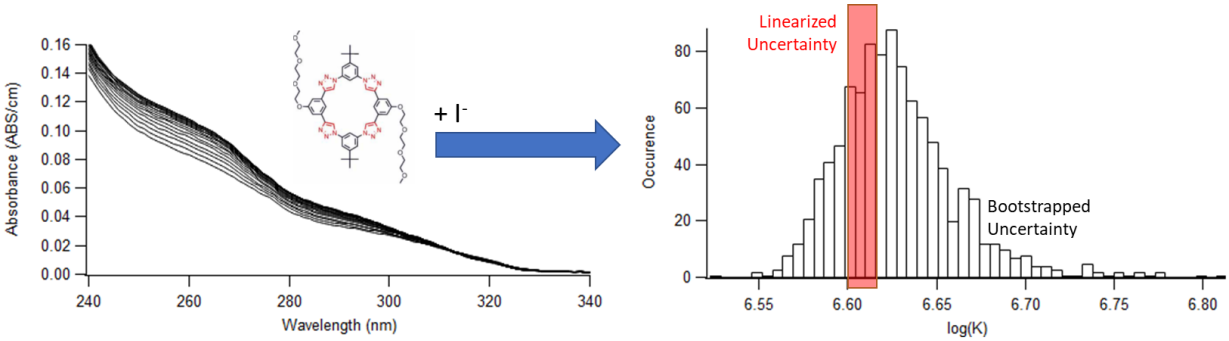
# Bootstrap methods for quantifying the uncertainty of binding constants in the hard modeling of spectrophotometric titration data

Nathanael P. Kazmierczak, Joyce A. Chew, Douglas A. Vander Griend, Calvin University




**Keywords**

Bootstrap, Multiway Data Analysis, Binding Constant, Confidence Limits, Hard Modeling, Uncertainty

**Abstract**

Equilibrium hard modeling of spectrophotometric titration data with binding parameters (ΔG° or logK values) involves nonlinear mathematical relationships and correlated experimental uncertainties. Therefore, uncertainty quantification techniques based on standard error computation substantially underestimate the true error for the calculated binding parameters. We show that the bootstrapping technique can provide accurate uncertainty quantification with no *a priori* knowledge of experimental error levels. Monte Carlo studies on simulated data show that bootstrapping the chemical solutions, whether on the data or residuals, handles absorbance error, transmittance error, and composition error well, producing asymmetric confidence intervals that correctly assess the true uncertainty. Additionally, stock solution error is handled well if it is present with other forms of error. Confidence interval bands for molar absorptivity curves can likewise be calculated. Analogous bootstrapping studies on real

datasets confirm that the 95% confidence intervals match the variance observed from experimental replicates, though bootstrapping on the residuals should be used for smaller datasets. Bootstrapping along the titration axis should be used to estimate uncertainty whenever binding parameters are ascertained from titration datasets.

**Introduction**

Binding constants are among the most important parameters in supramolecular and coordination chemistry [1]. They describe the fundamental equilibrium processes behind self-assembled nanostructures [2], chemical sensors [3,4], ion sequestration systems [5], and pharmaceutical activity [6], among many other applications. Interactions between multiple equilibrium binding events provide the pathway for realizing rational chemical design of complex systems on the nanoscale [7]. This nonlinear behavior permits chemical equilibrium systems to display a wide range of tunable, complex properties [8].

Unlike soft modeling [9], hard modeling of data aims to describe a system according to a parameterized physicochemical model and provides a powerful analytical technique for accessing binding constants from equilibrium spectrophotometric titrations [10], The simultaneous use of multiple wavelengths of absorbance data enables superior resolution over single-wavelength analysis [11]. Composite absorbance data is deconvoluted into molar absorptivity curves for multiple species according to the Beer-Lambert Law [12]. This can be conveniently depicted in matrix form as shown in Figure 1. By enforcing both mass balance and mass action, the binding constants become the key parameters defining the thermodynamic model [13]. Several computer programs exist for such analyses [14,15,16,17,18,19,20]. Hard modeling computes the optimal binding constants, equilibrium concentrations, and molar absorptivity values that best reproduce the raw absorbance data. The quality

of the fit is quantified by the root-mean-square error (RMSE) [21], the point-by-point residual difference between the modeled and measured absorbance data. The mathematical procedure of hard modeling has been described in detail in previous literature [22,23].

The least-squares fitting of the model to the data finds the optimal binding constant values in a matter of minutes if not seconds. Much of the error in the original data is sequestered in the matrix of residuals, but some remains embedded in the model. This implies that there is uncertainty in the optimal values for the binding constants, which must be quantified accurately. Importantly, the mathematical relationship between the binding constants and the data is very nonlinear [24]. For extremely exergonic or endergonic binding regimes, a shift in a binding constant will have vanishingly small impact on the modeled data. This property makes it challenging to accurately quantify the uncertainty in the binding constant values in the presence of experimental errors, as detailed below.

Four basic types of error occur in all experimental titration datasets: absorbance error (AE), transmittance error (TE), composition error (CE), and stock solution error (SE). The first two spring from imperfections in the optics and electronics of spectrometers. The latter two typically spring from uncertainties in reagent purity, volume measurements, and solvent loss over time. These latter sources of error have been referred to as “chemical” rather than “classical” sources of error [25]. As pointed out by the ISO Guide to the Expression of Uncertainty in Measurement (GUM), it is essential to account for the impact of chemical errors, and failure to do so is a common cause of poor uncertainty estimation [26]. A good spectrometer will limit the absorbance error to around 0.0001 units and the transmittance error to around 0.02% T [25]. Good experimental technique will limit errors in solution composition to less than 1% if not less than 0.1%. Except for absorbance error, each type of experimental error has a nonlinear impact on the data, which itself has a strongly nonlinear impact on the binding constants.

To quantify the uncertainty in the binding constants, the method of linearized standard errors has been commonly used [27]. This technique uses the Jacobian of the model to calculate symmetric errors on the binding parameters, as defined below [28].

$$SE = s\sqrt{(J^T J)^{-1}}$$

Here, J is the Jacobian matrix of the vector-valued model function evaluated at the optimized binding parameters. *s* is the mean squared error defined as

$$s = \frac{sum\ of\ squared\ residuals}{residual\ degrees\ of\ freedom}$$

and *SE* has the t distribution with the residual degrees of freedom.

Though computationally fast and implemented in many software packages, such as Matlab's *nlparci* function, linearized standard errors make several assumptions:

1. The predictor variables (i.e., compositions) contain no error.
2. The response variables (i.e., absorbance data) contain only normally distributed error.
3. The response variable errors are independent (i.e., uncorrelated) and identically distributed.
4. The error should be symmetric around the calculated parameter values.
5. The least-squares regression model behaves linearly at the optimized parameter values.

Assumption 1 is clearly violated in spectrophotometric titrations, as composition error and stock solution error can be as large as 1%. The relative error in the predictor variables is generally larger than the relative error in the response variables for titrations, constituting a major problem for the linearized standard error method. Assumption 2 is violated by transmittance error, which does not follow a Gaussian distribution because it impacts the negative logarithm of the absorbance rather than the absorbance itself. Assumption 3 may be violated by effects such as spectrometer baseline drift or

heteroskedasticity at short wavelengths. Such effects are challenging to quantify, however, so we do not include them in our computational studies. Assumptions 4 and 5 do not hold due to the strongly nonlinear nature of the equilibrium constant equation. Monte Carlo studies have shown that simulated experimental errors lead to asymmetric binding constant distributions in spectrophotometric and NMR titrations [29,30], implying that the corresponding confidence intervals must also be asymmetric. Due to these violated assumptions, we will show that linearized error range is inadequate for almost all real and simulated titration datasets.

To avoid these assumptions, a different approach must be used to ascertain the uncertainty: bootstrapping [31]. Bootstrapping is a statistical procedure that attempts to simulate noise in a dataset by iteratively reshuffling the dataset itself [32]. New datasets are constructed by randomly sampling from the original dataset with replacement, such that some values may be included multiple times and other values may not be included at all. This effectively weights the errors in the original dataset differently, thereby mimicking what other datasets collected under the same experimental conditions might have looked like. The new datasets will be the same size as the original and can be modeled in a parallel fashion according to the same set of chemical reactions [33]. The resulting outputs (e.g., binding parameters) will vary with each resampled dataset because the signals and noise in the original dataset have been randomly shuffled and consequently their impact on the outputs is altered. The distribution of the outputs can then be used to form an asymmetric confidence interval.

Unlike linearized standard errors, bootstrapping makes no assumptions about the mathematical form of error present in the data. It can handle error on both predictor and response variables, as well as non-normal distributions such as that of transmittance error. The only requirement is that randomly sampling from the data should return values with different error levels. If the error affects the entire dataset in systematic fashion, bootstrapping may not be able to detect the presence of error. These considerations suggest that bootstrapping will accurately address absorbance, transmittance, and

composition error, but may struggle with stock solution error. We address this point in detail throughout the manuscript.

Bootstrapping compares favorably to other advanced methods for producing binding constant uncertainty ranges. Thordarson and Hibbert have employed a Monte Carlo technique for estimating binding constant error from spectrophotometric titration data [30]. This approach is similar to bootstrapping in that many new datasets are iteratively generated. However, unlike bootstrapping, the Monte Carlo method requires the user to specify the amount of each type of experimental error through an uncertainty budget. As such, the Monte Carlo method is an error propagation technique, whereas bootstrapping is an *a priori* uncertainty estimation technique. While Monte Carlo error propagation is useful for its precise treatment of all uncertainties including stock solution error, the chosen error levels are somewhat subjective [10]. We show in this study that bootstrapping can handle all four types of error well for experimentally relevant conditions. Because bootstrapping does not require *a priori* assumption of uncertainty levels as Monte Carlo does, we believe that bootstrapping should be the method of choice.

It is important to acknowledge that bootstrapping is only effective if a sufficiently large number of iterations are achieved. This makes bootstrapping much more computationally expensive than calculating linearized uncertainty ranges. However, this computational expense is key to the success of bootstrapping, as it enables elimination of faulty assumptions about the mathematical relationships and experimental errors.

Bootstrapping is also used throughout the scientific world [34,35,36,37,38]. Taavitsainen *et al.* have contrasted linearized and bootstrapping approaches to ascertain uncertainty values for kinetic parameters, but they found that the two approaches gave similar results for their model [39]. Furusjö and Danielsson used bootstrapping to estimate the reliability of kinetic parameters determined in the

presence of spectrometer baseline drift [40]. In their 2006 review, Clifford and Maeder observed that confidence intervals derived from linearized regression statistics are often found to be unreliable in an experimental context [41]. They anticipated that bootstrapping should provide superior results but include no citations to existing literature studying this technique. Two existing papers have applied bootstrapping to binding constants calculated from experimental fluorescence and UV vis titration data [42,43]. However, they include no experimental replicates or simulations to evaluate the performance of the bootstrap, and we will also show that their implementation is not reliable for the highly correlated data found in multi-wavelength spectrophotometric titrations.

Our study answers the call to provide rigorous, computationally tractable methods for robust uncertainty estimation in equilibrium spectrophotometric titrations. We focus on the hard modeling of equilibrium data from spectrophotometric titrations. We consider uncertainty in both thermodynamic parameters and molar absorptivity values. The former will be represented as asymmetric 95% confidence intervals on the ΔG° of binding. The latter will be represented as molar absorptivity bands for each chemical absorber. We analyze both simulated and real datasets to ensure maximum relevance to experimental work.

**Methods**

Simulated UV-vis spectrophotometric titration datasets are constructed to emulate typical data of a chromophore host molecule whose absorbance shifts modestly upon complexation with a non-absorbing guest molecule. We assume a host concentration of 1 mM that remains constant throughout the titration and a 1:1 host:guest complex. Gaussian shaped molar absorptivity bands with a peak height of 1100 $M^{-1}cm^{-1}$ and full-width-at-half-max of 100 nm for both host and host:guest complex are staggered 17 nm apart (Figure 3a). The ΔG° value for the complexation reaction, *i.e.,* the true binding

parameter, is set to be 0.0, -20.0, -40.0, or -60.0 kJ/mol, which converts to a range of $\log K_a$ from 0 to 10.5 at 298 K. Because of the wide range of binding parameters from weak to strong, the 50 steps of each titration are designed to telescope out from the most crucial number of guest equivalents: $1 + 1/(KH_0)$, which equals to $1 + 1000/K$ since $[H]_0 = 0.001$ M. The first solution is zero equivalents; the sixth solution is $0.9 + 1000/(3K)$ equivalents; the $45^{th}$ solution is $1.1 + 3000/K$ equivalents; the $51^{st}$ solution is $1 + 5000/K$ or 2 equivalents, whichever is greater. The intervening solutions are spaced in equal volume addition increments over their respective intervals, assuming an initial 2.5 mL of analyte and a 50 µL analyte addition step size. Such a construction helps ensure that every dataset is optimally sensitive to the binding parameter with which it is constructed [29].

Four types of normally-distributed errors are introduced into our simulations: absorbance error (AE), where error with a mean of zero is added directly to the raw data; transmittance error (TE), where error with a mean of zero is added to the transmittance before converting back to absorbance for modeling; composition error (CE), where error with a mean of one is proportionally multiplied into the initial concentrations of both host and guest separately for each chemical solution, and stock solution error (SE), where both titrant and analyte stock solution concentrations are multiplied by random errors from a normal distribution with a mean of one, systematically shifting all the solution compositions. The ranges of the standard deviation of the transmittance and absorbance error from 0.00001 to 0.04 were chosen to span a range of instrument qualities, given that most research-grade spectrometers have a transmittance error rating around 0.0002 T [24]. Composition and stock solution error levels were chosen to cover a range from 0.01% to 10%, including both experimentally reasonable values and extreme cases to test the limits of the algorithm.

Each simulated dataset is then modeled to ascertain the value of a binding parameter that leads to the best fit of the data as measured by root-mean-square-error. Each optimization was carried out using the Levenberg-Marquardt algorithm starting at the true binding parameter with which the data were

constructed. Since the optimization starts in the global minimum convergence well, non-negativity does not need to be enforced on the model to avoid local minima [44].

Each dataset was then subjected to one of six different types of bootstrapping. For each bootstrap iteration, a new dataset, identical in size to the original, is generated from the original dataset according to one of the following methods.

1) Data column-wise bootstrapping: each column of the original dataset, corresponding to a chemical solution, is replaced with a random column from the original dataset. The resulting dataset may contain multiple identical columns. (On average, 63% of the original data are present because $(1-1/n)^n$ approaches 1/e = 36.8% as *n* increases and is already above 36% when n = 24).
2) Data row-wise bootstrapping: each row of the original dataset, corresponding to a wavelength, is replaced with a random row from the original dataset. The resulting dataset may contain multiple identical rows. On average, 63% of the original data are present.
3) Data column & row bootstrapping: first each column of the original dataset is replaced with a random column from the original dataset, then each row of this modified dataset is replaced with a random row from itself. On average, 63% × 63% = 40% of the original data are present.
4) Residual column-wise bootstrapping: a random column of the residual matrix, corresponding to a chemical solution, is added to each column of the reconstructed absorbance data (molar absorptivity × equilibrium concentrations) to generate a new dataset. 100% of the original data are present.
5) Residual row-wise bootstrapping: a random row of the residual matrix, corresponding to a wavelength, is added to each row of the reconstructed absorbance data to generate a new dataset. 100% of the original data are present.

6) Residual point-wise bootstrapping: a random entry of the residual matrix is added to each point of the reconstructed absorbance data to generate a new dataset. 100% of the original data are present.

Figure 1 shows how columns and rows are defined in relation to the dimensions of the matrices used to define the model for Beer's Law. Heretofore we will use the terms column-wise and row-wise in accord with this depiction.

It is important to distinguish the properties of data bootstrapping and residual bootstrapping. As shown above, new datasets obtained from data bootstrapping only contain a subset of the original data. If the original dataset is not sufficiently dense, these resampled datasets may produce poor estimates of the binding parameter due to a lack of crucial chemical information. On the other hand, residual bootstrapping retains all the original data, but implicitly assumes that all the error has the same structure, and therefore can fail when systematic errors are present in the data.

Each newly generated dataset is then modeled in a parallel manner as the original one to ascertain the binding parameter leading to the best fit for the new absorbance data, measured as usual by the RMSE figure of merit. The process is repeated *n* times, where we refer to *n* as the number of bootstrap iterations. The resulting binding parameters constitute a distribution that covers a numeric range, the width of which represents the uncertainty of the original binding parameter. We herein report the range from the 2.5% percentile to the 97.5% percentile of the bootstrapped $\Delta G^\circ$ values as an estimate of the 95% confidence intervals of the original optimized $\Delta G^\circ$. This technique is known as the percentile confidence interval [45]. A variety of more sophisticated methods have been developed for computing bootstrap confidence intervals once the bootstrap distribution has been obtained [46,47]. However, we find in practice that these refinements have little impact on the estimated uncertainty and may require excess computation time.

By definition, the true ΔG° of a chemical system should fall within the constructed confidence interval for 95% of equivalent datasets analyzed. This fact allows us to test the performance of a confidence interval using simulations. Monte Carlo studies are performed wherein random error is added to a single pristine dataset to produce 100 new datasets. Each of these Monte Carlo datasets possesses the same true ΔG° value but a unique optimized ΔG° value, owing to the unique randomized error signature added. Next, each Monte Carlo dataset is bootstrapped to generate 100 new bootstrap datasets per Monte Carlo dataset. The 95% confidence interval on the ΔG° value is calculated as described above. If the bootstrapping procedure is producing accurate 95% confident intervals, then the bootstrapped confidence interval should include the true ΔG° value for approximately 95 of the 100 Monte Carlo iterations. We refer to this number as the bootstrap coverage % and use it as the key figure of merit for evaluating confidence interval quality. An analogous procedure may be followed for linearized standard error confidence intervals. Figure 2 displays the process in schematic form.

To help interpret the ensuing tables it is important to evaluate some specific likelihoods. If a binary event with a 95% likelihood of success is probed 100 times, the most likely number of successes is 95. However, there is also an 82% likelihood that the number of successes differs from 95. Evaluation of the binomial cumulative distribution function shows there is a 99.6% chance that there will be between 89 and 100 successes in this scenario. If the same scenario is tested 1000 times, there is then a 99.7% chance that the number of successes ranges from 930 to 970. We have included shading in the tables to qualitatively indicate how unlikely that particular result would be if the 95% confidence intervals shown are reliable.

Besides building confidence intervals on the binding parameter, bootstrapping can also produce 95% confidence intervals on the molar absorptivity values of each chemical species at each wavelength. These are not equivalent to the molar absorptivity curves calculated at the extrema of the binding parameter confidence interval, but rather give a wavelength-by-wavelength analysis of the range of

molar absorptivity values. Others have constructed molar absorptivity uncertainty bands for self-modeling curve resolution algorithms [48,49], and the software package MCR-BANDS provides a flexible method to compute these for a variety of soft modeling scenarios in conjunction with MCR-ALS [50,51]. We find that the hard modeling absorptivity bands are significantly tighter than those of soft modeling, pointing to the extra chemical information and constraints employed in the hard modeling context.

Overall, bootstrapping is easy to implement computationally and starts to provide insightful quantification of uncertainty within seconds. More iterations take more computing power but modeling these single reaction systems without non-negativity takes only about 0.1 sec per iteration on a standard desktop computer. This means that bootstrapping a dataset 10,000 times takes about 20 minutes, and a 1000 iteration Monte Carlo study on 100 iteration bootstraps takes about 4 hours.

**Results**

Table 1 reports Monte Carlo confidence interval simulations for all six bootstrapping techniques under the four types of experimental error. Coverage values close to 95% indicate well-performing confidence intervals. Of the six types of bootstrapping tested, the two column-wise approaches are clearly superior for spectrophotometric titration data. Column-wise bootstrapping handles absorbance error, transmittance error, and composition error well when these types of error are present separately in a dataset. Graphs detailing each table entry can be found in the Supporting Information.

It is not surprising that column-wise bootstrapping is the only type that handles composition error, as composition error impacts each chemical solution of the titration in a unique, correlated manner. To correctly handle this correlation structure, it is necessary to treat each chemical solution as a unit when performing the bootstrapping process. Linearized uncertainty suffices for absorbance error because this type of error satisfies assumptions 1 - 4 of standard errors, as detailed in the introduction. For all other

types of error, linearized uncertainty estimations fall short. Column-then-row bootstrapping appears to overestimate the uncertainty range. This likely results from the fact that only 40% of the original data are present after bootstrapping, thereby essentially bootstrapping the data twice. None of the methods can handle stock solution error present by itself because this is a systematic error that is not significantly impacted by the reshuffling of the data or residuals. These results explain why a previous attempt to apply residuals-pointwise bootstrapping to multi-wavelength titrations found that the calculated confidence intervals were no different than those obtained by linearized standard error [39]. Considering these results, we focus our further studies on column-wise bootstrapping.

Table 2 investigates how the confidence interval coverage % varies with the magnitude of the added error and the binding strength, comparing the column-wise bootstrapping and linearized standard error methods. Graphs detailing each entry can be found in the Supporting Information. Again, linearized uncertainty estimations fall short for all types and levels of error except absorbance error. Column-wise bootstrapping of the data performs well for all but stock solution error, though performance deteriorates slightly in the strongest binding regime tested ($\Delta G^\circ$ = -60 kJ/mol, $K[H_0] = 3 \times 10^7$). Column-wise bootstrapping with residuals also performs well for all but stock solution error, though coverage % deteriorates slightly with high levels of transmittance error and composition error in weaker binding regimes ($\Delta G^\circ \geq$ -20 kJ/mol, $K[H]_0 \leq 3200$). Both column-wise bootstraps appear to handle stock solution error under the strongest binding regimes tested. This is likely because stock solution error produces systematic deviations from equilibrium behavior. In the case of highly exergonic binding, such deviations can produce an apparent extent of reaction greater than one, leading the optimizer to calculate a very large binding constant. Since these bootstraps will widen the confidence intervals dramatically, the confidence intervals will include the true binding parameter largely by accident. Similar phenomena are observed in Tables 3 and 4.

Tables 1 and 2 show that stock solution error is not handled well when it is the only error present. At an initial glance, this could be a serious issue for the bootstrapping technique, as all experimental datasets will inevitably contain at least some stock solution error. In experiments, however, stock solution error does not exist in isolation, but in combination with absorbance error, transmittance error, and composition error. We now show that these other forms of error constitute the limiting factor on the precision of binding constant calculation for experimentally reasonable error values. As a result, the bootstrapping method remains valid for experimental error combinations.

Table 3 demonstrates that when present with sufficient levels of other types of error, even stock solution error (SE) is handled well by column-wise bootstrapping on the data. Four distinct error regimes are delineated. Error Regime 1 demonstrates that if the level of SE outpaces the level of the other three types of error, then the bootstrapping confidence intervals are not reliable because they end up much too small. Error Regime 2 demonstrates that the confidence intervals fall slightly short if the level of composition error (CE) is the same as that of the SE and the absorbance/transmittance error (AE/TE) is too small. Error Regime 3 shows that stock solution error up to 1% can be handled well if the other types of error are similar in magnitude. Either sufficient CE or AE/TE can lead to reliable confidence intervals in the presence of SE. Specifically, 0.1% SE with 0.3% CE and minimal AE or TE (0.0003) leads to 92.7% confidence interval coverage, and 0.1% SE with no CE and 0.01 AE and TE leads to 99.4% confidence interval coverage. Bootstrapping for Error Regime 3 does appear to moderately overestimate the confidence intervals in some cases, as judged by the coverage success exceeding the expected 97%. However, bootstrapping remains the most helpful method for quantifying the uncertainty in this region, as linearized error ranges are too small. Finally, for extreme amounts of SE (Error Regime 4), the percentage of Monte Carlo successes not only exceed 97%, but the confidence intervals soar to uselessly high values. This occurs because the fit of the binding model to the data has deteriorated extensively.

We argue that common experimental conditions fall in either Error Regime 2 or 3, implying that data column-wise bootstrapping will yield 95% confidence intervals no worse than coverage in the high 80% range, and likely close to 95%. Careful solution preparation and handling should allow CE and SE to have comparable magnitudes around 0.1%. As has been noted, research grade spectrometers carry transmittance and absorbance error ratings between 0.0001 and 0.0003, consistent with simulations in both Error Regimes 2 and 3. However, a variety of additional factors may contribute to absorbance errors, including baseline drift and light scattering [25]. Such considerations could make higher-absorbance-error scenarios under Error Regime 3 plausible. Regardless, both scenarios show that the bootstrapping method can give reliable confidence intervals for experimental titrations with stock solution error. Under these idealized error conditions, bootstrapping with the data outperforms bootstrapping with the residuals, and both clearly outperform the linearized uncertainty estimations, which fall short for every combination of error tested at $\Delta G^\circ$ = -20 kJ/mol.

We now turn to experimental data to obtain the most realistic picture of bootstrapping performance. The bootstrapping technique is designed to detect all non-systematic errors in the data, regardless of the mathematical distribution, and experimental data may contain more errors than modeled in our simulations. Therefore, while the bootstrap confidence interval performance should be no worse than in our simulations, it may experimentally be much better. Eleven experimental datasets [52,53] were bootstrapped 1000 times to ascertain 95% confidence intervals on the binding parameters. Both data and residual column-wise bootstraps were used. Table 4 displays the results along with linearized confidence intervals for all eleven data sets. All the datasets could be fit adequately with a single chemical reaction producing a 1:1 host:guest complex except for the triazolophane fluoride system, which also produces a 2:1 triazolophane:fluoride complex [54]. The details of each fitted model can be found in the Supporting Information.

The first three rows of Table 4 constitute three experimental replicates on a single system, enabling us to judge the accuracy of the bootstrapped and linearized confidence intervals. The range of replicate ΔG° values obtained is 0.1 kJ/mol. Bootstrapped confidence intervals have an average range of 0.3 kJ/mol, about the right magnitude for a 95% confidence interval. The ranges overlap for all three replicates, whether bootstrapped on the data or with the residuals. Linearized confidence intervals, on the other hand, are vastly too small at 0.02 kJ/mol. The uncertainty on the ΔG° parameter is not adequately ascertained with linearized uncertainty estimations, as the range is too small to include the ΔG° values from the other replicates. Note that for asymmetric confidence intervals, it is not justifiable to combine the results of the three replications by simply averaging the values or combining the confidence limits without an explicitly defended mathematical model [55]. Instead, with multiple experimental datasets, it is advisable to concatenate them and optimize a single model with associated bootstrapped confidence intervals. If the resulting uncertainty range expands, then the datasets may possess mismatched baselines or stock solution errors. In this case, we recommend combining the lower and upper bounds, respectively, by taking the square root of the sum of squares and dividing by n.[56]

The magnitudes of the confidence ranges for the macrocyclic cage datasets match the expectations for a high-quality spectrometer with AE and TE around 0.0003 and CE and SE around 1%, which should lead to a range around 0.3 kJ/mol (Table 3). Indeed, the upper limits for the error in the macrocyclic cage dataset was estimated to be SE < 2%, CE < 1%, AE < 0.01, and TE < 0.01, which would make the confidence interval range < 1 kJ/mol. All the datasets represented in Table 4 were measured on benchtop spectrophotometers with AE and TE less than 0.001.

A very important observation arises from the relative magnitude of the linearized and bootstrapped confidence intervals. Throughout all datasets analyzed in Table 4, linearized confidence interval ranges are about an order of magnitude smaller than bootstrapped ranges. In our Table 3 simulations, however, the linearized ranges are only about a factor of two smaller than the bootstrapped ranges.

This strongly suggests that experimental datasets contain unknown errors not included in our Monte Carlo models. These real experimental errors are detected by the bootstrap but not the linearization, thereby widening the interval range disparity. This implies that the relative performance of bootstrapping is even better than what is suggested by our simulations. We take this to indicate the value of bootstrapping as an *a priori* uncertainty estimation technique. Investigations of confidence interval precision suggest that at least 1000 bootstrap iterations should be used on experimental datasets for acceptable reproducibility.

The confidence intervals generated from either data or residual column-wise bootstrapping match each other quite closely, with the latter usually being a bit smaller, especially on the exergonic side. This matches the results of our simulations from Table 3. The small discrepancy likely arises because bootstrapping on the data matrix itself inherently leaves ~37% of the data behind, unlike residuals bootstrapping. If key chemical solutions are not included in any bootstrapped dataset, a large variation in the calculated binding constant may result. For the Buckyball dataset listed at the bottom of Table 4, column-wise data bootstrapping results in about 5% of the bootstraps launching the binding parameter far in the exergonic direction, indicating the loss of sensitivity between the binding parameter and the bootstrapped dataset [24, 29]. Because this dataset contains only ten chemical solutions, the likelihood of leaving out essential information regarding the binding isotherm is increased. Therefore, residuals bootstrapping is recommended for small datasets. This issue would also be mitigated if the dataset contained more chemical solutions.

Finally, we consider uncertainty quantification for the calculated molar absorptivity spectra. A distribution of absorptivity bands can readily be built from the same bootstrap used to quantify the $\Delta G^\circ$ confidence interval. This entails storing the molar absorptivity curves for both host and complex for each of the bootstrap iterations, and then extracting the 2.5 and 97.5% percentile molar absorptivity values for each wavelength. This gives a very different answer than simply taking the optimized molar

absorptivity curves at the two edges of the binding constant confidence interval. The endpoint method does not necessarily even yield an interval. Figure 3b shows the molar absorptivity curves that correspond to the two edges of the binding constant interval. Notice that they cross each other around 560 nm. As the absorptivity value calculated at 560 nm has some uncertainty, this exposes the flaw in the endpoint method.

Figure 3c compares confidence interval bands calculated via bootstrapping to the symmetric bands calculated by linear regression standard errors. Though we have shown linear standard errors fail badly for quantifying binding parameter uncertainty, here we see they perform well for molar absorptivity uncertainty. This is because fewer assumptions are violated. The process of calculating the absorptivity values is indeed a linear regression relationship (satisfying assumption 5), and the true confidence intervals are approximately symmetric (satisfying assumption 4), as seen by comparison to the bootstrap. Thus, linearized confidence intervals and bootstrap confidence intervals both perform satisfactorily on the absorptivity values in our simulations. We still recommend bootstrapping for absorptivity values, however, to guard against experimental errors not captured in these simulations. The Supporting Information includes bootstrapped absorptivity bands for each entry in the previous three tables.

**Conclusions**

We have tested six different bootstrapping protocols for evaluating the uncertainty in the binding parameters obtained from modeling spectrophotometric titration data. While all protocols handle datasets with only absorbance error well, column-wise bootstrapping outperforms other types of bootstrapping, as well as linearized uncertainty estimates, whenever transmittance or composition error is present. Column-wise bootstrapping on the data columns seems to hold up best when multiple types

of error are present simultaneously. Stock solution error presents the greatest challenge for the bootstrapping approach. However, simulations matching experimental error conditions suggest that stock solution error is not the limiting uncertainty factor. Bootstrapping can regularly produce confidence intervals with ~85% coverage in the worst case to 95% coverage in the best case for simulated datasets. This highlights the importance for making stock solutions of precise concentration when performing spectrophotometric titrations.

The superiority of bootstrapping over linearization is especially clear with experimental datasets, wherein the confidence interval width is around ten times smaller for the latter. This suggests that experiments contain extra errors which are not captured in our simulations. Bootstrapping detects these errors while linearization does not, implying even better performance of bootstrapped confidence intervals for experimental data. To the extent that unidentified types of error exist in a dataset, bootstrapping is also superior to Monte Carlo error estimations, where the user must anticipate the mathematical form and magnitude of each type of error.

Both data and residual based column-wise bootstrapping are straightforward to implement, yielding comparable results in most cases. Data bootstrapping produces slightly more conservative confidence intervals. Residual bootstrapping seems more protected against launching, which occurs under strong binding conditions or with titrations containing few chemical solutions. Therefore, data column-wise bootstrapping should be used to estimate the uncertainty of the thermodynamic parameters ($\Delta G^\circ$ or logK) when hard-modeling spectrophotometric titration data unless the binding regime is strong ($K[H]_0 > 10^5$) or the number of chemical solutions per complex is small (< 10). It is always safe to conduct both types of column-wise bootstrapping and compare the results.

Our results provide an important advance in hard modeling of spectrophotometric titrations. Despite widespread recognition of the inadequacy of linearized standard error uncertainty quantification, this

inaccurate method has persisted in the literature. As a result, many published binding constants likely have estimated error ranges that are much too small. This study describes the first uncertainty quantification technique to rigorously treat the nonlinearities of the problem without requiring subjective estimates of the experimental uncertainties. The calculated ranges of error display appropriate asymmetry and match the binding constant variance between experimental replicates. The column-wise bootstrapping method is simple to code, allowing integration into a wide variety of software packages for analyzing titration data. Sivvu.org already has such bootstrapping methodology incorporated into the hard modeling of spectrophotometric data [16]. Widespread adoption of this technique will improve the reliability of binding constant information in all areas of chemistry.

**Supporting Information**

A data repository for all datasets, bootstraps, and fitted models, including graphical and textual output for the modeling of the datasets in Tables 1 – 4, as well as Matlab scripts and code used to generate them, can be found at DOI: 10.5281/zenodo.6103374. An instructional readme file to describe how to implement the scripts is also included in the repository.

**Acknowledgements**

The authors wish to thank Randall Pruim, professor of mathematics and statistics at Calvin University, and Tim Hesterberg, Senior Statistician at Google, for helpful discussions related to the theory of bootstrapping.

Funding: This work supported by the National Science Foundation through an RUI grant from the Division of Chemistry (2004005) and an MRI grant (1726260) for the supercluster (borg.calvin.edu),

which supported the computational aspects of this project. We are also grateful to the Arnold and Mabel Beckman Foundation for supporting Calvin University and Nathanael Kazmierczak as a Beckman Scholar.

**Figures**

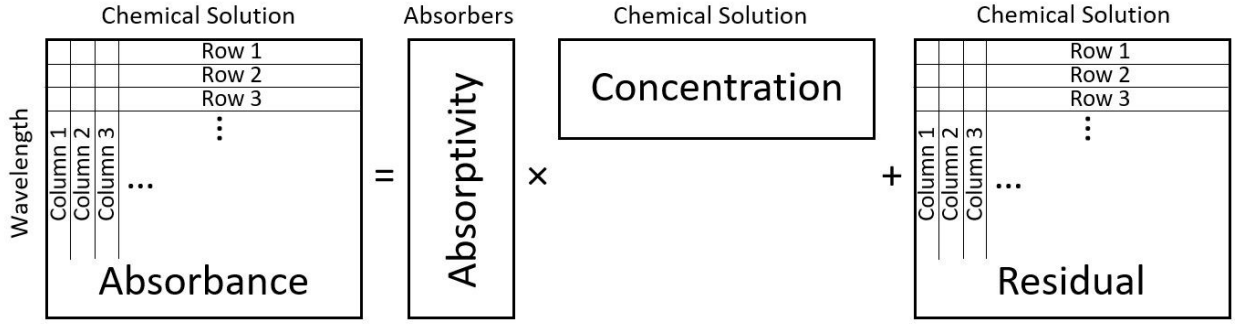



Figure 1. Schematic for the deconvolution of a data matrix of absorbance values into two smaller matrices, Absorptivity and Concentration, in accord with the hard modelling of spectrophotometric titration data according to the Beer-Lambert law for absorbance. The columns and rows are defined as shown throughout.

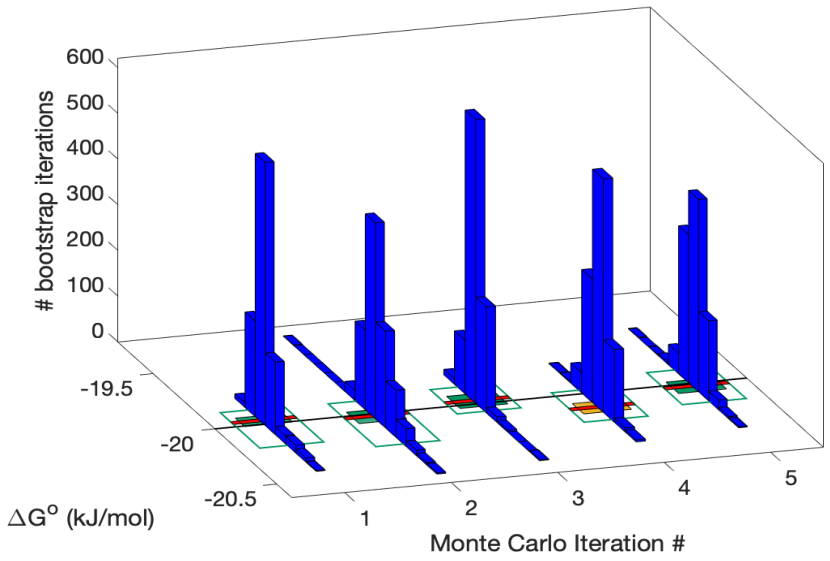



Figure 2. Five histograms each consisting of 1000 bootstrapped datasets all from a single simulated dataset with varied embedded error. The true ΔG° value is -20.0 kJ/mol. The optimized ΔG° value is shown in red for each Monte Carlo iteration. 95% confidence intervals for linearized error estimations are shown in shaded yellow or green rectangles. Open rectangles represent the 95% confidence intervals for the set of bootstraps. Notice that bootstrapped confidence intervals include the true ΔG° value for all five iterations shown here, but the linearized confidence interval for iteration #4 does not.

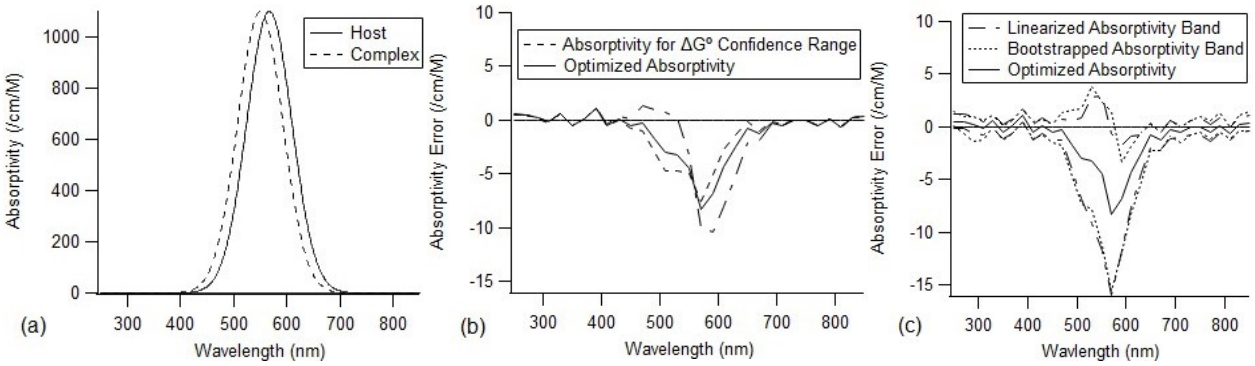



Figure 3. (a): Molar absorptivity curves for the dataset with AE = TE = 0.01 and CE = SE = 1% (the guest is non-absorbing). (b): Deviations in the molar absorptivity curves for 1:1 complex that correspond to the ΔG° 95% confidence interval values, shown with the optimized molar absorptivity error. (c): Bands for the error in the calculated molar absorptivity for the 1:1 complex, both bootstrapped and linearized 2.5% and 97.5% confidence intervals for every 20$^{th}$ wavelength.

**Tables**

Table 1. Confidence interval coverage % values for Monte Carlo simulations with 100 iterations. 100 bootstrap iterations are employed for each Monte Carlo iteration. Datasets are generated with a true $\Delta G^\circ$ = -20 kJ/mol, $[H]_0$ = 0.001, silent guest, 1:1 stoichiometry, and telescoping titration structure. Shading corresponds to inaccuracy of confidence interval estimation.

| Error Type (amount) | Linearized Uncertainty | Data Bootstrapping | | | Residuals Bootstrapping | | |
|---|---|---|---|---|---|---|---|
| | | Column | Row | Column & Row | Column | Row | Point |
| Absorbance Error (0.01) | 94.7* | 97 | 93 | 100 | 95 | 95 | 92 |
| Transmittance Error (0.01) | 80.7* | 100 | 95 | 100 | 97 | 84 | 77 |
| Composition Error (1%) | 52.3* | 96 | 55 | 99 | 89 | 39 | 51 |
| Stock Solution Error (1%) | 0* | 2 | 0 | 8 | 0 | 0 | 0 |

*average over six replications

Table 2. Confidence interval coverage % values for Monte Carlo simulations with 100 iterations. 100 bootstrap iterations are employed for each Monte Carlo iteration. Datasets are generated with $[H]_0$ = 0.001, silent guest, 1:1 stoichiometry, and telescoping titration structure. Shading corresponds to inaccuracy of confidence interval estimation. Each linearized error estimated twice.

| True ΔG°: | | ΔG° = 0 kJ/mol | | | ΔG° = -20 kJ/mol | | | ΔG° = -40 kJ/mol | | | ΔG° = -60 kJ/mol | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Error Level | Linearized Error Estimation | Column Bootstrap on Data | Column Bootstrap on Residuals | Linearized Error Estimation | Column Bootstrap on Data | Column Bootstrap on Residuals | Linearized Error Estimation | Column Bootstrap on Data | Column Bootstrap on Residuals | Linearized Error Estimation | Column Bootstrap on Data | Column Bootstrap on Residuals |
| Absorbance Error | 0.00001 | 89, 98 | 96 | 88 | 94, 93 | 95 | 91 | 93, 95 | 90 | 95 | 97, 94 | 90 | 91 |
| | 0.0001 | 95, 98 | 92 | 95 | 97, 99 | 98 | 94 | 94, 94 | 92 | 93 | 84, 89 | 89 | 97 |
| | 0.001 | 91, 96 | 94 | 91 | 97, 93 | 90 | 93 | 93, 97 | 90 | 94 | 74, 71 | 84 | 95 |
| | 0.01 | 97, 98 | 100 | 96 | 96, 94 | 98 | 93 | 93, 93 | 92 | 93 | 64, 66 | 92 | 99 |
| | 0.02 | 97, 94 | 100 | 97 | 92, 92 | 98 | 95 | 92, 91 | 96 | 95 | 69, 61 | 98 | 95 |
| | 0.04 | 93, 89 | 100 | 99 | 89, 89 | 100 | 98 | 89, 86 | 99 | 100 | 65, 54 | 97 | 99 |
| Transmittance Error | 0.00001 | 77, 79 | 90 | 86 | 82, 83 | 95 | 94 | 78, 90 | 90 | 94 | 86, 79 | 91 | 93 |
| | 0.0001 | 86, 84 | 94 | 93 | 89, 86 | 96 | 96 | 77, 78 | 90 | 96 | 69, 73 | 91 | 95 |
| | 0.001 | 87, 82 | 92 | 95 | 84, 77 | 93 | 87 | 80, 84 | 93 | 95 | 58, 70 | 85 | 94 |
| | 0.01 | 81, 83 | 98 | 91 | 80, 75 | 100 | 91 | 89, 76 | 95 | 93 | 51, 61 | 93 | 99 |
| Composition Error | 0.1% | 48, 60 | 92 | 87 | 67, 60 | 96 | 90 | 36, 34 | 93 | 93 | 41, 30 | 86 | 94 |
| | 0.4% | 59, 60 | 93 | 79 | 54, 59 | 96 | 87 | 32, 30 | 91 | 97 | 22, 26 | 86 | 95 |
| | 0.7% | 57, 51 | 93 | 90 | 53, 51 | 94 | 91 | 25, 35 | 90 | 94 | 15, 18 | 84 | 94 |
| | 1% | 62, 50 | 93 | 92 | 44, 59 | 93 | 96 | 39, 38 | 93 | 94 | 15, 12 | 76 | 93 |
| Stock Solution Error | 0.1% | 0, 0 | 0 | 0 | 0, 0 | 3 | 2 | 0, 0 | 57 | 97 | 23, 16 | 100 | 100 |
| | 0.4% | 0, 0 | 0 | 0 | 0, 0 | 3 | 0 | 0, 0 | 52 | 85 | 17, 16 | 91 | 93 |
| | 0.7% | 0, 0 | 0 | 0 | 0, 0 | 3 | 1 | 0, 0 | 56 | 61 | 7, 10 | 81 | 80 |
| | 1% | 0, 0 | 0 | 0 | 0, 0 | 8 | 0 | 0, 0 | 43 | 69 | 7, 8 | 70 | 63 |

Table 3. Confidence interval coverage % values for Monte Carlo simulations with 1000 iterations. 100 bootstrap iterations are employed for each Monte Carlo iteration. Datasets are constructed with a true $\Delta G^\circ$ = -20 kJ/mol, $[H]_0$ = 0.001, silent guest, 1:1 stoichiometry, and telescoping titration structure. Each row represents a combination of multiple types of error: stock solution error (SE), composition error (CE), absorbance error (AE), and transmittance error (TE). Shading corresponds to inaccuracy of estimation for 1000 iterations. Each linearized error estimated twice.

| | Error Levels | | | | Linearized Error Estimation (%) | Column Bootstrap on Data (%) | Column Bootstrap on Residuals (%) | Bootstrap Average Confidence Interval (kJ/mol) | Linearized Average Confidence Interval (kJ/mol) |
|---|---|---|---|---|---|---|---|---|---|
| Error Regime | SE (%) | CE (%) | AE (ABS) | TE (T) | | | | | |
| 1 | 1 | 0.1 | 0.0003 | 0.0003 | 8.1, 7.6 | 19.3 | 12.1 | 0.03 | 0.01 |
| | 1 | 0 | 0.003 | 0.003 | 46.4, 46.8 | 72.7 | 61.7 | 0.2 | 0.1 |
| 2 | 0.1 | 0.1 | 0.0003 | 0.0003 | 55.7, 52.6 | 85.6 | 75.4 | 0.04 | 0.02 |
| | 0.1 | 0.1 | 0.001 | 0.001 | 80.5, 79.4 | 92.1 | 89.8 | 0.08 | 0.05 |
| | 0.3 | 0.3 | 0.001 | 0.001 | 53.6, 57.8 | 86.4 | 76.9 | 0.1 | 0.06 |
| | 1 | 1 | 0.0003 | 0.0003 | 31.9, 36.8 | 83.9 | 74.9 | 0.3 | 0.1 |
| | 1 | 1 | 0.001 | 0.001 | 36.4, 38.6 | 85.8 | 75.0 | 0.4 | 0.1 |
| | 1 | 1 | 0.003 | 0.003 | 52.8, 50.4 | 89.1 | 76.1 | 0.4 | 0.2 |
| 3 | 0.01 | 0.03 | 0.0003 | 0.0003 | 81.9, 79.8 | 94.3 | 90.4 | 0.02 | 0.02 |
| | 0.03 | 0.03 | 0.0003 | 0.0003 | 77.1, 78.3 | 92.9 | 89.6 | 0.03 | 0.02 |
| | 0.03 | 0.1 | 0.001 | 0.001 | 83.8, 84.4 | 94.0 | 91.0 | 0.1 | 0.05 |
| | 0.03 | 0 | 0.01 | 0.01 | 83.3, 84.2 | 99.3 | 94.5 | 1 | 0.5 |
| | 0.1 | 0.3 | 0.0003 | 0.0003 | 52.7, 54.0 | 93.2 | 85.5 | 0.1 | 0.04 |
| | 0.1 | 0.3 | 0.003 | 0.003 | 83.9, 82.9 | 95.0 | 92.7 | 0.3 | 0.2 |
| | 0.1 | 0 | 0.01 | 0.01 | 84.7, 80.0 | 99.0 | 92.8 | 0.7 | 0.5 |
| | 0.3 | 0 | 0.01 | 0.01 | 83.2, 81.0 | 98.6 | 93.1 | 1 | 0.5 |
| | 0.3 | 0.3 | 0.003 | 0.003 | 77.1, 76.5 | 93.2 | 88.6 | 0.3 | 0.2 |
| | 0.3 | 1 | 0.01 | 0.01 | 77.4, 80.5 | 98.9 | 93.3 | 1 | 0.5 |
| | 1 | 0 | 0.01 | 0.01 | 77.9, 76.4 | 97.5 | 90.3 | 1 | 0.5 |
| | 1 | 1 | 0.01 | 0.01 | 74.7, 74.9 | 97.7 | 91.1 | 1 | 0.5 |
| 4 | 1 | 3 | 0.03 | 0.03 | 30.9, 31.4 | 97.3 | 86.9 | >1000 | >1000 |
| | 3 | 3 | 0.03 | 0.03 | 34.5, 32.3 | 96.2 | 89.6 | >1000 | >1000 |
| | 3 | 10 | 0.1 | 0.1 | 43.4, 45.3 | 99.9 | 70.4 | >1000 | >1000 |
| | 10 | 10 | 0.1 | 0.1 | 43.9, 46.4 | 99.6 | 72.2 | >1000 | >1000 |

Table 4. Uncertainty estimation of the ΔG° parameter for real datasets [52,53,54] using 1000 bootstrap iterations, modeled according to a 1:1 host:guest interaction (Triazolophane fluoride also has a $T_2F^-$ complex).

| Spectrophotometric Titration Dataset | | Silent Guest? | # of Solu-tions | RMSE ($x10^4$) | Calculated ΔG° (kJ/mol) | Linearized Uncertainty Estimation | Column Bootstrap on Data | Column Bootstrap on Residuals |
|---|---|---|---|---|---|---|---|---|
| Macrocyclic Cage (30 μM) with $Cl^-$ | | N | 21 | 3.0 | -17.59 | $^{+0.02}_{-0.02}$ | $^{+0.23}_{-0.07}$ | $^{+0.10}_{-0.05}$ |
| | | | 21 | 3.4 | -17.71 | $^{+0.01}_{-0.01}$ | $^{+0.06}_{-0.32}$ | $^{+0.06}_{-0.10}$ |
| | | | 20 | 3.5 | -17.56 | $^{+0.01}_{-0.01}$ | $^{+0.19}_{-0.14}$ | $^{+0.10}_{-0.10}$ |
| Triazolophane (13 μM) with $F^-$ | $T + F^- \rightleftharpoons TF^-$ | Y | 16 | 25 | -32.40 | $^{+0.17}_{-0.17}$ | $^{+0.69}_{-0.75}$ | $^{+0.34}_{-0.50}$ |
| | $TF^- + T \rightleftharpoons T_2F^-$ | | | | -25.28 | $^{+1.13}_{-1.13}$ | $^{+4.00}_{-2.76}$ | $^{+0.80}_{-0.59}$ |
| Triazolophane (12 μM) with $Cl^-$ | | Y | 17 | 37 | -36.45 | $^{+0.15}_{-0.15}$ | $^{+0.93}_{-0.98}$ | $^{+1.00}_{-1.49}$ |
| Triazolophane (6 μM) with $Cl^-$ | | Y | 16 | 13 | -36.66 | $^{+0.08}_{-0.08}$ | $^{+0.98}_{-1.10}$ | $^{+0.98}_{-1.18}$ |
| Triazolophane (3 μM) with $Cl^-$ | | Y | 17 | 4.8 | -37.91 | $^{+0.05}_{-0.05}$ | $^{+0.57}_{-0.66}$ | $^{+0.56}_{-0.63}$ |
| Triazolophane (1.5 μM) with $Cl^-$ | | Y | 17 | 3.6 | -37.99 | $^{+0.05}_{-0.05}$ | $^{+0.59}_{-0.46}$ | $^{+0.43}_{-0.42}$ |
| Triazolophane (1.1 μM) with $Br^-$ | | Y | 17 | 2.4 | -37.52 | $^{+0.05}_{-0.05}$ | $^{+0.34}_{-0.48}$ | $^{+0.26}_{-0.28}$ |
| Triazolophane (12 μM) with $I^-$ | | N | 14 | 107 | -21.86 | $^{+0.20}_{-0.20}$ | $^{+1.79}_{-3.10}$ | $^{+1.60}_{-2.86}$ |
| Buckyball (65 μM) with Macrocycle | | N | 10 | 4.0 | -19.21 | $^{+0.55}_{-0.55}$ | $^{+3.38}_{-1530}$ | $^{+3.02}_{-4.47}$ |

**References**


1. K. A. Connors, Binding Constants: The Measurement of Molecular Complex Stability, 1st ed.; Wiley-Interscience, 1987.

2. L. E. Manck, C. R. Benson, A. I. Share, H. Park, D. A. Vander Griend, A. H. Flood, Self-Assembly Snapshots of a 2 × 2 Copper(I) Grid, Supramol. Chem. 26 (2014) 267–279. https://doi.org/10.1080/10610278.2013.872780.

3. M. Agnes, A. Nitti, D. A. Vander Griend, D. Dondi, D. Merli, D. Pasini, A Chiroptical Molecular Sensor for Ferrocene, Chem. Commun. 52 (2016) 11492–11495. https://doi.org/10.1039/C6CC05937F.

4. M. Caricato, N. J. Leza, K. Roy, D. Dondi, G. Gattuso, L. S. Shimizu, D. A. Vander Griend, D. Pasini, A Chiroptical Probe for Sensing Metal Ions in Water, Eur. J. Org. Chem. 27 (2013) 6078–6083. https://doi.org/10.1002/ejoc.201300884.

5. E. M. Zahran, E. M. Fatila, C.-H. Chen, A. H. Flood, L. G. Bachas, Cyanostar: C–H Hydrogen Bonding Neutral Carrier Scaffold for Anion-Selective Sensors, Anal. Chem. 90 (2018) 1925–1933. https://doi.org/10.1021/acs.analchem.7b04008.

6. M. Meloun, S. Bordovska, L. Galla, The Thermodynamic Dissociation Constants of Four Non-Steroidal Anti-Inflammatory Drugs by the Least-Squares Nonlinear Regression of Multiwavelength Spectrophotometric PH-Titration Data, J. Pharm. Biomed. Anal. 45 (2007) 552–564.

7. L. N. Dawe, T. S. M. Abedin, L. K. Thompson, Ligand Directed Self-Assembly of Polymetallic [n x n] Grids: Rational Routes to Large Functional Molecular Subunits? Dalton Trans. 13 (2008) 1661–1675. https://doi.org/10.1039/b716114j.

8. E. M. Zahran, Y. Hua, S. Lee, A. H. Flood, L. G. Bachas, Ion-Selective Electrodes Based on a Pyridyl-Containing Triazolophane: Altering Halide Selectivity by Combining Dipole-Promoted Cooperativity with Hydrogen Bonding, Anal. Chem. 83 (2011) 3455–3461. https://doi.org/10.1021/ac200052q.

9. A. de Juan, E. Casassas, R. Tauler, Soft Modeling of Analytical Data. In *Encyclopedia of Analytical Chemistry*; John Wiley & Sons, Ltd, 2006. https://doi.org/10.1002/9780470027318.a5208.

10. P. Thordarson, Binding Constants and Their Measurement. In Supramolecular Chemistry; John Wiley & Sons, Ltd, 2012.

11. H. Gampp, M. Maeder, C. J. Meyer, A. D. Zuberbühler, Calculation of Equilibrium Constants from Multiwavelength Spectroscopic Data--I: Mathematical Considerations, Talanta 32 (1985) 95–101. https://doi.org/10.1016/0039-9140(85)80035-7.

12. E. R. Malinowski, Factor Analysis in Chemistry, 3rd edition.; Wiley: New York, 2002.

13. P. Thordarson, Determining Association Constants from Titration Experiments in Supramolecular Chemistry, Chem. Soc. Rev. 40 (2011) 1305–1323. https://doi.org/10.1039/C0CS00062K.

14. D.J. Leggett, W.A.E. McBryde, General computer program for the computation of stability constants from absorbance data, Anal. Chem. 1975, 47, 1065.

15. P. Thordarson, Supramolecular.org http://supramolecular.org/ (accessed March 2, 2022).

16. D. A. Vander Griend, M. Vermeer, M. Greeley, Y. Kim, N. Wang, D. Buist, C. Ulry, SIVVU.org http://sivvu.org/ (accessed March 2, 2022).

17. M. Garrido, F. X. Rius, M. S. Larrechi, Multivariate Curve Resolution–Alternating Least Squares (MCR-ALS) Applied to Spectroscopic Data from Monitoring Chemical Reactions Processes, Anal. Bioanal. Chem. 390 (2008) 2059–2066. https://doi.org/10.1007/s00216-008-1955-6.

18. ReactLab Equilibria | Jplus ReactLab http://jplusconsulting.com/products/reactlab-equilibria/ (accessed March 2, 2022).

19. C. Destefano, P. Princi, C. Rigano, Computer Analysis of Equilibrium Data in Solution: ESAB2M - An Improved Version of the ESAB Program, Annali di Chimica 77 (1987) 643–675.

20. L. Alderighi, P. Gans, A. Ienco, D. Peters, A. Sabatini, A. Vacca, Hyperquad Simulation and Speciation (HySS): A Utility Program for the Investigation of Equilibria Involving Soluble and Partially Soluble Species, Coordination Chemistry Reviews 184 (1999) 311–318.

21. M. Meloun, S. Bordovská, T. Syrový, A. Vrána, Tutorial on a Chemical Model Building by Least-Squares Non-Linear Regression of Multiwavelength Spectrophotometric PH-Titration Data, Analytica Chimica Acta 580 (2006) 107–121. https://doi.org/10.1016/j.aca.2006.07.043.

22. H. Gampp, M. Maeder, C. J. Meyer, A. D. Zuberbühler, Calculation of Equilibrium Constants from Multiwavelength Spectroscopic Data--II: Specfit: Two User-Friendly Programs in Basic and Standard Fortran 77, Talanta 32 (1985) 251–264. https://doi.org/10.1016/0039-9140(85)80077-1.

23. M. Meloun, Z. Ferenčíková, M. Javůrek, Reliability of Dissociation Constants and Resolution Capability of SQUAD(84) and SPECFIT/32 in the Regression of Multiwavelength Spectrophotometric PH-Titration Data, Spectrochim Acta A Mol Biomol. Spectrosc. 86 (2012) 305–314. https://doi.org/10.1016/j.saa.2011.10.041.

24. K. Hirose, Analytical Methods in Supramolecular Chemistry; Schalley, C. A., Ed.; Wiley-VCH: Weinheim, 2007.

25. L. Sooväli, E.-I. Rõõm, A. Kütt, I. Kaljurand, I. Leito, Uncertainty Sources in UV-Vis Spectrophotometric Measurement, Accred Qual Assur 11 (2006) 246–255. https://doi.org/10.1007/s00769-006-0124-x.

26. Joint Committee for Guides in Metrology. JCGM 100: Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement.; JCGM, 2008.

27. N. Draper and H. Smith (1981) Applied Regression Analysis, 2nd ed., Wiley. p. 94.

28. A. Sen, M. Srivastava, Regression Analysis: Theory, Methods, and Applications; Springer: New York, NY, 1990.

29. N. P. Kazmierczak, J. A. Chew, A. R. Michmerhuizen, S. E. Kim, Z. D. Drees, A. Rylaarsdam, T. Thong, L. Laar, D. A. Vander Griend, Sensitivity Limits for Determining 1:1 Binding Constants from Spectrophotometric Titrations via Global Analysis, J. Chemom. 33 (2019) 3119–3131. https://doi.org/10.1002/cem.3119.

30. D. B. Hibbert, P. Thordarson, The Death of the Job Plot, Transparency, Open Science and Online Tools, Uncertainty Estimation Methods and Other Developments in Supramolecular Chemistry Data Analysis, Chem. Commun. 52 (2016) 12792–12805. https://doi.org/10.1039/C6CC03888C.

31. R. Wehrens, H. Putter, L. M. C. Buydens, The Bootstrap: A Tutorial, Chemom. Intell. Lab. Sys. 54 (2000) 35–52. https://doi.org/10.1016/S0169-7439(00)00102-7.

32. B. Efron, R. Tibshirani, An Introduction to the Bootstrap, 1st edition.; Chapman and Hall/CRC: New York, 1993.

33. E. Carlstein, K.-A. Do, P. Hall, T. C. Hesterberg, H. R. Künsch, Matched-Block Bootstrap for Dependent Data, Bernoulli 4 (1998) 305–328.

34. A. C. Davison, D. V. Hinkley, Bootstrap Methods and Their Application, 1st edition.; Cambridge University Press: Cambridge; New York, NY, USA, 1997.

35. A. Kulesa, M. Krzywinski, P. Blainey, N. Altman, Sampling Distributions and the Bootstrap, Nat. Methods 12 (2015) 477–478. https://doi.org/10.1038/nmeth.3414.

36. H. Babamoradi, F, van den Berg, Å. Rinnan, Bootstrap Based Confidence Limits in Principal Component Analysis — A Case Study, Chemom. Intell. Lab. Sys. 120 (2013) 97–105. https://doi.org/10.1016/j.chemolab.2012.10.007.

37. S. J. Dixon, Y. Xu, R. G. Brereton, H. A. Soini, M. V. Novotny, E. Oberzaucher, K. Grammer, D. J. Penn, Pattern Recognition of Gas Chromatography Mass Spectrometry of Human Volatiles in Sweat to Distinguish the Sex of Subjects and Determine Potential Discriminatory Marker Peaks, Chemom. Intell. Lab. Sys. 87 (2007) 161–172. https://doi.org/10.1016/j.chemolab.2006.12.004.

38. R. Ríos-Reina, S. M. Azcarate, J. Camiña, R. M. Callejón, J. M. Amigo, Application of Hierarchical Classification Models and Reliability Estimation by Bootstrapping, for Authentication and Discrimination of Wine Vinegars by UV–Vis Spectroscopy, Chemom. Intell. Lab. Sys. 191 (2019) 42–53. https://doi.org/10.1016/j.chemolab.2019.06.001.

39. V.-M. Taavitsainen, H. Haario, M. Laine, Rapid Estimation of Chemical Kinetics by Implicit Calibration II, J. Chemom. 17 (2003) 140–150. https://doi.org/10.1002/cem.779.

40. E. Furusjö, L.-G. Danielsson, Uncertainty in Rate Constants Estimated from Spectral Data with Baseline Drift, J. Chemom. 14 (2000) 483–499. https://doi.org/10.1002/1099-128X(200009/12)14:5/6<483::AID-CEM620>3.0.CO;2-J.

41. S. Norman, M. Maeder, Model-Based Analysis for Kinetic and Equilibrium Investigations, Crit. Rev. Anal. Chem. 36 (2006) 199–209. https://doi.org/10.1080/10408340600969619.

42. E. Almansa López, J. M. Bosque-Sendra, L. Cuadros Rodríguez, A. M. García Campaña, J. J. Aaron, Applying Non-Parametric Statistical Methods to the Classical Measurements of Inclusion Complex Binding Constants, Anal. Bioanal. Chem. 375 (2003) 414–423. https://doi.org/10.1007/s00216-002-1693-0.

43. É. D'M Costa, E. B. Ferreira, D. A. Rodrigues, M. H. dos Santos, Acid–Base Equilibrium of Guttiferone-A in Ethanol–Water Mixtures: Modeling and Bootstrap-Based Evaluation of Uncertainties, Chemom. Intell. Lab. Sys. 198 (2020) 103938. https://doi.org/10.1016/j.chemolab.2020.103938.

44. N. P. Kazmierczak, D. A. Vander Griend, Properly Handling Negative Values in the Calculation of Binding Constants by Physicochemical Modeling of Spectroscopic Titration Data, J. Chemom. 33 (2019) 3183–3196. https://doi.org/10.1002/cem.3183.

45. R. Pruim, Foundations and Applications of Statistics: An Introduction Using R, 2nd edition.; American Mathematical Society: Providence, R.I, 2011.

46 . T. C. Hesterberg, Bootstrap Tilting Confidence Intervals and Hypothesis Tests. Hypothesis 1999, 6 (8), 4–5.

47. T. C. Hesterberg, Bootstrap. Wiley Interdisciplinary Reviews: Computational Statistics 2011, 3 (6), 497–526.

48. P. J. Gemperline, Computation of the Range of Feasible Solutions in Self-Modeling Curve Resolution Algorithms, Anal. Chem. 71 (1999) 5398–5404. https://doi.org/10.1021/ac990648y.

49. H. Abdollahi, M. Maeder, R. Tauler, Calculation and Meaning of Feasible Band Boundaries in Multivariate Curve Resolution of a Two-Component System, Anal. Chem. 81 (2009) 2115–2122. https://doi.org/10.1021/ac8022197.

50. J. Jaumot, R. Tauler, MCR-BANDS: A User-Friendly MATLAB Program for the Evaluation of Rotation Ambiguities in Multivariate Curve Resolution, Chemom. Intell. Lab. Sys. 103 (2010) 96–107. https://doi.org/10.1016/j.chemolab.2010.05.020.

51. J. Jaumot, A. de Juan, R. Tauler, MCR-ALS GUI 2.0: New Features and Applications, Chemom. Intell. Lab. Sys. 140 (2015) 1–12. https://doi.org/10.1016/j.chemolab.2014.10.003.

52. S. I. Etkind, D. A. Vander Griend, T. M. Swager, An Electroactive Anion Receptor with High Affinity for Arsenate. J. Org. Chem. 85 (2020) 10050-10061. https://doi.org/10.1021/acs.joc.0c01206.

53. M. Caricato, C. Coluccini, D. Dondi, D. A. Vander Griend, D. Pasini, Nesting Complexation of $C_{60}$ with Large, Rigid $D_2$ Symmetrical Macrocycles, Org. Biomol. Chem. (2010) 3272-3280. https://doi.org/10.1039/c004379f.

54. Y. Li, D. A. Vander Griend, A. H. Flood, Modelling Triazolophane-Halide Binding Equilibria Using Sivvu Analysis of UV-Vis Titration Data Recorded under Medium Binding Conditions, Supramol. Chem. 21 (2009) 111–117.

55. R. Barlow, Asymmetric Errors, eConf 2003, C030908, WEMT002.

56. Garcia, D. *Harris Quantitative Chemical Analysis 8th Edition*. New York: W.H. Freeman and Company, 2010, p. 57.