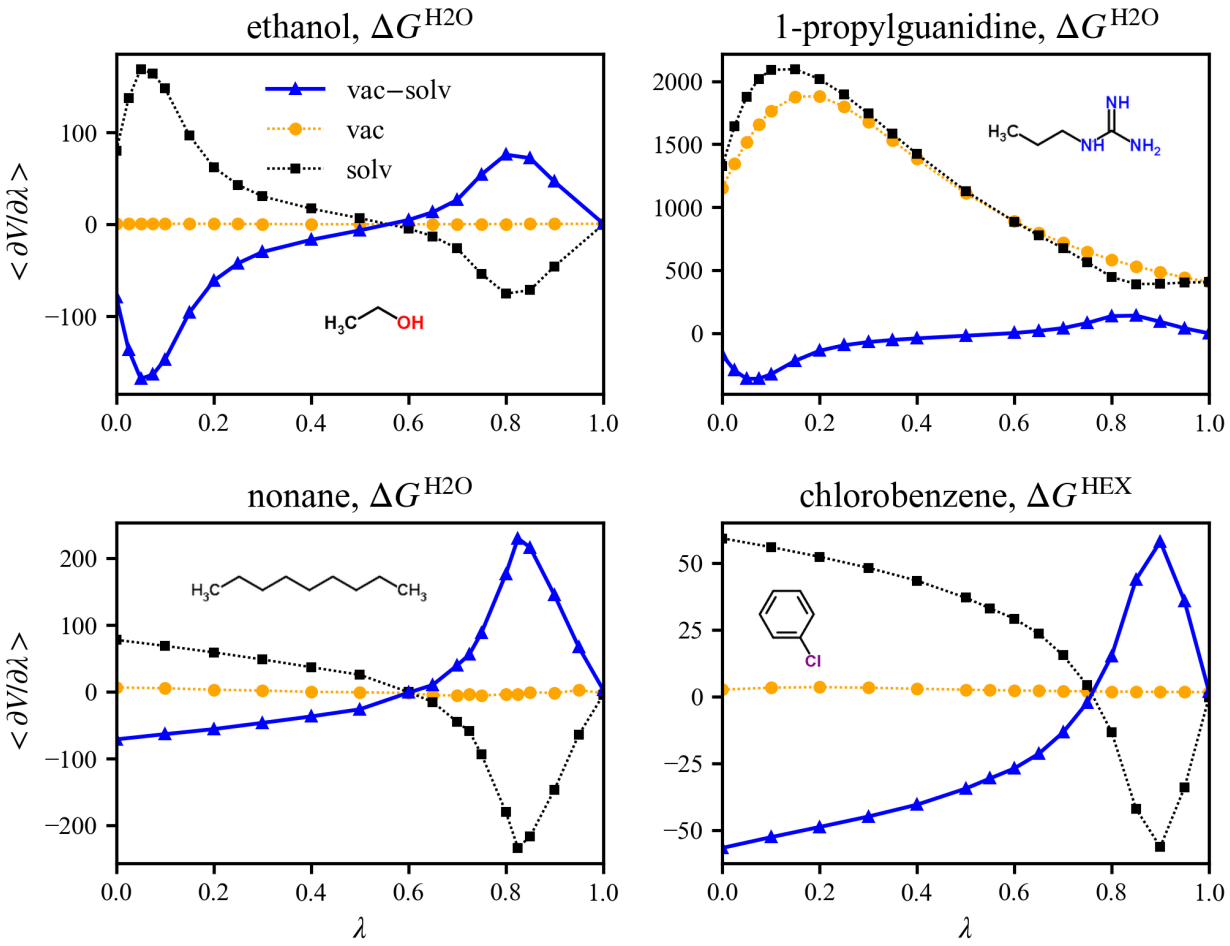

Figure 2.9. Example $<\partial V/\partial\lambda>$ with respect to λ curves used in the calculations of ΔG^{H2O} and ΔG^{HEX} by TI. $<\partial V/\partial\lambda>$ values are shown for calculations in solvent (solv, black squares), vacuum (vac, orange circles), their difference ($<\partial V/\partial\lambda>_{vac} - <\partial V/\partial\lambda>_{solvent}$) is shown in blue triangles. All values have been converged to within 1 kJ mol^{-1}.

There is a stark contrast in the magnitude of $<\partial V/\partial\lambda>$ between 1-propylguanidine and the other 3 cases, due to the significant intermolecular non-bonded interactions. Note however that the magnitude of the difference curve is comparable to the other cases, this illustrates the benefit of refining the error of the difference curve directly rather than the solvent and vacuum terms independently. Also clear from Figure 2.9 is the difference in the shape of the curves for cases with significant electrostatic interactions (ethane and 1-propylguanidine in water) and those without (nonane in water and chlorobenzene in hexane). Those with significant electrostatic interactions contain an additional peak in the solvent curve between 0.0 and 0.2 and therefore require more points to accurately integrate. To achieve an error estimate below 1 kJ mol^{-1} for ethane and 1-propylguanidine required 20 and 21 points, refined with 3 and 4 iterations respectively. As compared to the nonane and chlorobenzene cases which required 17 and 16 points, refined with 3 and 2 iterations respectively, to achieve the same target error.

The examples shown in Figure 2.9 demonstrate that a different number and distribution of points are required in order to obtained a similar precision in ΔG^{solv} for individual cases. By utilising an adaptive error refinement scheme, we are able to accommodate these differences and therefore significantly improve both calculation efficiency and precision.

2.3.2 Validation and Analysis

The performance of the automated TI protocol was assessed by analysing 1201 ΔG^{solv} calculations performed on 501 unique molecules in either water, hexane and cyclohexane solvents. (The specific results of these calculations are presented as a part of Chapter 3, Chapter 6 and Chapter 7.) The analysis consisted of taking the final calculated values as a reference (ΔG_{ref}) and comparing these reference values to the intermediate values (ΔG_{interm}) obtained during error refinement. All values of ΔG_{ref} were converged to have an estimated uncertainty of either 1 or 0.5 kJ mol^{-1}. The validation of any method by comparing intermediate with final values is not entirely foolproof. This is because any underlying bias will cancel out and give a false impression of the accuracy. It does however demonstrate internal consistency which is a necessary, albeit not a sufficient, validation property.

Figure 2.10 shows a plot of ΔG_{interm} as a function of ΔG_{ref} for 1201 ΔG^{solv} calculations performed using the automated protocol outlined in the previous section (2.3.1). The error estimates for ΔG_{interm} (Error$_{pred}$) calculated during each iteration (with either Eq. 19 or 21) are shown in colour. Figure 2.10 shows that the 1201 calculations contain molecules which have a wide range of ΔG^{solv} values, from less than –70 kJ mol^{-1} to about 15 kJ mol^{-1}. Overall there is a correlation between Error$_{pred}$ (shown in colour) and the distance to the blue line corresponding to perfect agreement between ΔG_{interm} and ΔG_{ref}. This is analysed in more detail below.

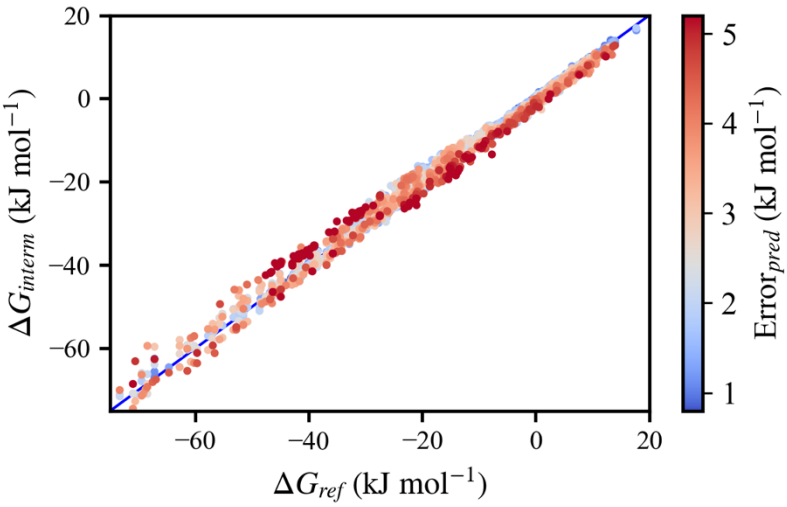

Figure 2.10. Comparison of 1201 final ΔG^{solv} values calculated by TI (ΔG_{ref}) with the values predicted during error refinement (ΔG_{interm}). The estimated overall error in ΔG_{interm} ($Error_{pred}$) for each iteration is shown in colour. The blue line represents a one-to-one agreement between ΔG_{ref} and ΔG_{interm}.

The error predictions ($Error_{pred}$) shown in Figure 2.10, strongly indicate that there is a systematic bias for $Error_{pred}$ values greater than about 2.0 kJ mol^{-1}. This is evident from the concentration of red dots below the blue line for ΔG^{solv} greater than about –25 kJ mol^{-1} and below the blue line for ΔG^{solv} less than about –25 kJ mol^{-1}. Specifically, molecules with ΔG^{solv} greater than –25 kJ mol^{-1} converge to the reference value from below, while molecules with ΔG^{solv} less than –25 kJ mol^{-1} converge to the reference from above. This result is due to the conserved nature of the shape of the $\langle \partial V/\partial \lambda \rangle_{vac} - \langle \partial V/\partial \lambda \rangle_{solvent}$ curve for molecules with similar ΔG^{solv} values. As noted above, molecules that have significant electrostatic interactions with the solvent have an additional region of high curvature in $\langle \partial V/\partial \lambda \rangle_{solvent}$ between λ=0 and λ=0.2, which has the opposite sign to the conserved peak between λ=0.6 and λ=1. All calculations in this dataset which had a ΔG^{solv} less than about –25 kJ mol^{-1} were for solvation in water. The molecules in these cases are all highly polar and thus the interval truncation error is dominated by the region of high curvature between λ=0 and λ=0.2 which has the opposite sign to the other cases which are dominated by truncation errors between λ=0.6 and λ=1. Note that the shape of the $\langle \partial V/\partial \lambda \rangle_{vac} - \langle \partial V/\partial \lambda \rangle_{solvent}$ curve is dependent on the specific soft-core[40] interaction functions. The particular bias identified here is therefore not transferrable to cases for which a different λ dependence was used. However, unless the curve is symmetric or the integration truncation errors are very small, some degree of bias in such calculations is unavoidable.

The bias in errors for ΔG^{solv} calculated by TI is of particular concern when using such calculations to parameterise force fields. As discussed in more detail in the following chapter, systematic errors cannot be considered in the same way as unbiased errors as they have different distributions and therefore different cancelation properties. Note however that for Trapezoidal rule error predictions less than ~2.0 kJ mol^{-1} this bias is small and significantly less that the contributions from the point uncertainty.

A more detailed analysis of the error predictions was performed by comparing the intermediate errors predicted on each refinement iteration ($Error_{pred}$) with the difference between the corresponding intermediate ΔG^{solv} value (ΔG_{interm}) and the final ΔG^{solv} value (ΔG_{ref}) which satisfied the target uncertainty of either 1 or 0.5 kJ mol^{-1}. That is, a reference error was defined to be $Error_{ref} = |\Delta G_{interm} - \Delta G_{ref}|$, and this quantity was compared to $Error_{pred}$ calculated during refinement iterations. The results of this analysis, with and without the addition of the heuristic correction (described in section 2.2.4), are shown in Figure 2.11. The corresponding summary statistics are presented in Table 2.2.

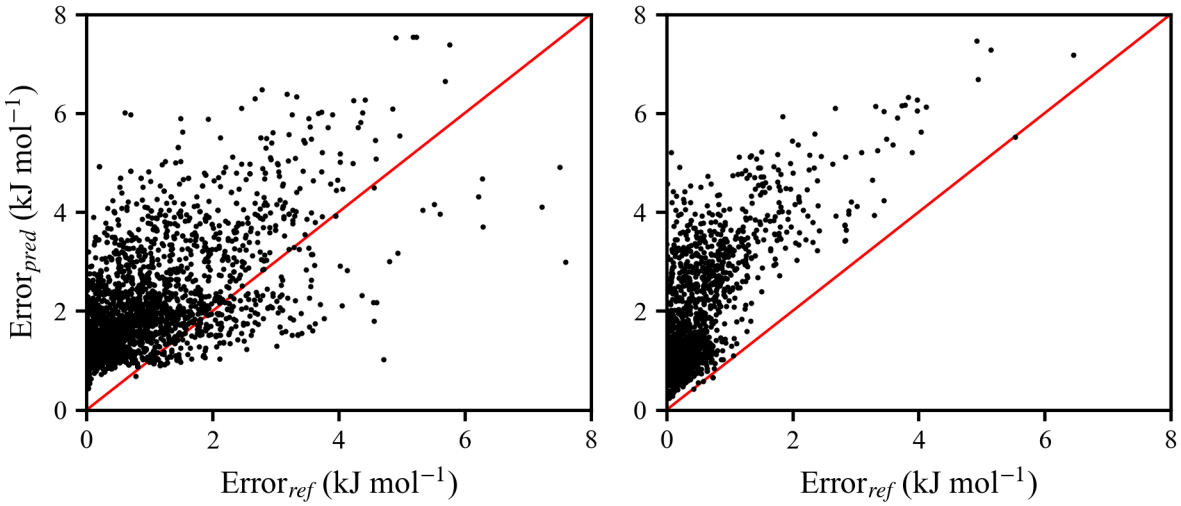

Figure 2.11. Trapezoidal rule integration error comparison between intermediate error predictions ($Error_{pred}$) and the difference between intermediate and final ΔG^{solv} values ($Error_{ref}$) when calculated using Eq. 19 (left) and Eq. 21 (right). The target error for all cases shown in the left panel was 1 kJ mol^{-1} while the right panel contains a combination of 1 and 0.5 kJ mol^{-1}.

The error prediction analysis shown in the left panel of Figure 2.11 correspond to ΔG^{solv} calculations performed with integration error estimates predicted using Eq. 19 (without heuristic correction), first row of Table 2.2. It was found that in the vast majority of instances $Error_{pred}$ was

slightly overestimated. This is likely due to two factors: (1) always taking the largest of forward and backward difference methods when summing interval truncation errors, and (2) that the interval and point error contributions are simply added and therefore not allowed to cancel. In about 2% of cases however the error was underestimated by more than 1 kJ mol^{-1} and in one extreme case by 5.4 kJ mol^{-1}. These failures were due to inappropriate cancellation of interval truncation errors. A simple heuristic correction was proposed to address these cases, this consisted of the addition of the maximum interval truncation error to the final sum (Eq. 21), see section 2.2.4 for details. The addition of this term eliminated all such failures as can be seen from the right panel of Figure 2.11 and the second row of Table 2.2. These results suggest that Eq. 21 is both conservative and robust and appears to give reliable estimate of the uncertainty in ΔG^{solv} calculated by TI.

Table 2.2. Summary of error prediction results for Trapezoidal rule error analysis methods. Under estimate (est.) is defined as $Error_{pred} < |\Delta G_{pred} - \Delta G_{ref}|$.

	$N_{molecules}$	$N_{TI\ calc.}$	N_{points}	$N_{predictions}$	$N_{under\ est.}$	under est. rate (%)	under est. > 1 kJ mol^{-1}	max. under est. (kJ mol^{-1})
basic error analysis, Eq. 19	501	575	18 ± 3	2,805	214	7.6	46	5.4
+ error heuristic, Eq. 21	110	626	16 ± 4	1,590	5	0.3	0	0.1

As noted above, comparing intermediate with final values does not provide an ideal validation of the method. A more comprehensive validation would involve the calculation of a large numbers of ΔG^{solv} values for a diverse set of molecules and solvent combinations using many more λ-points than required with each $\langle \partial V/\partial \lambda \rangle_{\lambda}$ value calculated from an extensive MD simulation. However, such a validation would be extremely computationally expensive and is most likely not warranted considering that the error analysis has been almost entirely analytically derived.

An analysis was also performed on the length and distribution of MD simulations used in the calculation of ΔG^{solv} by TI. This analysis was performed on 14707 individual $\langle \partial V/\partial \lambda \rangle_{\lambda,solvent}$ simulations taken from 845 ΔG^{solv} calculations that were performed on 267 unique molecules. This data is the subset (for which all MD simulation data was available) of the results used above. Figure 2.12 shows the distribution of simulation length required to obtain appropriately converged ensemble averages as dictated by the automated protocol. The distribution has a clear exponential dependence with over half the ensembles converging within the initial 500 ps simulation. This suggests that significant efficiency may be gained by reducing the initial simulation length. Figure 2.12 also shows that ~10% of ensemble averages require simulations of 2 ns or longer to appropriately converge.

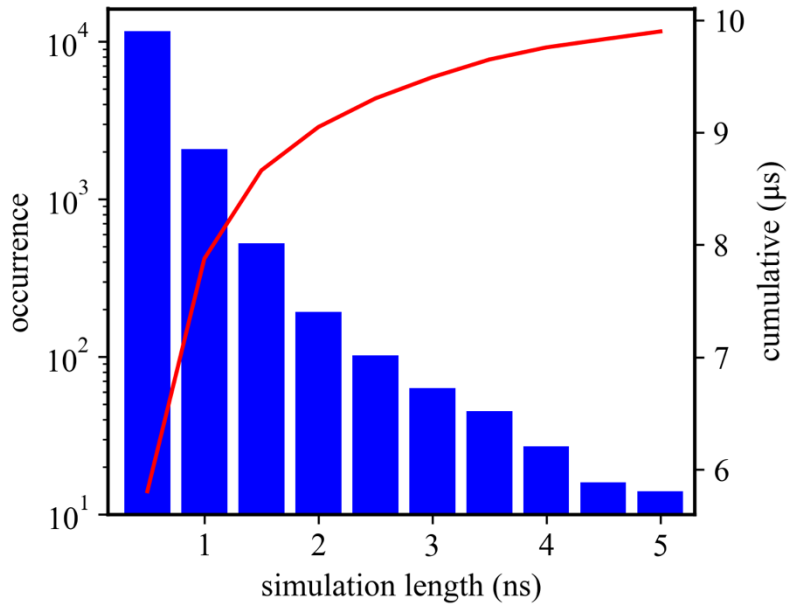

Figure 2.12. Distribution of individual $\langle \partial V/\partial \lambda \rangle_{\lambda,solvent}$ simulation lengths. Analysis was performed on 14707 individual simulations taken from 845 ΔG^{solv} calculations.

We also found that the MD simulation length required to obtain converged $\langle \partial V/\partial \lambda \rangle_{\lambda,solvent}$ values was strongly dependent on the value of λ. Figure 2.13 shows the average converged MD simulation length as a function of λ for the same set of 845 ΔG^{solv} calculations used to generate Figure 2.12. Figure 2.13 shows that to obtain converged ensemble averages for values of λ between 0.7 and 0.9 requires, on average, more than double the simulation time of other λ values. This concentration of simulation time around λ=0.8 is due to the inherent instability of the system as solvent molecules begin to enter the molecular cavity with greater probability i.e. as the repulsive Lennard-Jones term decreases. This causes an increase in the fluctuations of $\partial V/\partial \lambda$ and therefore a longer time series is required to obtain a reliable average.

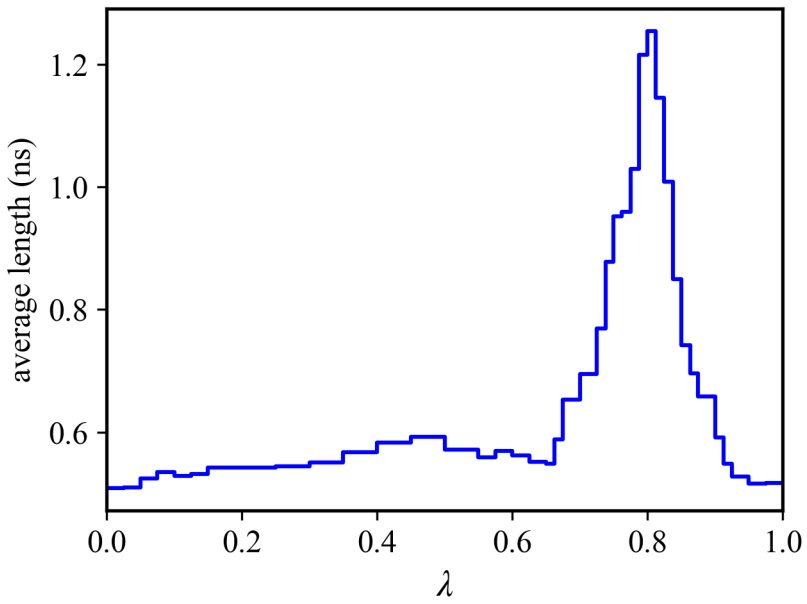

Figure 2.13. Distribution of average simulation length as a function of λ. Analysis was performed on 14707 individual simulations taken from 845 ΔG^{solv} calculations.

For this 845 ΔG^{solv} calculation dataset, the average total simulation time required to obtain an overall uncertainty within 1 kJ mol^{-1} was 12 ± 5 ns. However, some took up to ~30 ns to obtain ΔG^{H2O} within 1 kJ mol^{-1}, such cases were for large and/or flexible molecules e.g. Pyrene and 2,2,5-Trimethylhexane. The total simulation time required was found to be highly dependent on the particular molecule and solvent combination. These results clearly show that it is either inefficient or inaccurate to run uniform length MD simulations when calculating ΔG^{solv} by TI.

In this chapter we have presented a general method for estimating the uncertainty in the calculation of ΔG by integration based approaches. We have also shown that through the rigorous analysis of the specific sources of error, the uncertainty in ΔG can be systematically refined to an arbitrary precision by targeted minimisation of the error in specific contributions. The method was demonstrated by analysing the results from over 1000 fully automated TI calculations of ΔG^{solv}, performed on a wide range of molecules in multiple solvents. ΔG^{solv} error estimates were found to be self-consistent and conservative, since intermediate predictions significantly overestimated the uncertainty with respect to the final estimated values. The method was also found to adapt to the convergence requirements of particular λ-values as well as different molecule and solvent combinations. Overall the approach outlined is a much more efficient and reliable approach to the calculation of ΔG by TI than performing a pre-defined set of fixed length MD simulations at evenly distributed points between λ=0 and λ=1.