Evaluating crystallographic likelihood functions using numerical quadratures

Petrus H. Zwart[a,b]* and Elliott D. Perryman[a,b,c]

[a]*Center for Advanced Mathematics in Energy Research Applications, Computational Research Division, Lawrence Berkeley National Laboratory, 1 Cyclotron road, Berkeley, CA 94720 USA,* [b]*Molecular Biophysics and Integrative Bioimaging Division, Lawrence Berkeley National Laboratory, 1 Cyclotron road, Berkeley, CA 94720 USA, and* [c]*The University of Tennessee at Knoxville, Knoxville, TN 37916 USA. E-mail: PHZwart@lbl.gov*

Maximum Likelihood,

Abstract

Intensity-based likelihood functions in crystallographic applications have the potential to enhance the quality of structures derived from marginal diffraction data. Their usage however is complicated by the ability to efficiently compute these targets functions. Here a numerical quadrature is developed that allows for the rapid evaluation of intensity-based likelihood functions in crystallographic applications. By utilizing a sequence of change of variable transformations, including a non-linear domain compression operation, an accurate, robust and efficient quadrature is constructed.

1. Introduction

The estimation of model parameters from experimental observations plays a central role in the natural sciences, and the use of likelihood-based methods have shown to yield robust estimates of 'best guess' values and their associated confidence intervals (Rossi, 2018). Maximum-likelihood estimation goes back to sporadic use in the early 1800s by Gauss (Gauss, 1809; Gauss, 1816; Gauss, 1823) and Hagen (1867), and was further developed by Fisher (1915), Wilks (1938) and Neyman and Pearson (Neyman *et al.*, 1948; Pearson, 1970). In the crystallographic community, Beu *et al.* (1962) were the first to explicitly use maximum likelihood estimation, applying it on lattice parameter refinement in powder diffraction. In a late reaction to this work, Wilson (1980) states that "the use of maximum likelihood is unnecessary, and open to some objection", and subsequently recasts the work of Beu *et al.* (1962) into a more familiar least-squares description. The use of maximum likelihood based methods in structural sciences really took off after making significant impacts in the analysis of macromolecules. For these type of samples structure solution and refinement problems were often problematic due to very incomplete and/or low quality starting models, making standard least squares techniques under perform. In the 1980's and 1990's, likelihood based methods became mainstream, culminating in the ability to routinely determine and refine structures that were previously thought of as problematic (Lunin & Urzhumtsev, 1984; Read, 1986; Bricogne & Gilmore, 1990; de La Fortelle & Bricogne, 1997; Pannu & Read, 1996; Murshudov *et al.*, 1997). A key ingredient to the success was the development of cross-validation techniques to reduce bias in the estimation of hyper parameters that govern behavior of the likelihood functions (Lunin & Skovoroda, 1995; Pannu & Read, 1996). In the beginning of the 21st century, Read and coworkers extended the likelihood formalism to molecular replacement settings as well, resulting in a significant improvement in the ability to solve structures from

marginal starting models (McCoy *et al.*, 2005; Storoni *et al.*, 2004; Read, 2001). The first use of approximate likelihood methods for the detection of heavy atoms from anomalous or derivative data originates from Terwilliger & Eisenberg (1983) who used an origin removed Patterson correlation function for substructure solution, an approach that is equivalent to a second-order approximation of a Rice-based likelihood function (Bricogne, 1997). A more recent development is the use of a more elaborate likelihood formalism in the location of substructures (Bunkóczi *et al.*, 2015), showing a dramatic improvement in the ability to locate heavy atoms.

As the above examples illustrate, impressive progress has been made by the application of likelihood based methods to a wide variety of crystallographic problems. In all scenarios described, key advances were made by deriving problem-specific likelihood functions and applying them on challenging structure determination problems. In the majority of those cases though, a thorough treatment of experimental errors have taken only a secondary role, resulting in approximations that work well in cases of medium or low noise settings. The principal challenge in the handling of random noise in crystallographic likelihood functions, is how to efficiently convolute Rice-like distribution functions modeling the distribution of a structure factor from an incomplete model with errors with a normal-like distribution that models experimental noise. In this manuscript we develop quadrature approaches to overcome said difficulties. The approach derived has direct applications in structure refinement and molecular replacement, while the general methodology can be extended to other crystallographic scenarios as well. In the remainder of this paper we will provide a general introduction into likelihood based methods, list relevant background into numerical integration techniques, develop a adaptive quadrature approach, apply it to a Rice-type likelihood functions and validate its results.

1.1. Maximum Likelihood Formalism

The estimation of model parameters $\boldsymbol{\theta}$ given some data set $\mathcal{X} = \{x_1, \cdots, x_j, \cdots, x_N\}$ via the likelihood formalism is done in the following manner. Given the probability density function (PDF) $f(x_j|\boldsymbol{\theta})$ of a single observation x_j given a model parameter $\boldsymbol{\theta}$, the joint probability of the entire data set is, under the assumption of independence of the observations, equal to the product of the individual PDFs:

$$f(\mathcal{X}|\boldsymbol{\theta}) = \prod_{j=1}^{N} f(x_j|\boldsymbol{\theta}) \tag{1}$$

The probability of the data $\mathcal{X}$ given the model parameters $\boldsymbol{\theta}$ is known as the likelihood of the model parameters given the data:

$$L(\boldsymbol{\theta}|\mathcal{X}) = f(\mathcal{X}|\boldsymbol{\theta}) \tag{2}$$

A natural choice for the *best estimate* of the model parameters is done by finding that $\boldsymbol{\theta}$ that maximizes the likelihood function. This choice is called the maximum likelihood estimate (MLE). The likelihood function itself $L(\boldsymbol{\theta}|\mathcal{X})$ can be seen as a probability distribution, allowing one to obtain confidence limit estimates on the MLE (Rossi, 2018). The determination of the MLE is typically performed by optimizing the log-likelihood, as this is numerically more stable:

$$\log L(\boldsymbol{\theta}|\mathcal{X}) = \sum_{j=1}^{N} \log f(x_j|\boldsymbol{\theta}) \tag{3}$$

Often, the distribution needed for the likelihood function has to be obtained via a process known as marginalization, in which a so-called nuisance parameter is integrated out:

$$f(x|\boldsymbol{\theta}) = \int_{-\infty}^{\infty} f(x, y|\boldsymbol{\theta}) dy \tag{4}$$

where

$$f(x, y|\boldsymbol{\theta}) = f(x|\boldsymbol{\theta}) f(y|x) \tag{5}$$

Depending on the mathematical form of the distributions involved, this marginalization can range from anywhere between a trivial analytical exercise, to a numerically challenging problem. In likelihood functions in a crystallographic setting, this marginalization is required to take into account the effects of experimental noise, and its efficient calculation the focus of this communication.

1.2. Motivation

The most common likelihood function used in crystallographic applications specifies the probability of the *true* structure factor amplitude given the value of a calculated structure factor originating from a model with known imperfections (Sim, 1959; Luzzati, 1952; Woolfson, 1956; Lunin & Urzhumtsev, 1984):

$$f_a(F|F_C,\alpha,\beta) = \frac{2F}{\epsilon\beta}\exp\left[-\frac{F^2+\alpha^2F_C^2}{\epsilon\beta}\right] I_0\left(\frac{2\alpha F F_C}{\epsilon\beta}\right) \quad (6)$$

$$f_c(F|F_C,\alpha,\beta) = \left(\frac{2}{\epsilon\pi\beta}\right)^{1/2}\exp\left[-\frac{F^2+\alpha^2F_C^2}{2\epsilon\beta}\right] \times \cosh\left(\frac{2\alpha F F_C}{2\epsilon\beta}\right) \quad (7)$$

f_a and f_c are the distributions for acentric and centric reflections, ϵ is a symmetry enhancement factor, F the true structure factor amplitude, F_C is the current model structure factor amplitude, while α and β are likelihood distribution parameters (Lunin & Urzhumtsev, 1984). For the refinement of structures given experimental data, the likelihood of the model-based structure factor amplitudes given the experimental data is needed, and can be obtained from a marginalization over the unknown, error-free structure factor amplitude:

$$L(F_C|I_o) = f(I_o|F_C,\alpha,\beta,\sigma_I^2) = \int_0^\infty f(I_o|\sigma_I^2,F)f(F|F_C,\alpha,\beta)dF \quad (8)$$

where $f(F|F_C,\alpha\beta)$ is given by expressions (6) or (7) and $f(I_o|\sigma_I^2,F)$ is equal to a normal distribution with mean F^2 and variance σ_I^2. This integral is equivalent to

the MLI target function derived by Pannu & Read (1996). Because there is no fast converging series approximation or simple closed form analytical expression for this integral, various approaches have been developed, excellently summarized by Read & McCoy (2016), including a method-of-moments type approach to find reasonable analytical approximations to the intensity-based likelihood function.

In this work, we investigate the use of numerical integration methods to obtain high-quality approximations of integral (8), in the hope that these target functions can provide an additional performance boost when working with very marginal data, such as obtained from time-resolved serial crystallography or Free Electron Laser data in which experimental errors are typically worse than obtained using standard rotation-based methods (Brewster *et al.*, 2019). Furthermore, high-quality datasets are rarely resolution limited by the diffraction geometry alone, indicating that a lot more marginal data is readily available that can potentially incredase the quality of the final models if appropriate target functions are available. In the remainder of this manuscript, we develop and compare a number of numerical integration schemes aimed at rapidly evaluating an intensity based likelihood functions and its derivatives that take into account the presence of experimental errors.

2. Methods

In order to evaluate a variety numerical integration schemes and approximation methods, the integration is first recast into a normalized structure factor amplitudes E and normalized intensities Z framework, and the use of the σ_A formulation of the distributions involved, assuming a P1 space group, such that $\epsilon = 1$ (Read, 1997). The joint probability distribution of the error-free structure factor amplitude E and experimental intensity Z_o, given the calculated normalized structure factor E_C, the model

quality parameter σ_A and the standard deviation of the observation σ_Z reads:

$$\begin{aligned} f_a(E, Z_o|E_C, \sigma_A, \sigma_Z^2) &= \frac{2E}{1-\sigma_A^2} \exp\left[-\frac{E^2+\sigma_A^2 E_C^2}{1-\sigma_A^2}\right] \\ &\times I_0\left(\frac{2\sigma_A E E_C}{1-\sigma_A^2}\right) \\ &\times \frac{1}{\sqrt{2\pi\sigma_Z^2}} \exp\left[-\frac{(E^2-Z_o)^2}{2\sigma_Z^2}\right] \end{aligned} \tag{9}$$

for acentric reflections, and

$$\begin{aligned} f_c(E, Z_o|E_C, \sigma_A, \sigma_Z^2) &= \left(\frac{2}{\pi(1-\sigma_A^2)}\right)^{1/2} \exp\left[-\frac{E^2+\sigma_A^2 E_C^2}{2-2\sigma_A^2}\right] \\ &\times \cosh\left(\frac{\alpha E E_C}{1-\sigma_A^2}\right) \\ &\times \frac{1}{\sqrt{2\pi\sigma_Z^2}} \exp\left[-\frac{(E^2-Z_o)^2}{2\sigma_Z^2}\right] \end{aligned} \tag{10}$$

for centrics. The above joint probability distributions need to be marginalized over E in $\mathbb{R}^+$ to obtain the distribution of interest:

$$f_\cdot(Z_o|E_C, \sigma_A, \sigma_Z^2) = \int_0^\infty f_\cdot(E, Z_o|E_C, \sigma_A, \sigma_Z^2) dE \tag{11}$$

2.1. Variance inflation

A common approach to avoid performing the integration specified above, is to inflate the variance of the Rice function $(1-\sigma_A^2)$ by the variance of the "observed structure factor amplitude" (Green, 1979). This approach circumvents the need to perform an integration, but is suboptimal in a number of different ways. First of all, one doesn't observe amplitudes and we are thus required to estimate it and its variance from observed intensity data. A common way to perform the intensity to amplitude conversion is via a Bayesian estimate (French & Wilson, 1978) under the assumption of a uniform distribution of atoms throughout the unit cell. Although this so-called Wilson prior is used in most cases, it requires a number of additional numerical steps and assumptions as compared to the approach take here. Here we assume a constant,

improper prior on the possible values of the structure factor amplitudes on the positive half-line (Sivia & David, 1994), which results in an intensity to amplitude conversion that is unaffected by effects of pseudo-symmetry, diffraction anisotropy or twinning, all while fitting within a single tweet:

$$E_{\rm o} = \left(\frac{1}{2}\left(Z_{\rm o} + \sqrt{Z_{\rm o}^2 + 2\sigma_Z^2}\right)\right)^{1/2} \tag{12}$$

$$\sigma_E^2 = \left[\frac{Z_{\rm o} + 3\sqrt{Z_{\rm o}^2 + 2\sigma_Z^2}}{\sigma_Z^2} + \frac{2}{\left(Z_{\rm o} + \sqrt{Z_{\rm o}^2 + 2\sigma_Z^2}\right)}\right]^{-1} \tag{13}$$

Further details are listed in E. While this procedure allows for a straightforward intensity to amplitude conversion, even when intensities are negative, and can be subsequently used to inflate the variance of the Rice function, it is no substitute for the full integration. Given the simplicity of variance inflation approach and its wide usage in a number of crystallographic applications, it will be used here as a benchmark.

2.2. Approaches to numerical integration

Several conventional numerical integration approximations exist for improper integrals such as expression (8). Standard methods include trapezoidal based methods with a truncated integration range, the use of Laplace's Method or Monte Carlo based methods or approaches based on orthogonal polynomials (Davies & Rabinowitz, 1984). Whereas a straightforward use of a trapezoidal integration scheme is tempting, the shape of the integrand for certain combinations of distribution parameters will result in a fair chance of missing the bulk of the mass of the function, unless a very fine sampling grid is used. When using the Laplace approximation, where the integrand is approximated by an appropriately scaled and translated Gaussian function, the integrand can deviate significantly from a Gaussian, also resulting in a poor performance. These challenges are summarized in Fig.1 where a number of typical integrand shapes are visualized for different parameter choices. A number of numerical integration and

approximation methods will be outlined below, including a discussion on how *ground truth* is established as a basis for comparison.

2.3. Change of variables.

Analytical and numerical integration is often greatly simplified by a change of variables of the integrand (Davies & Rabinowitz, 1984). The change of variable theorem relates the integral of some function $g(u)$ under a change of variables $u = \psi(x)$:

$$\int_a^b g(u)du = \int_{\psi^{-1}(a)}^{\psi^{-1}(b)} g(\psi(x))\frac{d\psi(x)}{dx}dx \tag{14}$$

Considering for example the function $\exp(-Z)$ and the change of variables

$$Z = T^\gamma \tag{15}$$

One obtains

$$\int_a^b \exp(-Z)dZ = \int_{a^{1/(\gamma)}}^{b^{1/(\gamma)}} \gamma T^{\gamma-1}\exp[-T^\gamma] \tag{16}$$

Although the functional form of the integrand in this example suits itself for an analytical evaluation, a visual inspection indicates that different choices of γ drastically change the shape of the function to the point that a local normal approximation has a reasonable fit, Fig 2.

2.4. The Laplace approximation.

The modified shape of the integrand by the change of variable theorem make the use of the so-called Laplace approximation appealing. In a Laplace approximation, the integrand is approximated by a scaled squared exponential function with suitably chosen mean and length scale (Peng, 2018). The Laplace approximation can be derived

from truncated Taylor expansion of the logarithm of the integrand:

$$\begin{aligned}\int_a^b f(x)dx &\approx \int_{-\infty}^{\infty} e^{g(x)}dx \\ &= \int_{-\infty}^{\infty} \exp\left[\sum_{n=0}^{\infty} \frac{g^{(n)}(x_0)}{n!}(x-x_0)^n\right]dx \\ &\approx \int_{-\infty}^{\infty} e^{g(x_0)} e^{-\frac{1}{2}|g''(x_0)|(x-x_0)^2}dx \end{aligned} \tag{17}$$

where $g(x) = \log(f(x))$ and x_0 is the maximum of f such that $g'(x_0) = 0$. This yields

$$\int_a^b f(x)dx \approx f(x_0)\sqrt{\frac{2\pi}{-g''(x_0)}} \tag{18}$$

The effectiveness of this approximation hinges on the location of x_0 (it should be contained within the original integration domain), the magnitude of $g''(x_0)$ and how rapidly higher order derivatives of $g(x)$ vanish around x_0. The change of variable strategy outlined above can aid in a rapid and high-quality approximation of expression (8).

2.5. Quadrature methods

Even though the change of variables approach combined with the Laplace approximation has the potential to yield accurate integral approximations, obtaining reasonable estimates of the derivative of the log-likelihood, as needed for difference maps or first or higher-order optimization methods, is less straightforward. The difficulty arises in the need to obtain the derivative of the location of the maximum of integrand. For this reason, the use of a quadrature approach is of interest, which provides a way to increase the precision of integral and the associated derivatives by increasing the number of sampling points. Quadrature approaches have however been assumed to need a large number of terms to get sufficient precision (Read & McCoy, 2016), making them an unattractive target for practical crystallographic applications. In order to combine the benefits from a Laplace approximation and a numerical quadrature

and overcome drawbacks associated with both approaches, an adaptive integration scheme is developed. Here we construct a quadrature on the basis of a power transform followed by a logistic transform of the integrand, that maps the domain from $[0, \infty)$ onto $[0, 1]$, non-linearly compressing low-mass integrand regions on relatively small line segments, while approximately linearly transforming high-mass areas of the integrand to the middle of the new integration domain, Fig. 2. Once a quadrature has been established, the logarithm of the integral can be recast as the log of the sum of weighted Rice functions as outlined in Appendix D:

$$\begin{aligned} Q(E_C|Z_o, \sigma_A, \sigma_Z^2) &= \log L_\cdot(E_C|Z_o, \sigma_A, \sigma_Z^2) \\ &= \log \left[\sum_{j=1}^{N} w_j f_\cdot(E_j|E_C, \sigma_A) \right] \end{aligned} \tag{19}$$

where E_j are the quadrature sampling points and w_j the associated weights and are dependent on E_c, Z_o, σ_A and σ_Z. The quadrature sampling used can either be an N-point power-transformed hyperbolic quadrature, or a single-point quadrature on the basis of a (power transformed) Laplace approximation. Further details are given in Appendix A–D.

2.6. Derivatives

The practical use of a likelihood based target function requires the calculation of its derivatives so that it can be used in gradient-based optimization methods. From expression 19, derivates with respect to $Y \in \{E_C, \sigma_A\}$ can be obtained as follows:

$$\begin{aligned} Q'_Y(E_C|Z_o, \sigma_A, \sigma_Z^2) &= \frac{d}{dY} Q(E_C|Z_o, \sigma_A, \sigma_Z^2) \\ &= \exp\left[-Q(E_C|Z_o, \sigma_A, \sigma_Z^2)\right] \\ &\times \sum_{j=1}^{N} \frac{dw_j f_a(E_j|E_C, \sigma_A)}{dY} \end{aligned} \tag{20}$$

The derivatives of $f_\cdot(E_j|E_C, \sigma_A)$ with respect to Y are readily obtained if one assumes that the derived quadrature derived simultaneously samples the high-mass regions of

$f_\cdot(E, Z_o|E_C, \sigma_A, \sigma_Z)$ and its derivative with respect to Y. Given that the Laplace approximation only samples $f_\cdot(E, Z_o|E_C, \sigma_A, \sigma_Z)$ at its maximum, it is unlikely that decent derivatives $Q'_Y(E_C|Z_o, \sigma_A, \sigma_Z^2)$ can be obtained in cases where the mass of the integrand is spread widely, as is the case when σ_Z is large. The derivatives of the Rice components with respect to E_C are listed in Appendix C.

3. Results and Discussion

The first step to evaluate the proposed integration methods is to establish ground truth of the integral we wish to approximate. To this extend, an equispaced, non power-transformed trapezoidal quadrature was constructed integrating the function from $E = 0$ to $E = 6$ using 50000 sampling points using all combination of distribution parameters as listed in Table 1. A number of different integration schemes were tested, comparing results using the mean relative error in the log-integrand over all parameter settings. Zero integrand values are set to machine precision so that a logarithm can be taken. Because the variance inflation approximation doesn't actually perform a marginalization, but performs a more *ad hoc* correction to low fidelity measurements, its relative error against the log likelihood is not a fair measure of its performance, nor does it provide insights in its strengths and drawbacks. Instead, we will compare gradients with respect to E_C for all approximations, as this is more indicative of the quality of difference maps and possible progress in refinement.

3.1. Comparing integration methods

Results of the calculation of the intensity based likelihood function with experimental error are summarized in Tables 2 and 3, where the mean and standard deviation of the relative error in the log integrant are reported (in percent). The Laplace approximation behaves relative poorly for the centric case, where without the use of a power

transform a large spread of relative errors is observed. The Laplace approximation for the non-power transformed acentric case ($\gamma = 1$) yields both a mean error and standard deviation below 1%. In order to obtain a similar level of accuracy and precision in the centric case, a 5-point hyperbolic quadrature on a power-transformed likelihood function with γ set to 2 is required.

3.2. Comparing log-likelihood gradients

In order to obtain basic insights in the behavior of the derived likelihood function, we directly compare the gradients of log likelihood functions using the different approximations derived above for a selected set of parameter combinations only. Numerical tests indicate that gradients computed using an 1500-point hyperbolic quadrature of the power-transformed function (with *gamma* set to 1 and 2 for acentric and centric distributions respectively) are indistinguishable from finite difference gradients computed with an fifty-thousand point trapezoidal approach. In order to investigate the quality of the various approximations under common refinement scenarios, we construct a dataset consisting of triplets of $E_{\rm true}$, E_C and Z_o for a given value of σ_A and a fixed $E^2_{\rm true}/\sigma_Z$ ratio, using random sampling techniques. E_{true} was drawn from the acentric and centric amplitude distributions, from which a Z_o was generated by adding Gaussian noise to $E^2_{\rm true}$ at the specified level. For the acentric case, E_C values were generated by adding noise in the complex plane on a complex vector of length $E_{\rm true}$, and subsequently computing the resulting vector length. A similar procedure was followed for the centric test set, only considering perturbations along the real line. This approach allows us to construct a synthetic dataset with a given experimental noise level (Z_0/σ_Z = constant) and model quality (σ_A). The quality of the gradients are gauged by the ratio

$$R_{\rm method} = 100 \times \frac{\langle |Q'_{\rm method} - Q'_{\rm true}| \rangle}{\langle |Q'_{\rm true}| \rangle} \tag{21}$$

The averaging in the above expression is carried out over ten-thousand simulated data points for a 5 and 11 point hyperbolic quadrature, the power-transformed Laplace and variance inflation approximations. The results are shown tables 4–5 and Fig 3. They indicate that for strong data, all gradients calculation methods converge to those obtained using the full intensity-based likelihood function with experimental errors, but that for high and intermediate noise, both the Laplace approximation and the variance inflation method behave poorly. A considerable better performance is obtained when using a quadrature approach.

4. Conclusions

Numerical procedures for efficient determination of the intensity based likelihood function and its gradient are developed and compared. Whereas the Laplace approximation behaves reasonably well for the estimation of the likelihood function itself, our results show that the associated gradients of this approach can be significantly improved by using a numerical quadrature. Given that the derivative of the log-likelihood function are the key ingredient in gradient-based methods and are used to compute difference maps for structure completion, the proposed approach could improve the convergence of existing refinement and model building methods. Although it is unclear what the optimal quadrature order should be in a practical case, our results suggest that it is likely below 15 sampling points, and that it can be tailored to the noise level of an observation. Algorithmicly, the most costly operation is the iterative procedure for finding the maximum of the integrant. The proposed Newton-based method typically converges well within 50 function evaluations, even in absence of predetermined good starting point of the line search. The construction of the hyperbolic quadrature doesn't require any optimization, nor does the subsequent calculation of associated gradient and function values. Given the large additional overhead in refinement or

other crystallographic maximum-likelihood applications, the use of this target function will likely only have a minimal impact on the total runtime of the workflow, while providing a fast converging approximation to the full intensity-based likelihood that takes experimental errors into account. Although only a full integration into a crystallographic software package can determine under what situations a practical benefit can be obtained from using these target functions, their computation should no longer be seen as a limiting factor.

The above algorithms are implemented in a set of python3 routines, and are available upon request. Some parts of this work was prepared in partial fulfillment of the requirements of the Berkeley Lab Undergraduate Research (BLUR) Program, managed by Workforce Development & Education at Berkeley Lab. This research was supported, in part, by the Advanced Scientific Computing Research and the Basic Energy Sciences programs, which are supported by the Office of Science of the US Department of Energy (DOE) under Contract DE-AC02-05CH11231. Further support originates from the National Institute Of General Medical Sciences of the National Institutes of Health (NIH) under Award 5R21GM129649-02. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Appendix A
A hyperbolic quadrature

Given a function $g(x)$, with $x \geq 0$, we seek to compute its integral over the positive half line:

$$G = \int_0^\infty g(x)dx \tag{22}$$

Set

$$h(x) = \log g(x) \tag{23}$$

$$h'(x) = \frac{d}{dx}h(x) \tag{24}$$

$$h''(x) = \frac{d^2}{dx^2}h(x) \tag{25}$$

Define the supremum of $g(x)$ by x_0 such that $h'(x_0) = 0$. For the class of functions we are interested in, $g(0)$ is equal to 0, for instance due to the power transform outlined in in the main text, and $\lim_{x\to\infty} g(x)$ is 0 as well. Define the following change of variables on the basis of a shifted and rescaled logistic function:

$$t = \frac{\exp[kx] - 1}{\exp[kx] + \exp[kx_0]} \tag{26}$$

Note that $t(x = 0) = 0$, and $\lim_{x\to\infty} t(x) = 1$. The inverse function is

$$x(t) = x_0 - \frac{1}{k}\log\left[\frac{\exp(x_0k)(1-t)}{1 + t\exp(kx_0)}\right] \tag{27}$$

and has a derivative with respect to t equal to

$$x'(t) = \frac{dx(t)}{dt} = \frac{\exp(-kt)(\exp(kx_0) + \exp(kt))^2}{k(\exp(kx_0) + 1)} \tag{28}$$

An N-point quadrature can now be constructed by a uniform sampling between 0 and 1:

$$t_j = \frac{j}{N+1} \tag{29}$$

for $1 \leq j \leq N$. Given that both $g(0)$ and $\lim_{x\to\infty} g(x)$ are zero, the integral G can now be computed via a trapezoidal integration rule:

$$G = \frac{1}{N+1}\sum_{j=1}^{N} g(x(t_j))x'(t_j) \tag{30}$$

If k is chosen to be

$$k = \sqrt{\frac{-2h''(x_0)}{\pi}} \tag{31}$$

the above quadrature for $N = 1$ yields the Laplace approximation when $x_0|h''(x_0)|^{1/2}$ is large, as $|x_0 - x(1/2)|$ goes to zero. If a hyperbolic quadrature is constructed on a distribution of power-transformed variable, these derived weights can be multiplied by the Jacobian of that transformation, so that the final numerical evaluation can be carried out in the original set of variables.

Appendix B
Finding x_0

As outlined in the main text, the numerical integration via the hyperbolic quadrature is greatly assisted by the change of variables

$$E = x^{\gamma} \tag{32}$$

with Jacobian

$$\frac{dE(x)}{dx} = \gamma x^{\gamma-1} \tag{33}$$

The location of the maximum value of the integrand, x_0, is found using a straightforward application of Newton root finding algorithm. The logarithm and the derivatives of the integrant are for the acentric case equal to

$$\begin{aligned} h_a(E|E_C, \sigma_A, Z_0, \sigma_z) &= \log 2 + \log E - \log(1-\sigma_A^2) \\ &- \frac{(E - \sigma_A E_C)^2}{1-\sigma_A^2} \\ &+ \log\left[eI_0\left(\frac{2\sigma_A E E_C}{1-\sigma_A^2}\right)\right] \\ &- \frac{1}{2}\log 2\pi - \log\sigma_Z \\ &- \frac{\left(E^2 - Z_o\right)^2}{2\sigma_Z^2} \end{aligned} \tag{34}$$

$$
\begin{aligned}
h'_a(E|E_C,\sigma_A,Z_0,\sigma_z) &= \frac{1}{E} - \frac{2(E-\sigma_A E_C)}{1-\sigma_A^2} \\
&+ \frac{2E_C\sigma_A}{1-\sigma_A^2}\left(\frac{I_1\left[\frac{2EE_C\sigma_A}{1-\sigma_A^2}\right]}{I_0\left[\frac{2EE_C\sigma_A}{1-\sigma_A^2}\right]} - 1\right) \\
&- \frac{2E(E^2-Z_o)}{\sigma_Z^2}
\end{aligned} \tag{35}
$$

$$
\begin{aligned}
h''_a(E|E_C,\sigma_A,Z_0,\sigma_z) &= -\frac{1}{E^2} - \frac{2}{1-\sigma_A^2} \\
&+ \frac{2Z_o - 6E^2}{\sigma_Z^2} \\
&+ \left(\frac{2E_C\sigma_A}{1-\sigma_A^2}\right)^2 \\
&- \left(\frac{2E_C\sigma_A}{1-\sigma_A^2}\right)^2 \left(\frac{I_1\left[\frac{2EE_C\sigma_A}{1-\sigma_A^2}\right]}{I_0\left[\frac{2EE_C\sigma_A}{1-\sigma_A^2}\right]}\right)^2 \\
&- \frac{2E_C\sigma_A}{E(1-\sigma_A^2)}\left(\frac{I_1\left[\frac{2EE_C\sigma_A}{1-\sigma_A^2}\right]}{I_0\left[\frac{2EE_C\sigma_A}{1-\sigma_A^2}\right]}\right)
\end{aligned} \tag{36}
$$

and for the centric case

$$
\begin{aligned}
h_c(E|E_C,\sigma_A,Z_0,\sigma_z) &= \frac{1}{2}\log(2) - \frac{1}{2}\log(\pi(1-\sigma_A^2)) \\
&- \frac{E^2 + (\sigma_A E_C)^2}{2(1-\sigma_A^2)} \\
&+ \log\cosh\left[\frac{\sigma_A E E_C}{(1-\sigma_A^2)}\right] \\
&- \frac{1}{2}\log 2\pi - \log\sigma_Z \\
&- \frac{\left(E^2 - Z_o\right)^2}{2\sigma_Z^2}
\end{aligned} \tag{37}
$$

$$\begin{aligned} h_c'(E|E_C,\sigma_A,Z_0,\sigma_z) &= \frac{\sigma_A E_C}{1-\sigma_A^2}\tanh\left[\frac{\sigma_A E E_C}{1-\sigma_A^2}\right] \\ &- \frac{E}{1-\sigma_A^2} \\ &- \frac{2E(E^2-Z_o)}{\sigma_Z^2} \end{aligned} \tag{38}$$

$$\begin{aligned} h_c''(E|E_C,\sigma_A,Z_0,\sigma_z) &= \frac{1}{\sigma_A^2-1} \\ &+ \left(\frac{\sigma_A E_C}{1-\sigma_A^2}\right)^2 \\ &- \left(\frac{\sigma_A E_C}{1-\sigma_A^2}\right)^2\left(\tanh\left[\frac{\sigma_A E_c E}{1-\sigma_A^2}\right]\right)^2 \\ &+ \frac{2Z_o-6E^2}{\sigma_Z^2} \end{aligned} \tag{39}$$

Because we are interested in finding the maximum of the integrand after the power transform, we need derivatives with respect to x:

$$\begin{aligned} \hat{h}_\cdot(x|E_C,\sigma_A,\gamma) &= \log\gamma + (\gamma-1)\log x + h_\cdot(E=x^\gamma|E_C,\sigma_A) \\ \hat{h}_\cdot'(x|E_C,\sigma_A,\gamma) &= \frac{(\gamma-1)}{x} + \gamma x^{\gamma-1} h_\cdot'(E=x^\gamma|E_C,\sigma_A) \\ \hat{h}_\cdot''(x|E_C,\sigma_A,\gamma) &= \frac{(1-\gamma)}{x^2} \\ &+ \gamma^2 x^{2\gamma-2} h_\cdot''(E=x^\gamma|E_C,\sigma_A) \\ &+ \gamma(\gamma-1)x^{\gamma-2} h_\cdot'(E=x^\gamma|E_C,\sigma_A) \end{aligned} \tag{40}$$

Decent starting values for the Newton search can be found by performing a single Newton-based update on a set of (say) 15 equispaced values of x, sampled between 0 and $x = 6^{1/\gamma}$. The integrand-weighted mean of the resulting updated sampling points is typically refines within 10 iterations to the supremums.

Appendix C
The Rice function and its derivatives

Given the individual Rice function

$$\begin{aligned} f_a(E|E_C,\sigma_A) &= \frac{2E}{1-\sigma_A^2}\exp\left[-\frac{E^2+\sigma_A^2F_C^2}{1-\sigma_A^2}\right] \\ &\times I_0\left(\frac{2\sigma_A E E_C}{1-\sigma_A^2}\right) \end{aligned} \tag{41}$$

$$\begin{aligned} f_c(E|E_C,\sigma_A) &= \left(\frac{2}{\pi(1-\sigma_A^2)}\right)^{1/2}\exp\left[-\frac{E^2+\sigma_A^2E_C^2}{2(1-\sigma_A^2)}\right] \\ &\times \cosh\left(\frac{\sigma_A E E_C}{1-\sigma_A^2}\right) \end{aligned} \tag{42}$$

the derivatives with respect to E_C are equal to

$$\begin{aligned} f'_{a,E_C}(E|E_C,\sigma_A) &= \frac{4E\sigma_A}{(1-\sigma_A^2)^2}\exp\left[-\frac{E^2+\sigma_A^2E_C^2}{1-\sigma_A^2}\right] \\ &\times \left[EI_1\left(\frac{2\sigma_A E E_C}{1-\sigma_A^2}\right)\right. \\ &- \left.\sigma_A E_C I_0\left(\frac{2\sigma_A E E_C}{1-\sigma_A^2}\right)\right] \end{aligned} \tag{43}$$

$$\begin{aligned} f'_{c,E_C}(E|E_C,\sigma_A) &= \left(\frac{2}{\pi(1-\sigma_A^2)^3}\right)^{1/2}\sigma_A\exp\left[-\frac{E^2+\sigma_A^2E_C^2}{2(1-\sigma_A^2)}\right] \\ &\times \left[E\sinh\left(\frac{\sigma_A E E_C}{1-\sigma_A^2}\right)\right. \\ &- \left.\sigma_A E_C\cosh\left(\frac{\sigma_A E E_C}{1-\sigma_A^2}\right)\right] \end{aligned} \tag{44}$$

These functions can be used to compute the derivatives of the intensity-based likelihood function with experimental errors. Derivatives with respect to σ_A can be derived analogously if desired.

Appendix D
Likelihood synthesis

Using the above approaches, the full likelihood function can be expressed as a sum of weighted Rice functions (41,42), where E is sampled on the basis of a quadrature derived from a power transformed variable $E = x^\gamma$ using the hyperbolic sampling

scheme outlined above. Taking into account the combination of the power transform and the hyperbolic quadrature, the sampling nodes of the quadrature are equal to

$$\begin{aligned} E_j &= x_j^{\gamma} && (45) \\ x_j &= x_0 - \frac{1}{k}\log\left[\frac{\exp(x_0 k)(1-t_j)}{1+t_j\exp(kx_0)}\right] && (46) \end{aligned}$$

where t_j, x_0 and k are defined and computed as outlined in Appendices A and B and $1 \leq j \leq N$. The quadrature weights can be now set to absorb the the hyperbolic sampling, the power transform and the observed intensity and its associated standard deviation

$$\begin{aligned} w_j &= \gamma x_j^{\gamma-1} \\ &\times \frac{\exp(-kt_j)(\exp(kx_0)+\exp(kt_j))^2}{k(\exp(kx_0)+1)} \\ &\times \frac{1}{\sqrt{2\pi\sigma_Z^2}}\exp\left[-\frac{\left(E_j^2 - Z_o\right)^2}{2\sigma_Z^2}\right] \\ &\times \frac{1}{N+1} \end{aligned} \quad (47)$$

which yields a sum of weighted Rice functions that approximates the full likelihood function:

$$L.(E_C|Z_o,\sigma_A,\sigma_Z^2) = \sum_{j=1}^{N} w_j f.(E_C|E_j,\sigma_A) \quad (48)$$

where $f.(\cdots)$ is defined in Appendix C.

When the likelihood function is approximated using the power-transformed Laplace approximation instead of using the quadrature approach, we get a weighted Rice function

$$L.(E_C|Z_o,\sigma_A,\sigma_Z^2) = w_0 f.(E_C|E_0,\sigma_A) \quad (49)$$

with the weight given by

$$\begin{aligned} w_0 &= \gamma x_0^{\gamma-1} \\ &\times \frac{1}{\sqrt{2\pi\sigma_Z^2}} \exp\left[-\frac{(E_0^2 - Z_0)^2}{2\sigma_Z^2}\right] \\ &\times \sqrt{\frac{2\pi}{\hat{h}''(x_0)}} \end{aligned} \tag{50}$$

where $E_0 = x_0^{\gamma}$, and $\hat{h}''(x_0)$ is defined in expression (40).

Appendix E
Bayesian estimation of structure factor amplitudes

In order to use an inflated variance modification as an approximation to the full numerical integration, we need to be able to estimate reflection amplitudes and their standard deviations from observed intensities and their standard deviation. While this process is normally performed using a standard French-Wilson estimation procedure, an other route can be adopted following an approach developed by Sivia & David (1994). Assume a uniform, improper prior on E, such that

$$f(E) = \begin{cases} 1 & E \geq 0 \\ 0 & E < 0 \end{cases} \tag{51}$$

resulting in a conditional distribution

$$f(E|Z_o, \sigma_Z^2) \propto E \exp\left(-\frac{(E^2 - Z_o)^2}{2\sigma_Z^2}\right) \tag{52}$$

A normal approximation to this distribution can be obtained by the method of moments or, as done here, by a maximum *a-posteriori* approximation with a mean equal to the mode of the above distribution and a standard deviation estimated on the basis of the second derivative of the log likelihood at the location of the mode:

$$E_o = \sup_E \log\left[f(E|Z_o, \sigma_Z^2)\right] \tag{53}$$

$$\sigma_E^2 = -\left[\frac{d^2}{dE^2} \log\left[f(E|Z_o, \sigma_Z^2)\right]\right]^{-1}_{E=E_o} \tag{54}$$

An analytic expression is readily obtained, resulting in

$$E_o = \left(\frac{1}{2}\left(Z_o + \sqrt{Z_o^2 + 2\sigma_Z^2}\right)\right)^{1/2} \tag{55}$$

$$\sigma_E^2 = \left[\frac{Z_o + 3\sqrt{Z_o^2 + 2\sigma_Z^2}}{\sigma_Z^2} + \frac{2}{\left(Z_o + \sqrt{Z_o^2 + 2\sigma_Z^2}\right)}\right]^{-1} \tag{56}$$

Table 1. *Parameter bounds for comparing integration methods.*

Parameter	Start	End	Sampling points
E_C	0.1	6.0	20
σ_A	0.0	0.95	10
Z_o	-5.0	50.0	20
Z/σ_Z	0.5	10.0	20

Table 2. *Integration results: acentric distribution. Reported are the mean error and standard deviation of the relative log-likelihood over the full parameter range (in percent).*

Method	$\gamma = 1$	$\gamma = 2$	$\gamma = 3$
Laplace approximation	-0.142 / 0.874	0.294 / 0.971	0.485 / 1.135
Quadrature ($N = 3$)	0.191 / 0.778	0.152 / 0.831	0.281 / 1.058
Quadrature ($N = 5$)	0.130 / 0.377	0.126 / 0.481	0.196 / 0.627
Quadrature ($N = 7$)	0.085 / 0.218	0.074 / 0.309	0.116 / 0.428

Table 3. *Integration results: centric distribution. Reported are the mean error and standard deviation of the relative log-likelihood over the full parameter range (in percent). Quadrature results for $\gamma = 1$ are absent because the function is not guaranteed to be zero at the origin as required by the hyperbolic quadrature scheme.*

Method	$\gamma = 1$	$\gamma = 2$	$\gamma = 3$
Laplace approximation	-1.766 / 8.170	0.357 / 1.729	0.738 / 1.841
Quadrature ($N = 3$)	–	0.30 / 1.617	0.725 / 1.850
Quadrature ($N = 5$)	–	0.391 / 0.990	0.438 / 1.183
Quadrature ($N = 7$)	–	0.269 / 0.750	0.311 / 0.943

Table 4. *Comparing likelihood gradients for acentric distribution via R_{method} as outlined in main text. The value of σ_A was set to 0.7 as a representation of an intermediate quality structure. For the hyperbolic quadrature and the Laplace approximation, γ was set to 1.*

Method	$Z/\sigma_Z = 1$	$Z/\sigma_Z = 3$	$Z/\sigma_Z = 7$
Variance inflation	67.3	24.6	4.50
Laplace approximation	17.9	11.4	2.24
Quadrature ($N = 5$)	6.45	2.71	0.19
Quadrature ($N = 11$)	1.83	0.91	0.01

Table 5. *Comparing likelihood gradients for centric distribution via R_{method} as outlined in main text. The value of σ_A was set to 0.7 as a representation of an intermediate quality structure. For the hyperbolic quadrature and the Laplace approximation, γ was set to 2.*

Method	$Z/\sigma_Z = 1$	$Z/\sigma_Z = 3$	$Z/\sigma_Z = 7$
Variance inflation	71.24	23.39	3.24
Laplace approximation	41.18	19.5	2.89
Quadrature ($N = 5$)	15.2	7.18	0.34
Quadrature ($N = 11$)	7.10	4.24	0.05

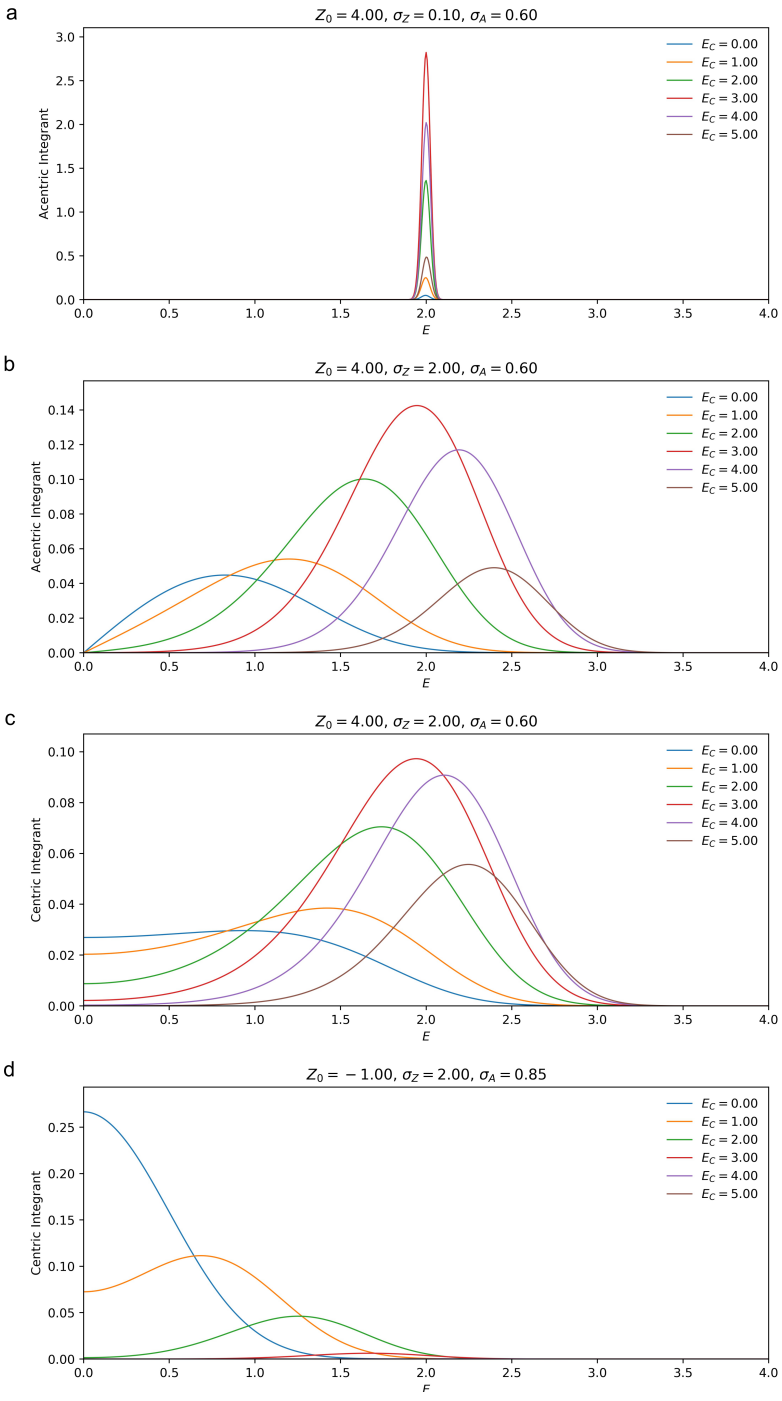


Fig. 1. Integrant shapes for the acentric and centric distribution for different parameter settings show the variety of function shapes that occur when computing marginal likelihood. When the error on the experimental data is relative small, the bulk of the integrants mass if concentrated in a small area (a). When the experimental errors are larger, the resulting integrands display a large variety of shapes (b–d). The variety of these shapes make the uniform application of a standard quadrature or Laplace approximation inefficient and suboptimal.

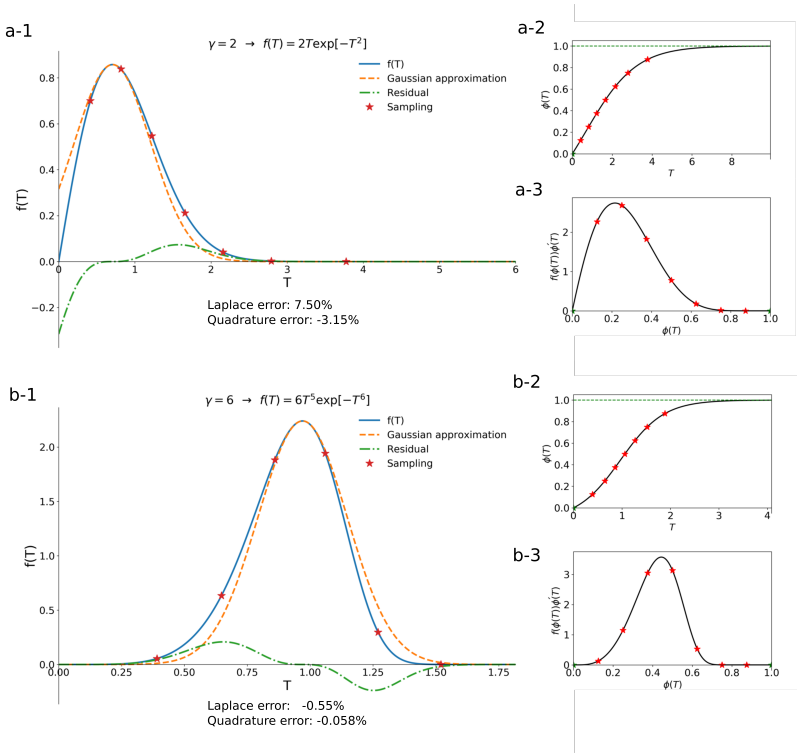


Fig. 2. A change of variables of an integrand drastically changes the shape of the function such that it is more amenable to standard approximation methods such as the Laplace approximation. Using the transformation as outlined in expression 15 with γ equal to 2 (a-1) and 6 (b-1), yields progressively more Gaussian-shaped integrands. The influence of change of variable on the quality of the Laplace approximation is significant as seen from the difference (green) between the true function (blue) and the Gaussian approximation (orange) and the relative error of the integral as obtained using the Laplace approximation. The quadrature approach to evaluate this integral via the outlined nonlinear domain compression (b-2,c-2) method results in a non-uniform sampling of the integrand (red stars) in areas where the integrand contains significant mass.

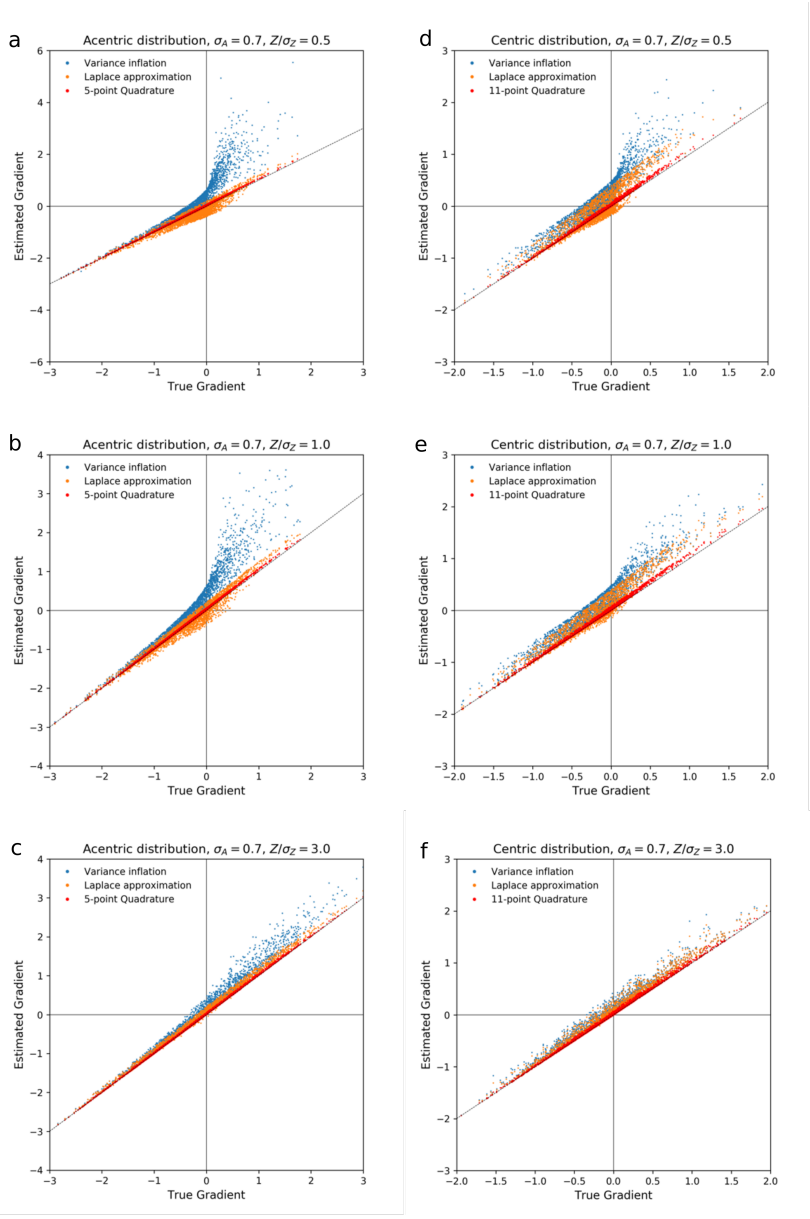


Fig. 3. A comparison of gradients computed using different approximation schemes. Plots a–c and d–f depict the behavior of the gradient approximations with a decreasing noise level for acentric and centric reflections respectively. While modest improvements in the gradients are obtained at $Z_0/\sigma_Z = 3$ using the quadrature approach, a significant performance enhancement is seen at less favorable signal to noise levels.

References

Beu, K. E., Musil, F. J. & Whitney, D. R. (1962). *Acta Crystallographica*, **15**(12), 1292–1301.

Brewster, A. S., Bhowmick, A., Bolotovsky, R., Mendez, D., Zwart, P. H. & Sauter, N. K. (2019). *Acta Crystallogr D Struct Biol*, **75**(Pt 11), 959–968.

Bricogne, G. (1997). *Proceedings of the CCP4 Study Weekend*.

Bricogne, G. & Gilmore, C. J. (1990). *Acta Crystallographica Section A*, **46**(4), 284–297.

Bunkóczi, G., McCoy, A. J., Echols, N., Grosse-Kunstleve, R. W., Adams, P. D., Holton, J. M., Read, R. J. & Terwilliger, T. C. (2015). *Nat Methods*, **12**(2), 127–30.

Davies, P. J. & Rabinowitz, P. (1984). *Methods of Numerical Integration*. Elsevier.

Fisher, R. A. (1915). *Biometrika*, **10**(4), 507–521.

French, S. & Wilson, K. (1978). *Acta Crystallographica Section A*, **34**(4), 517–525.

Gauss, C. F. (1809). *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*, vol. 7. Perthes et Besser.

Gauss, C.-F. (1823). *Theoria combinationis observationum erroribus minimis obnoxiae*, vol. 1. Henricus Dieterich.

Gauss, K. F. (1816). *Zeitschrift für Astronomie und verwandte Wissenschaften, herausgegeben von B. von Lindenau und JG F. Bohnenberger*, **1**(158), 1816.

Green, E. A. (1979). *Acta Crystallographica Section A*, **35**(3), 351–359.

Hagen, G. (1867). *Grundzüge der Wahrscheinlichkeits-Rechnung*. Ernst & Korn.

de La Fortelle, E. & Bricogne, G. (1997). In *Macromolecular Crystallography Part A*, vol. 276 of *Methods in Enzymology*, pp. 472 – 494. Academic Press.

Lunin, V. Y. & Skovoroda, T. P. (1995). *Acta Crystallographica Section A Foundations of Crystallography*, **51**(6), 880–887.

Lunin, V. Y. & Urzhumtsev, A. G. (1984). *Acta Crystallographica Section A*, **40**(3), 269–277.

Luzzati, V. (1952). *Acta Crystallographica*, **5**(6), 802–810.

McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Crystallographica Section D*, **61**(4), 458–464.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Crystallogr D Biol Crystallogr*, **53**(Pt 3), 240–55.

Neyman, J., Scott, E. L. *et al.* (1948). *Econometrica*, **16**(1), 1–32.

Pannu, N. S. & Read, R. J. (1996). *Acta Crystallographica Section A Foundations of Crystallography*, **52**(5), 659–668.

Pearson, E. (1970). *Studies in the History of Statistics and Probability (London, 1970)*, pp. 411–13.

Peng, R. D. (2018). *Advanced Statistical Computing*. https://leanpub.com/advstatcomp.

Read, R. J. (1986). *Acta Crystallographica Section A*, **42**(3), 140–149.

Read, R. J. (1997). In *Macromolecular Crystallography Part B*, vol. 277 of *Methods in Enzymology*, pp. 110 – 128. Academic Press.
URL: *http://www.sciencedirect.com/science/article/pii/S0076687997770095*

Read, R. J. (2001). *Acta Crystallographica Section D*, **57**(10), 1373–1382.

Read, R. J. & McCoy, A. J. (2016). *Acta Crystallographica D Structural Biology*, **72**, 375–387.

Rossi, R. J. (2018). *Mathematical Statistics: An Introduction to Likelihood Based Inference*. Wiley.

Sim, G. A. (1959). *Acta Crystallographica*, **12**(10), 813–815.

Sivia, D. S. & David, W. I. F. (1994). *Acta Crystallographica Section A*, **50**(6), 703–714.

Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Crystallographica Section D*, **60**(3), 432–438.

Terwilliger, T. C. & Eisenberg, D. (1983). *Acta Crystallographica Section A*, **39**(5), 813–817.

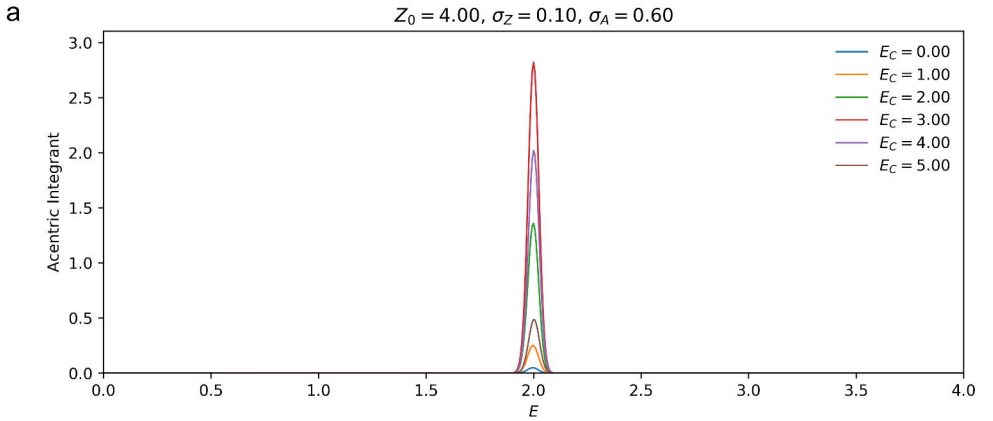
Wilks, S. S. (1938). *Ann. Math. Statist.* **9**(1), 60–62.

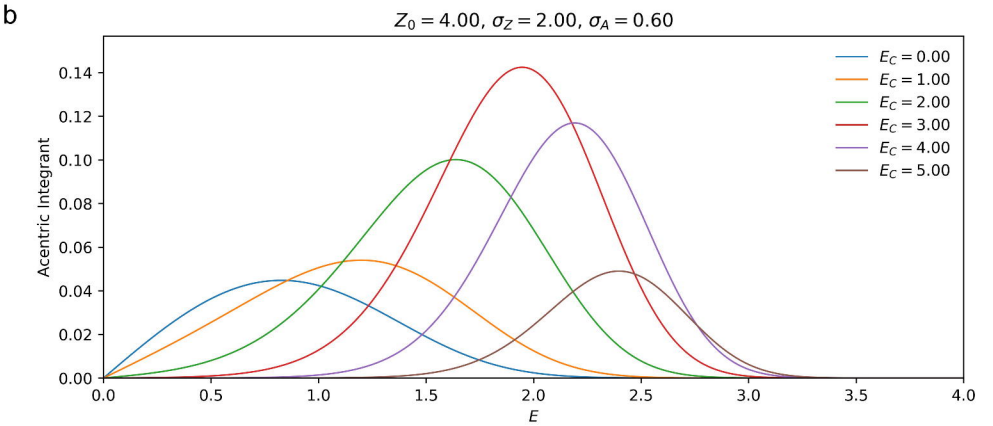
Wilson, A. J. C. (1980). *Acta Crystallographica Section A*, **36**(6), 937–944.

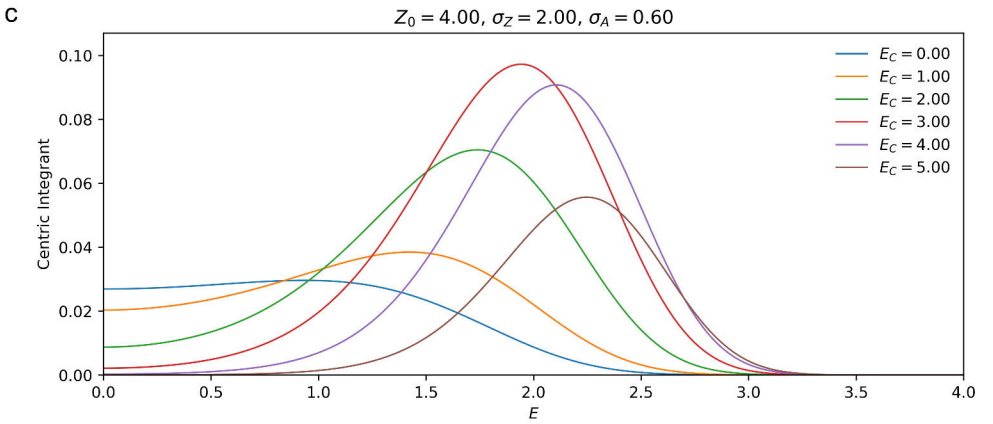
Woolfson, M. M. (1956). *Acta Crystallographica*, **9**(10), 804–810.

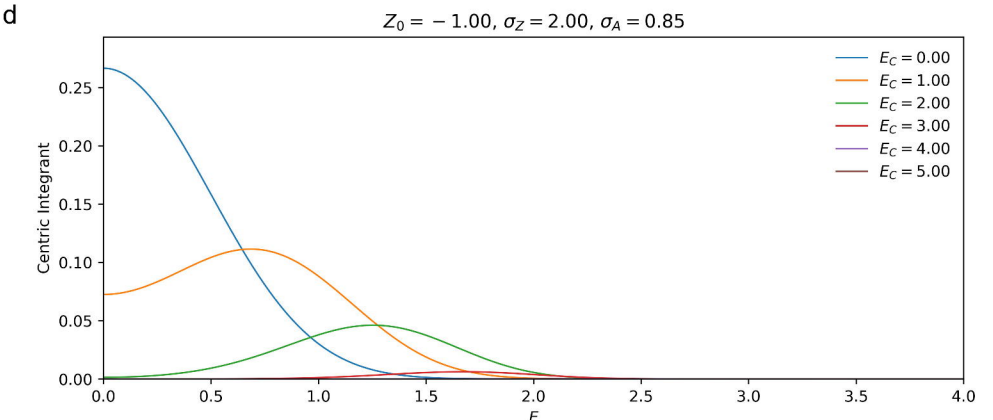
Synopsis

A quadrature is developed that allows for the efficient evaluation of an intensity-based likelihood target function that includes experimental errors.

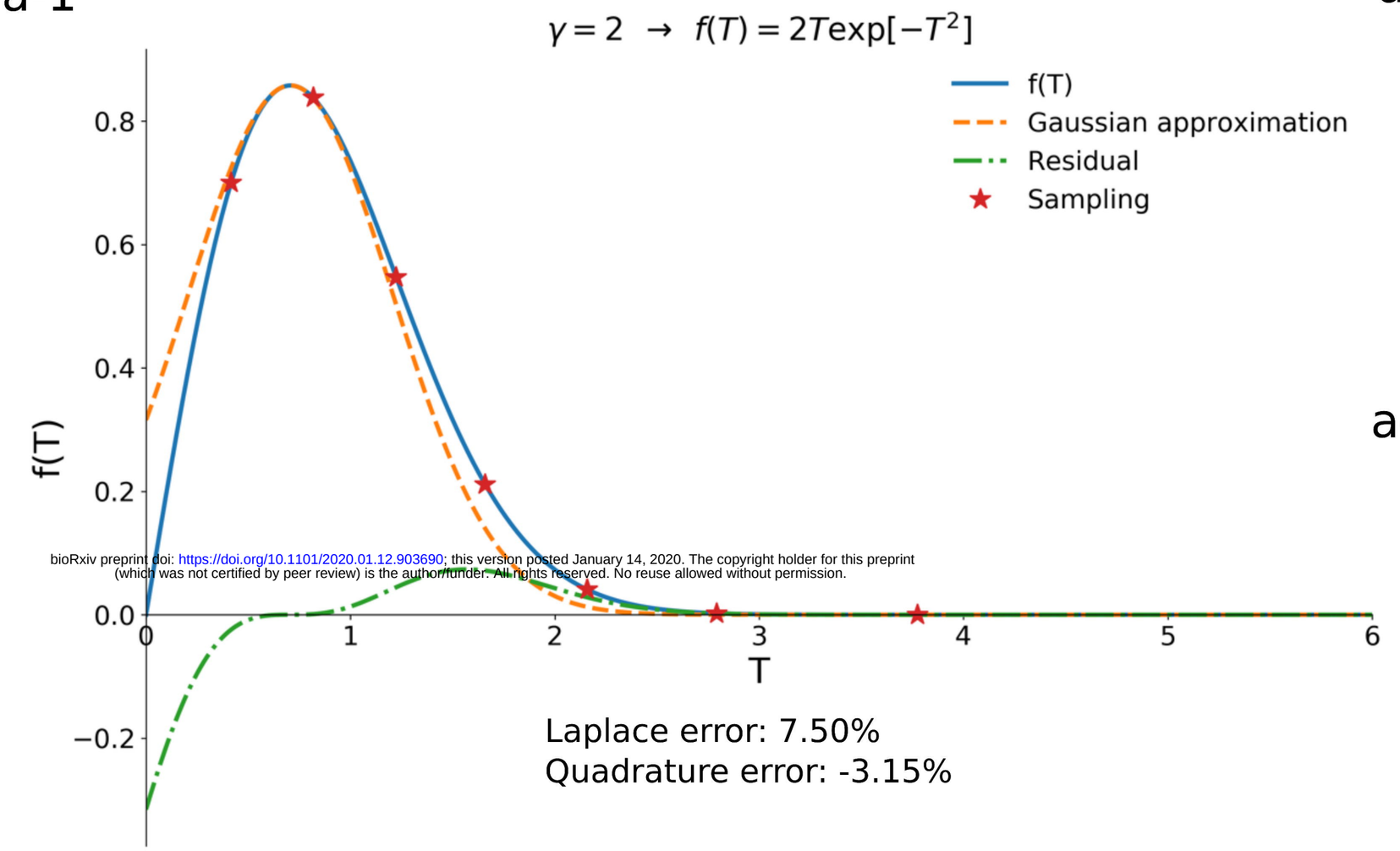

a
$Z_0 = 4.00$, $\sigma_Z = 0.10$, $\sigma_A = 0.60$
Acentric Integrant
E
$E_C = 0.00$
$E_C = 1.00$
$E_C = 2.00$
$E_C = 3.00$
$E_C = 4.00$
$E_C = 5.00$


b
$Z_0 = 4.00$, $\sigma_Z = 2.00$, $\sigma_A = 0.60$
Acentric Integrant
E
$E_C = 0.00$
$E_C = 1.00$
$E_C = 2.00$
$E_C = 3.00$
$E_C = 4.00$
$E_C = 5.00$

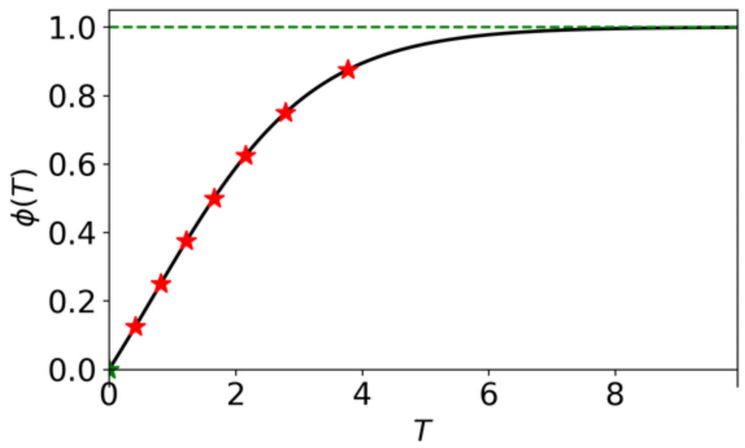

c
$Z_0 = 4.00$, $\sigma_Z = 2.00$, $\sigma_A = 0.60$
Centric Integrant
E
$E_C = 0.00$
$E_C = 1.00$
$E_C = 2.00$
$E_C = 3.00$
$E_C = 4.00$
$E_C = 5.00$


d
$Z_0 = -1.00$, $\sigma_Z = 2.00$, $\sigma_A = 0.85$
Centric Integrant
E
$E_C = 0.00$
$E_C = 1.00$
$E_C = 2.00$
$E_C = 3.00$
$E_C = 4.00$
$E_C = 5.00$

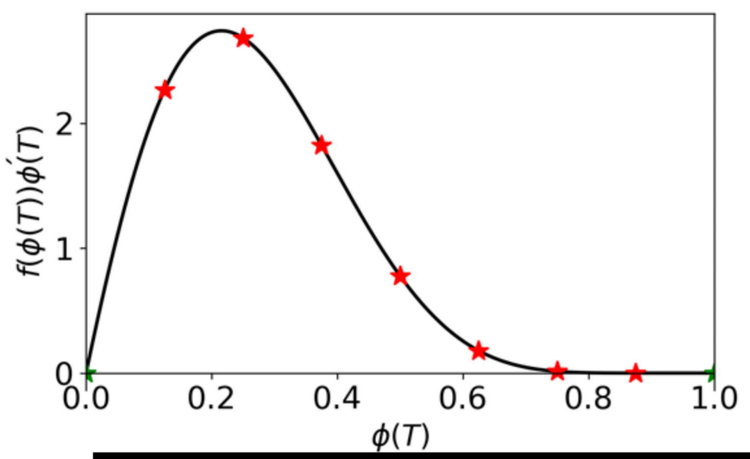
a-1



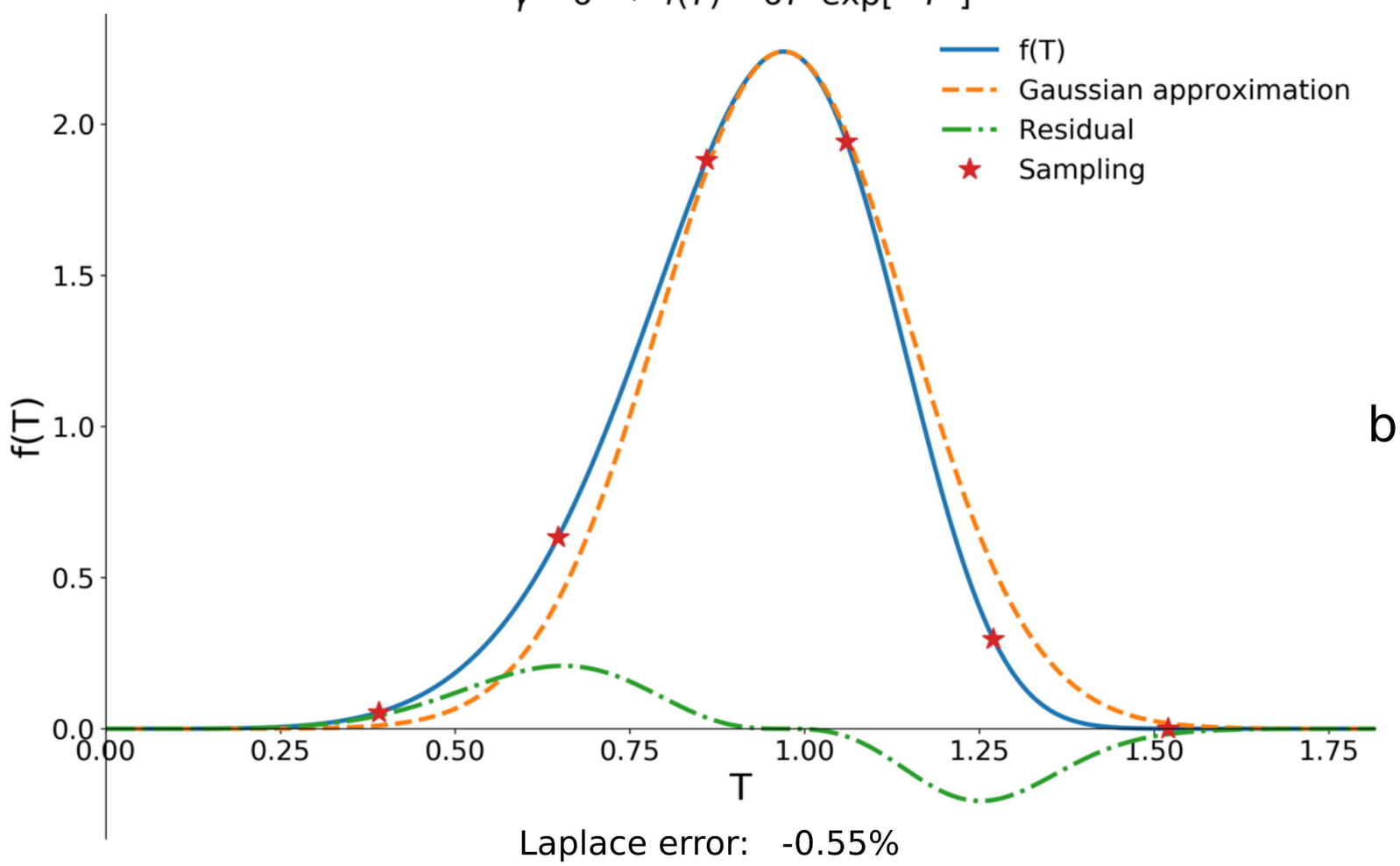
a-2



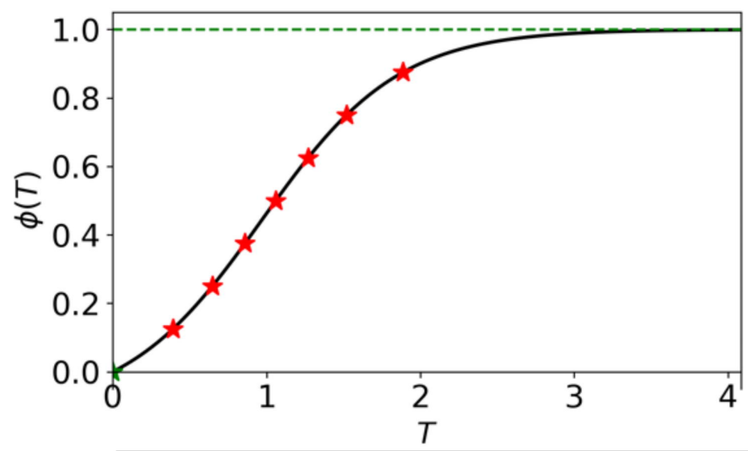
a-3



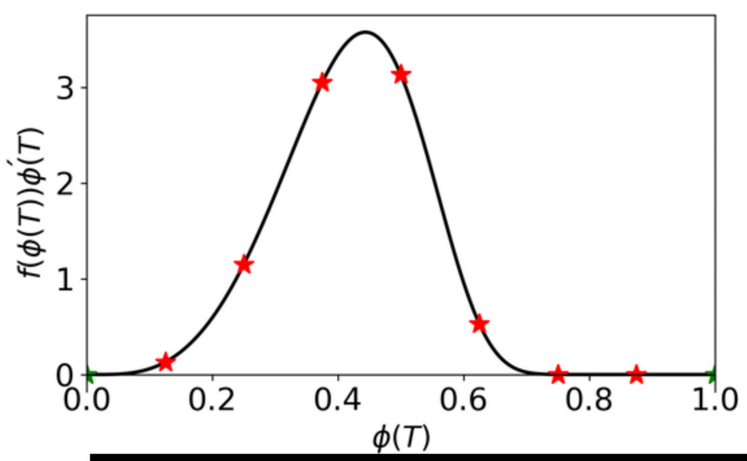
b-1

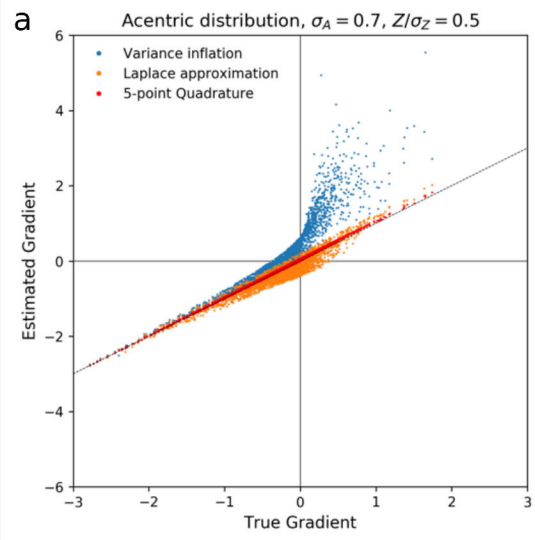


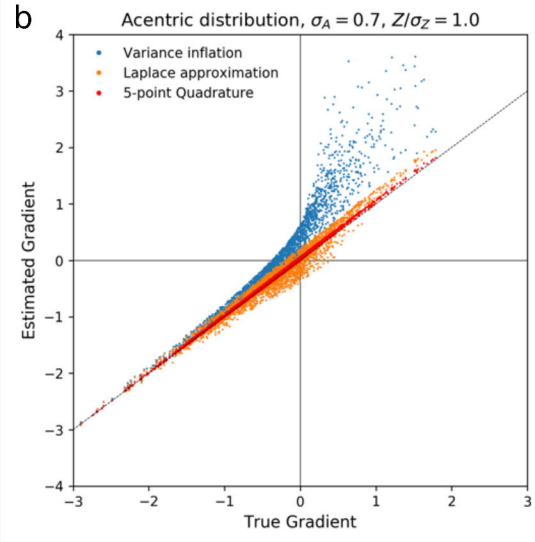
b-2

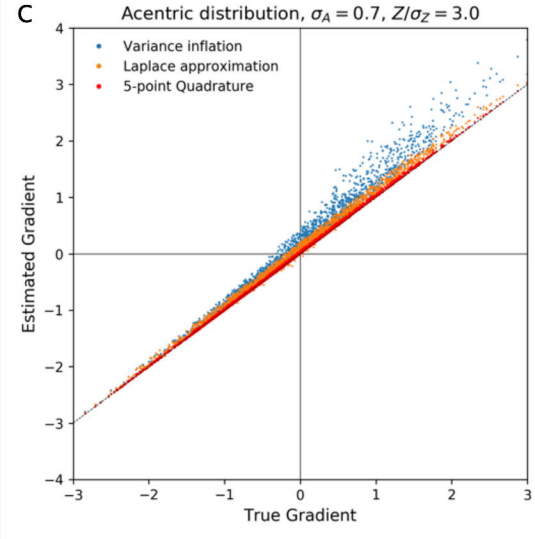


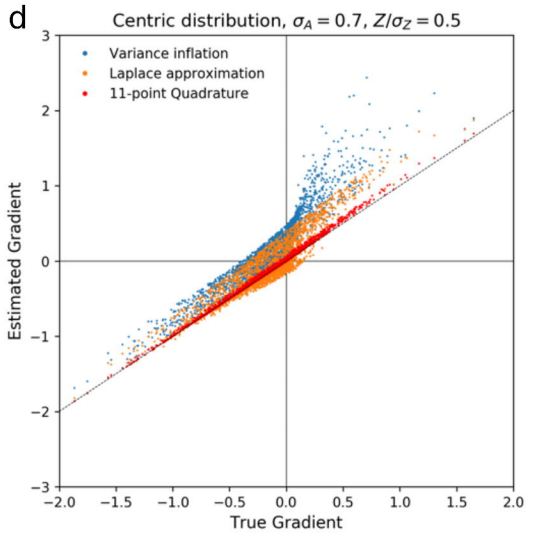
b-3

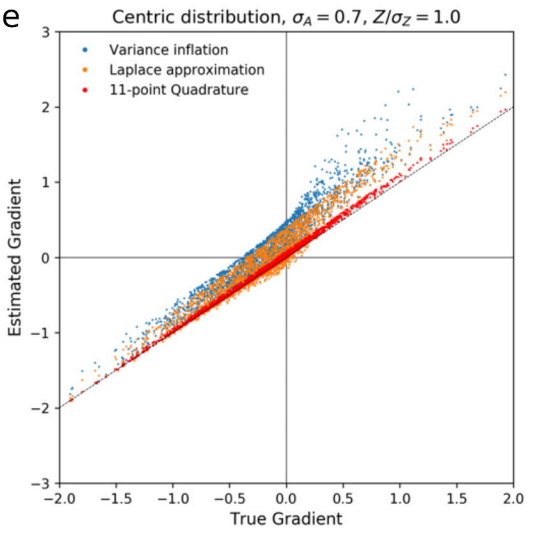


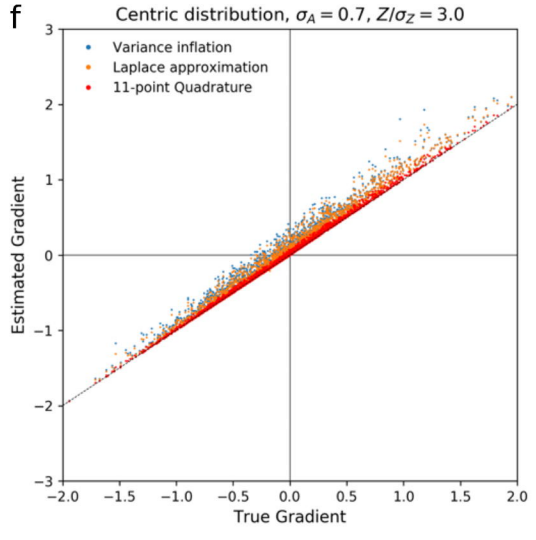

a
Acentric distribution, $\sigma_A = 0.7$, $Z/\sigma_Z = 0.5$
Variance inflation
Laplace approximation
5-point Quadrature
Estimated Gradient
True Gradient


b
Acentric distribution, $\sigma_A = 0.7$, $Z/\sigma_Z = 1.0$
Variance inflation
Laplace approximation
5-point Quadrature
Estimated Gradient
True Gradient


c
Acentric distribution, $\sigma_A = 0.7$, $Z/\sigma_Z = 3.0$
Variance inflation
Laplace approximation
5-point Quadrature
Estimated Gradient
True Gradient


d
Centric distribution, $\sigma_A = 0.7$, $Z/\sigma_Z = 0.5$
Variance inflation
Laplace approximation
11-point Quadrature
Estimated Gradient
True Gradient


e
Centric distribution, $\sigma_A = 0.7$, $Z/\sigma_Z = 1.0$
Variance inflation
Laplace approximation
11-point Quadrature
Estimated Gradient
True Gradient


f
Centric distribution, $\sigma_A = 0.7$, $Z/\sigma_Z = 3.0$
Variance inflation
Laplace approximation
11-point Quadrature
Estimated Gradient
True Gradient