PAPER • OPEN ACCESS

A general approach to maximise information density in neutron reflectometry analysis

View the article online for updates and enhancements.

MACHINE LEARNING
Science and Technology

PAPER

A general approach to maximise information density in neutron reflectometry analysis

Andrew R McCluskey[1,2] , Joshaniel F K Cooper[3], Tom Arnold[4,2,3] and Tim Snow[1,5]

[1] Diamond Light Source, Rutherford Appleton Laboratory, Harwell Science and Innovation Campus, Didcot, OX11 0DE, United Kingdom
[2] Department of Chemistry, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom
[3] ISIS Pulsed Neutron and Muon Source, Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0QX, United Kingdom
[4] European Spallation Source, SE-21100 Lund, Sweden
[5] School of Chemistry, University of Bristol, Bristol, BS8 1TS, United Kingdom

E-mail: andrew.mccluskey@diamond.ac.uk and a.r.mccluskey@bath.ac.uk

Abstract
Neutron and x-ray reflectometry are powerful techniques facilitating the study of the structure of interfacial materials. The analysis of these techniques is ill-posed in nature requiring the application of model-dependent methods. This can lead to the over- and under- analysis of experimental data when too many or too few parameters are allowed to vary in the model. In this work, we outline a robust and generic framework for the determination of the set of free parameters that are capable of maximising the information density of the model. This framework involves the determination of the Bayesian evidence for each permutation of free parameters; and is applied to a simple phospholipid monolayer. We believe this framework should become an important component in reflectometry data analysis and hope others more regularly consider the relative evidence for their analytical models.

1. Introduction

Neutron and x-ray reflectometry methods offer an important insight into the structure of interfacial materials; such as surfactant systems [1], biological membranes [2] and polymeric photovoltaics [3]. However, the analysis of reflectometry data is a non-trivial task, due in part to the phase problem. This is where a direct inversion of the reflectometry profile, to recover the scattering length density profile, $\rho(z)$ is not possible since the phase of the reflected radiation is lost [4]. While there has been much work to retrieve the phase information from a reflectometry profile [5]; from using reference layers [4, 6–8], polarized neutrons [9], a dwell-time approach [10] and from the investigation of multiple contrasts [7, 11–13], these methods are non-trivial and often require a specific experimental setup.

Where it is not possible to retrieve the phases, a model-dependent analysis method can be used to determine the structure of the experimental system [14–20]. This is where a model, typically a series of contiguous layers, is created and the model reflectometry is determined using the Abelès matrix formalism for stratified media [21] or Parratt recursive method [22]. However, the use of model-dependent methods is ill-posed in nature, where the solution may not necessarily be unique [23, 24]. This necessitates the integration of *a priori* information about the underlying physics and chemistry of the experimental system, such as assumptions about molecular volume [1, 25, 26]. This is typically achieved through the re-parameterisation of the model, to describe the system in terms of physiochemical descriptors from which a set of layers can be determined [27–29]. The inclusion of this prior information acts to reduce the dimensionality of the optimisation space and therefore reduces the number of possible structures that would offer satisfactory agreement with the data.

The inclusion of prior information is often achieved by taking literature values from other techniques and applying them inflexibly to constrain values in the re-parameterised model, such as molecular volumes

from computational simulation or solution scattering [30, 31]. However, as was shown in the work of Campbell *et al* [26] and McCluskey *et al* [1], these values are not always strictly representative when the physical or chemical environment is varied. This has lead to the use of Markov-chain Monte Carlo (MCMC) methods for the analysis of reflectometry data, resulting in the inclusion of these methods in many common reflectometry analysis packages [20, 32, 33]. The inclusion of these methods enables the prior probabilities for a series of variables to be outlined, which are considered in a Bayesian fashion during the sampling of the parameter space [1, 34]. This allows for more rational inclusion of the prior information about the experimental system in the analysis, in addition to, the ability to obtain information about the sampled posterior probabilities that are allowed from the experimental data.

The ability to include ranges of values, in the form of uninformative uniform prior probabilities, or normally distributed priors centred on a literature value in the analysis of an experimental dataset is an important development. However, 'with great power, there must also come—great responsibility!' [35] and the ability to sample or fit all of the parameters in an experimental analysis can lead to a high risk of over-fitting [36]. As a result, it is necessary to investigate methods to quantify the information that is available from a given experimental dataset and to enable the comparison of different analytical models; including the number of free parameters that may be considered. A specific approach to this problem is outlined in early works from Sivia and co-workers [37–39], where a two-parameter model was compared with a three-parameter model; with the additional parameter being related to the scaling of the model reflectometry. Sivia and co-workers assumed that the likelihood function for both of these model would have 'only one significant maximum', allowing numerical simplifications to be performed and therefore the problem to be solved analytically. This may not be possible for all models investigated by reflectometry, due again to the ill-posed nature of the analysis problem. Furthermore, the work of Hughes and co-workers [40] more recently applied the use of Bayesian model comparison to determine the mixing of lipids in a model bacterial membrane.

In this work, we outline a general Bayesian framework for the comparison of different analytical models for the rationalisation of reflectometry data. This method can allow the maximum information to be obtained from the analysis of a given dataset, without risking over-fitting. In the supplementary information for this paper, we provide detailed code showing how this framework may be applied to a problem leveraging exclusively open-source software packages.

2. Methods

2.1. Bayesian model selection framework

Bayesian model selection involves the comparison of the posterior probabilities $p(H|\mathbf{D})$ [6], for two models, H_x and H_y, given some dataset, $\mathbf{D}$. The posterior probability for a given model, H_x, can be determined from Bayes theorem [34],

$$p(H_x|\mathbf{D}) = \frac{p(\mathbf{D}|H_x)p(H_x)}{p(\mathbf{D})}, \tag{1}$$

where $p(\mathbf{D}|H_x)$ is the evidence for the model given the data, $p(H_x)$ is the prior probability for the model and $p(\mathbf{D})$ is the probability associated with the measured data, which is equal to the sum of the evidences for every potential model (which is potentially large if not infinite),

$$p(\mathbf{D}) = \sum_{x=1} p(\mathbf{D}|H_x). \tag{2}$$

Therefore, in order to describe the evidence as a truly conditional probability, it is necessary to evaluate very possible model that maybe applied to the experimental data. This is not feasible, therefore in this work we will use the notation $p(\mathbf{D}|H_x)$ for the evidence of a given model, however this is not a strict conditional probability. When comparing the models, concerning the same data, the following relationship can be applied [41],

$$\frac{p(H_x|\mathbf{D})}{p(H_y|\mathbf{D})} = \frac{p(\mathbf{D}|H_x)}{p(\mathbf{D}|H_y)} \times \frac{p(H_x)}{p(H_y)}, \tag{3}$$

where the probabilities associated with the data only cancel out and the relative posterior probabilities depend only on the evidence and prior probabilities for each model. Throughout this work, we have used the

---

[6]In this work, a reduced notation has been used, where the background information parameter, *I* is implied.

Table 1. Interpretation of the Bayes factor, B_{xy} between models x and y.

$2\ln B_{xy}$	Evidence against H_y
$[0, 2)$	Not worth more than a bare mention
$[2, 6)$	Positive
$[6, 10)$	Strong
$[10, \infty)$	Very strong

uninformative assumption that the prior probability for all models is 1. Therefore, it is only the evidence for each model, for the given data, that influences the posterior probability.

Nested sampling is a method for Bayesian evidence determination that was developed by Skilling [42] and is implemented in the Python package `dynesty` [43]. This sampling method allows for the evidence determination through the estimation of the, possibly multi-dimensional, integral under the likelihood, $\mathcal{L}(\boldsymbol{X}|H)$ and parameter specific prior $p(\boldsymbol{X}|H)$ [44],

$$p(\mathbf{D}|H) = \iint_{\boldsymbol{R}} \mathcal{L}(\boldsymbol{X}|H)p(\boldsymbol{X}|H)\, \mathrm{d}^M\boldsymbol{X}, \tag{4}$$

where $\boldsymbol{X}$ is a vector of length M of the varying parameters in the model, $\boldsymbol{R}$ is a $2 \times M$ matrix that describes the range over which the integral should be evaluated and $p(\boldsymbol{X}|H)$ is the parameter specific prior. Nested sampling involves 'intelligently' sampling this integral until some stopping criterion is achieved. This stopping criterion is when the following is less than a given value,

$$\triangle\ln\{\hat{p}(\boldsymbol{D}|H)_i\} = \ln\{\hat{p}(\boldsymbol{D}|H)_i - \Delta\hat{p}(\boldsymbol{D}|H)_i\} - \ln\hat{p}(\boldsymbol{D}|H)_i, \tag{5}$$

where $\hat{p}(\boldsymbol{D}|H)_i$ is the current estimated evidence and $\Delta\hat{p}(\boldsymbol{D}|H)_i$ can be estimated from the prior volume of the previous point in the sampling. In this work, a stopping criteria of $\triangle\ln\{\hat{p}(\boldsymbol{D}|H)\} = 0.5$ was used.

The likelihood is the difference between the data and model and is defined in this work as it is implemented in the `refnx` reflectometry package [20, 45],

$$\ln\mathcal{L} = -\frac{1}{2}\sum_{i=1}^{N}\left(\left\{\frac{R(q_i) - R(q_i)_m}{\delta R(q_i)}\right\}^2 + \ln[2\pi\{\delta R(q_i)\}^2]\right), \tag{6}$$

where, N is the number of measured q-vectors, q_i is the ith q-vector, $R(q_i)$ is the experimental reflected intensity at the ith q-vector, $\delta R(q_i)$ is the uncertainty in the experimental reflected intensity at the ith q-vector and $R(q_i)_m$ is the model reflected intensity at the ith q-vector determined from the implementation of the Abelès matrix formalism for stratified media [21, 22] in `refnx`.

The evidence allows for the comparison of two models, where the data are the same. It is only possible to compare a given pair of models for the same data, as in equation (3) the data-only dependent probabilities were cancelled out as they were equal. The Bayes factor, B_{xy}, is the name given to the ratio between the evidence for two models,

$$\ln B_{xy} = \ln\left[\frac{p(\boldsymbol{D}|H_x)}{p(\boldsymbol{D}|H_y)}\right] = \ln[p(\boldsymbol{D}|H_x)] - \ln[p(\boldsymbol{D}|H_y)]. \tag{7}$$

The interpretation of the Bayes factor was outlined by Kass and Raftery [46], which is given in table 1. This method is easily generalised to any model-dependent experiment and has been applied in areas such as systems biology, astronomy, computational chemistry and indeed, in a less general fashion than the current work, to neutron reflectometry [39, 41, 47, 48].

The Bayesian approach allows for the comparison of models where there are different numbers of free parameters, without the risk of overfitting and 'fitting an elephant' [36]. This is due to the Bayesian evidence being derived from an integral in parameter space and therefore scaling with the number of parameters. This means that the addition of an additional free parameter to the model must offer a significant improvement to the likelihood determined. However, it is important also to note that the accuracy of the determined evidence depends on the prior probabilities chosen for each of the free parameters and therefore care should be taken to ensure that these are meaningful.

Table 2. The shorthands used for the different combinations of phospholipid monolayer and water investigated.

Shorthand	Phospholipid contrast	Water contrast
h-D_2O	h-DSPC	D_2O
d_{13}-D_2O	d_{13}-DSPC	D_2O
d_{13}-ACMW	d_{13}-DSPC	ACMW
d_{70}-D_2O	d_{70}-DSPC	D_2O
d_{70}-ACMW	d_{70}-DSPC	ACMW
d_{83}-D_2O	d_{83}-DSPC	D_2O
d_{83}-ACMW	d_{83}-DSPC	ACMW

Table 3. The scattering lengths, b_i, for each of the DSPC phospholipid components in this work; where the subscript indicates either the head, *h*, or tail, *t*.

Phospholipid contrast	b_h/fm	b_t/fm
h-DSPC	6.01	-3.58
d_{13}-DSPC	19.54	-3.58
d_{70}-DSPC	11.21	69.32
d_{83}-DSPC	24.75	69.32

2.2. Neutron reflectometry measurements

The neutron reflectometry measurements investigated in this work were previously published by Hollinshead *et al* [49] and subsequently re-analysed by McCluskey *et al* [50]. The measurements concern the study of a monolayer of 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC) at the air/water interface and were conducted on seven different isotopic contrasts of the phospholipid and water. These contrasts were made up from four phospholipid types; fully-hydrogenated (h-DSPC), head-deuterated (d_{13}-DSPC), tail-deuterated (d_{70}-DSPC) and fully deuterated (d_{83}-DSPC). These were paired with two water contrasts; D_2O and air-contrast matched water (ACMW, where D_2O and H_2O are mixed such that the resulting scattering length density is zero). The pairing of the fully-hydrogenated phospholipid with ACMW was not investigated, most likely due to the lack of the scattering available from such a system. Table 2 gives the shorthands that are used in this work to refer to the different contrast pairings investigated. The data analysed in this work was taken at a surface pressure of 30 mNm^{-1}. Additional details of the data collection may be found in the original publication [49].

2.3. Analytical model

In the model-dependent analysis of neutron reflectometry data, it is necessary to have a description of the experimental system to which the Abelès matrix formalism may be applied. Campbell *et al* investigated different models for the study of phospholipid monolayers and found a two-layer description that differentiates between the phospholipid heads, h, and tails, t, to be the best at reproducing the reflected intensity [26]. The head layer is adjacent to the solvent, while the tail layer is adjacent to the air. The thickness of the tail layer is defined as d_t and the head layer is d_h. At each of the three interfaces, there is some interfacial roughness, σ, which is modelled with an error function [51]. In this work, the roughness was taken to be conformal, which is to say that it does not vary between interfaces, in agreement with the work of Campbell *et al* [26]. This final assumption is reasonable given the fact that this system consists of a monolayer of a single lipid type, however, this may not be reasonable in more complex systems.

The scattering length density, ρ_i, for a layer *i*, from which the model reflectometry profile is obtained, is calculated from the volume fraction of the chemical component, ϕ_i, the scattering length of the component, b_i (found in table 3) and its molecular volume, V_i,

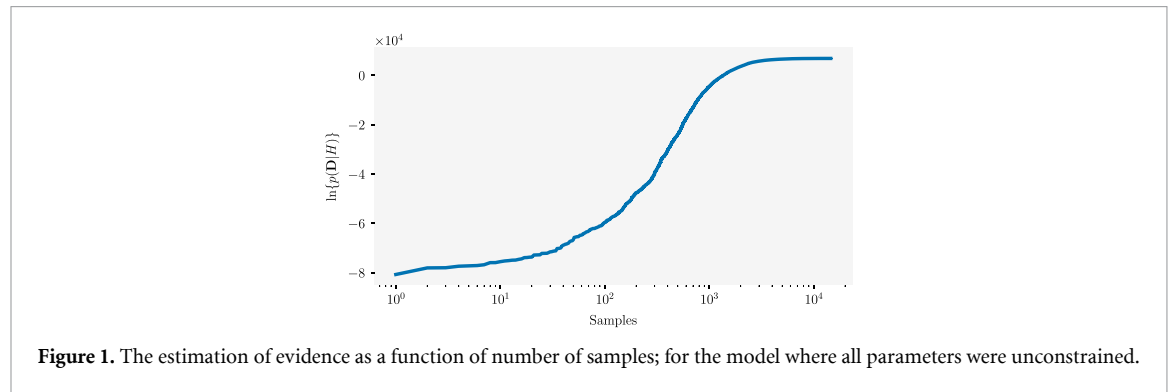$$\rho_i = \frac{b_i}{V_i}\phi_i + \rho_{H_2O}(1 - \phi_i). \tag{8}$$

where ρ_{H_2O} is the scattering length density of water, which is 0Å^{-2} for ACMW and 6.35×10^{-6}Å^{-2} for D_2O. To preserve the chemical bonding between the heads and tails of the phospholipid, the following constraint is included in the model,

$$\phi_h = \frac{d_t V_h \phi_t}{V_t d_h}. \tag{9}$$

This ensures that the number of head groups is equal to the number of pairs of tail groups.

Table 4. The constrained and uniform prior probability range for each of the free parameters in the analytical model.

Parameter	Constrained Value	Prior Range
$d_h/\text{Å}$	10.0^7	$[8.0, 16.0)$
$V_h/\text{Å}^3$	339.5^8	$[300.0, 380.0)$
$d_t/\text{Å}$	21.0^9	$[10.0, 26.0)$
ϕ_t	1.0^{10}	$[0.5, 1.0)^{11}$
$V_t/\text{Å}^3$	850.4^{12}	$[800.0, 1\,000.0)$
$\sigma/\text{Å}$	2.9^{13}	$[2.9, 5.0)^{14}$

Figure 1. The estimation of evidence as a function of number of samples; for the model where all parameters were unconstrained.

The model outlined above has up to six possible free parameters; the two-layer thicknesses, d_t and d_h, the two molecular volumes, V_t and V_h, the volume fraction of the tail layer, ϕ_t and the conformal interfacial roughness, σ. In this work, the evidence as a function of these parameters is of interest as it is possible to use this evidence and Bayes factor to determine the best set of free parameters for a given set of data. Therefore, in table 4 the uniform prior probability range and value when constrained is given for each of the parameters. These six parameters can give rise to 63 permutations where there is at least one free parameter. These permutations constitute every possible combination of the six parameters, either held constant or being allowed to vary within the prior probability. This allows for a complete understanding of which combinations of constrained and free parameters give rise to the highest evidence and therefore represents the maximum information density that may be obtained from the given data. We note again, that both the chosen prior and the constrained values are not infallible and will influence the result.

All seven isotopic neutron contrasts were analysed with a single structural model. In doing so we make the common assumption that there is no effect on the structure as a result of the contrast variation method [16]. Additionally, the reflected intensity from the model was scaled and a uniform background was added based on values obtained in the previous analysis of this dataset [50]. We accept that these parameters may influence the resulting evidence, as was shown in the work of Sivia and Webster [39].

3. Results & discussion

3.1. All parameters

Figure 1 shows the variation of the estimated evidence as a function of iterations, for the model where all six parameters were unconstrained. This plot shows clearly the ability for the nested sampling method to determine a good estimate of the evidence, with the plot plateauing at $6\,859.9 \pm 0.3$ after more than $10\,000$ samples. This was the slowest model to converge, as it was sampling the high-dimensionality parameter space, on a workstation computer this nested sampling process took around 20 min to converge leading to the results provided here.

In addition to the ability for the nested sampling method to determine the evidence for a given model, it also allows for the posterior probabilities of the analysis to be evaluated. Figure 2 gives the posterior probabilities and cross-correlations between each of the six parameters. It is clear that the volume fraction of the tail layer, ϕ_t and the interfacial roughness, σ, both are not completely represented by the defined prior probabilities. Hence, there lacks a full Gaussian distribution and an extremely skewed probability distribution is visible. However, maximum for ϕ_t and minimum for σ both come from physical limitations on the system; the former being that the volume fraction cannot be greater than 1 and the latter being that a surfactant monolayer will increase the surface roughness compared to water [57], which is ~ 2.9 Å [55, 56]. The skew-ness of $p(\phi_t)$ is to be expected given the hydrophobic interaction between the phospholipid tails which create the situation where it is highly unlikely for water to be intercalated into the tail structure. The
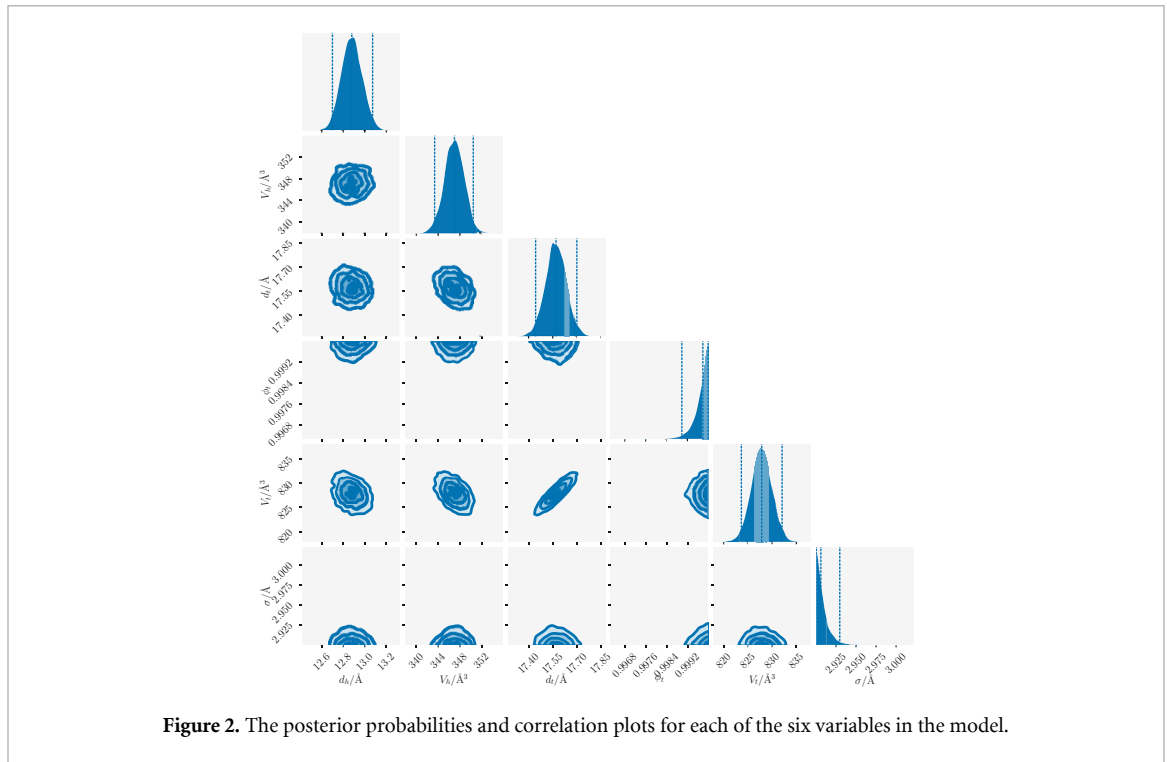
Figure 2. The posterior probabilities and correlation plots for each of the six variables in the model.
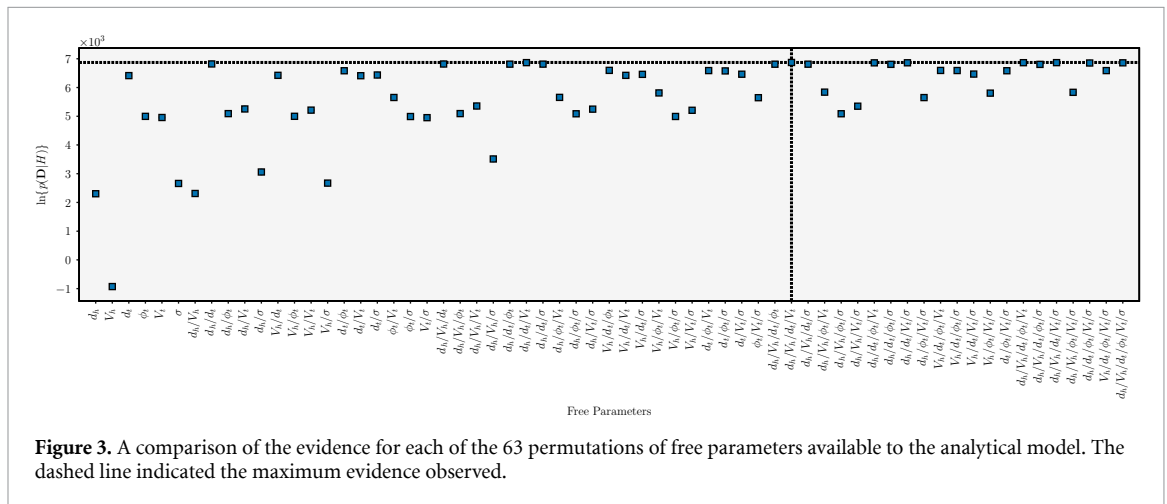

Figure 3. A comparison of the evidence for each of the 63 permutations of free parameters available to the analytical model. The dashed line indicated the maximum evidence observed.

nature of $p(\sigma)$ is similar in that the expected value of 2.9Å is the minimum value available to the distribution. However, it may be the case that by fitting all possible parameters, the data is being overfitted. Luckily, the Bayesian model comparison method allows for the evidence comparison to fit the model which offers the maximum information to be obtained, without over-fitting to the data.

3.2. Greatest evidence

The evidence for all of the 63 possible permutations of the free parameters was determined. This allows for the comparison of these to determine which model (set of varying parameters) offers the most information from the dataset, without the over-fitting to the data that may be occurring when all six parameters are varied. This comparison plotted in figure 3. The dashed black line indicates the set of free parameters that offer the greatest evidence for the co-refinement of the seven contrasts of neutron reflectometry data, which in this case is $d_h/V_h/d_t/V_t$, which has an evidence of 6 872.1 $\pm$ 0.2. The set of parameters with the next greatest evidence is $d_h/d_t/V_t$, however, the evidence for this set of parameters is 6 866.8 $\pm$ 0.2. The interpretation of the Bayes factor between these two sets of parameters gives $2\ln B_{xy} = 10.7 \pm 0.7$ where, $d_h/V_h/d_t/V_t$ is x and $d_h/d_t/V_t$ is y. This indicates that there strong or very strong evidence for the more complex model and that the maximum information density that can be obtained from the analysis at hand, without over-fitting, is the fitting of $d_h/V_h/d_t/V_t$.
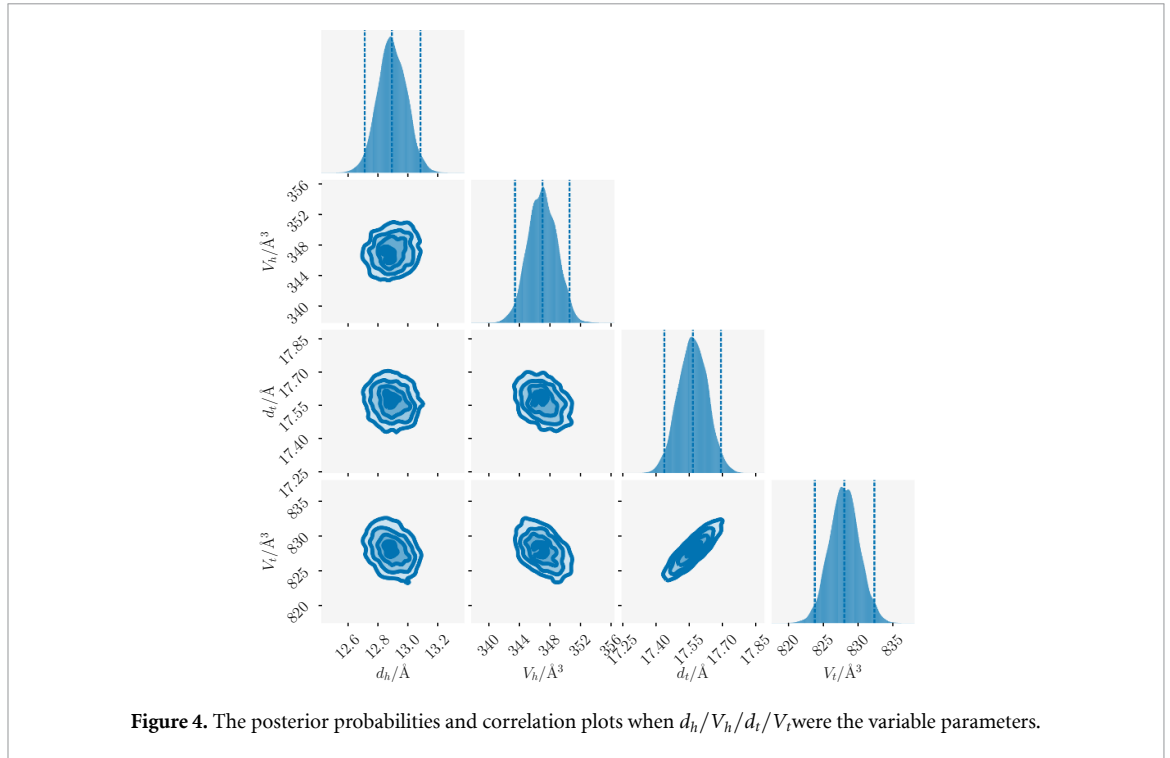
Figure 4. The posterior probabilities and correlation plots when $d_h/V_h/d_t/V_t$ were the variable parameters.

Table 5. The median values and 95% confidence intervals for each of the free parameters in the model with the greatest evidence.

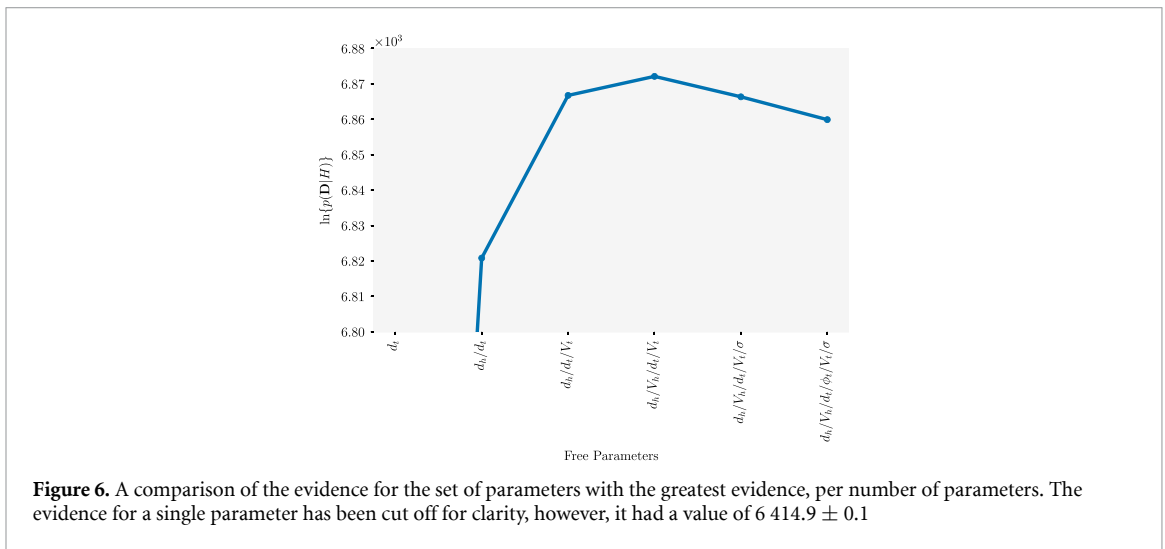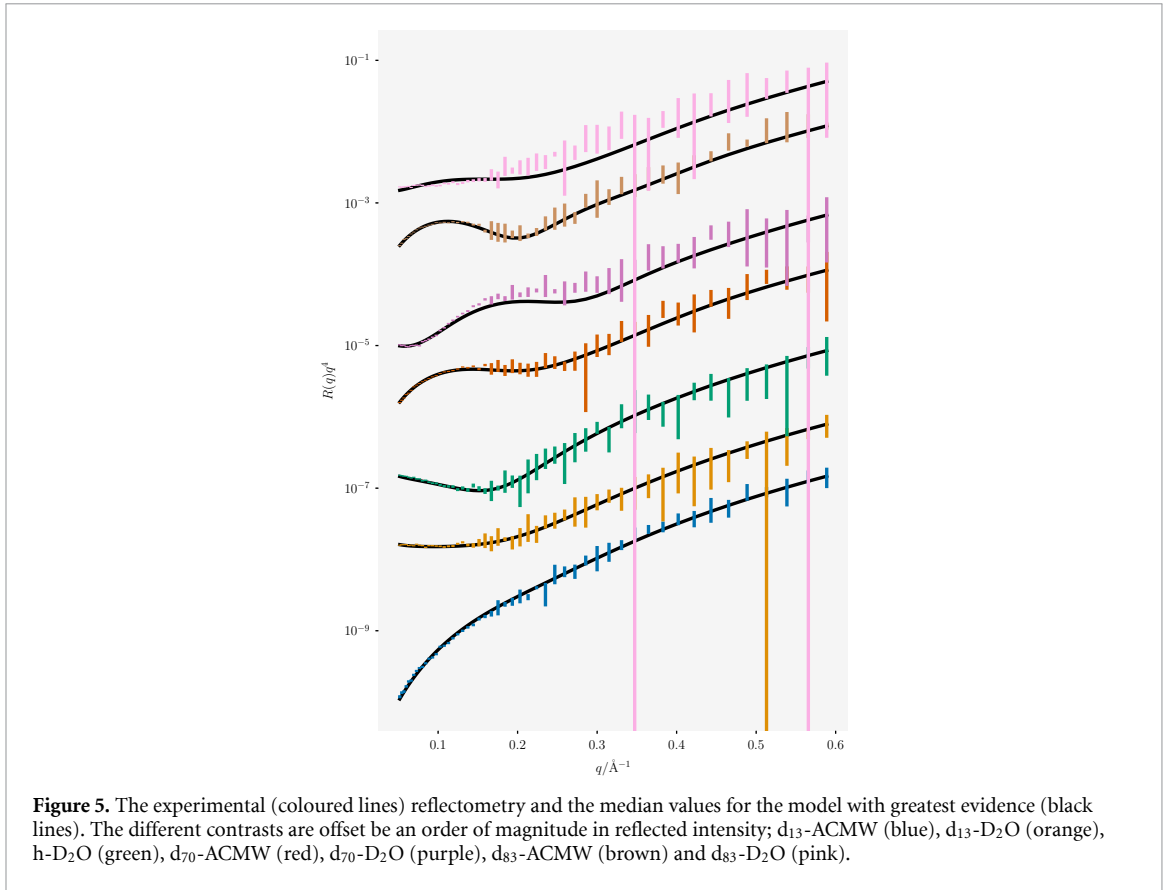| Free Parameter	Value
$d_h/\text{Å}$	$12.9^{+0.2}_{-0.2}$
$V_h/\text{Å}^3$	$347.0^{+3.6}_{-3.6}$
$d_t/\text{Å}$	$17.6^{+0.1}_{-0.1}$
$V_t/\text{Å}^3$	$828.0^{+4.3}_{-4.2}$

Figure 4 shows the posterior probability distributions that were obtained from the $d_h/V_h/d_t/V_t$ set of parameters. The median values for each of the probability distributions and their 95% confidence intervals are given in table 5. The value found here for the phosphatidylcholine is slightly larger than the 10Å used by Campbell *et al* [26] and the 10.5Å quoted by Li *et al* [58]. The thickness of the tail layer that is observed is less than that suggested previously [26]. However, it agrees well with the results from the original analysis of this data [49], which itself was consistent with previous x-ray reflectometry measurements [59–61]. The tail molecular volume agrees well with the compression of $\sim$12%, when compared to the values from other experimental techniques [30, 31], suggested by Campbell *et al*.

The reflectometry profiles for the median values are compared with the experimental data in figure 5. There is clear agreement between the experiment and the model across all of the contrasts. This indicates that it is likely that there is little variation due to the isotopic contrast variation that was used and the assumption that is discussed in 2.3 is valid.

3.3. Comparison with previous analysis

As mentioned above, this particular experimental dataset was previously analysed by McCluskey *et al* [50]. The analytical model used in this previous work was similar to that used herein, however, the tail volume of the phospholipid was constrained based on the assumption that the measured area per molecule was accurate and the volume fraction of the tail layer is 1. That model had three free parameters, namely d_h, d_t and σ.

This previous analysis made no effort to evaluate different models of analysis or numbers of free parameters. Therefore, it is not clear if the three parameters chosen offered the maximum information density available. While not completely comparable, as the constrained values in this work were different to those used previously, the suggestion from this work that having $d_h/V_h/d_t/V_t$ as the free parameters do not match with those probed previously. In future, we hope that by outlining this general methodology, others will consider the evidence of their particular model and the free parameter being used during their analysis.

Figure 5. The experimental (coloured lines) reflectometry and the median values for the model with greatest evidence (black lines). The different contrasts are offset be an order of magnitude in reflected intensity; d_{13}-ACMW (blue), d_{13}-D_2O (orange), h-D_2O (green), d_{70}-ACMW (red), d_{70}-D_2O (purple), d_{83}-ACMW (brown) and d_{83}-D_2O (pink).

Figure 6. A comparison of the evidence for the set of parameters with the greatest evidence, per number of parameters. The evidence for a single parameter has been cut off for clarity, however, it had a value of $6\,414.9 \pm 0.1$

3.4. Best per number of parameters

Finally, we will briefly present the best set of parameters for each number of parameters. These are shown in figure 6, which has the characteristic shape found in the previous work by Sivia and others [37, 39, 44, 62]. This plot can be seen to peak when there are four free parameters, before dropping off slowly. This drop off indicates that when more parameters are included there is an indication that over-fitting is occurring. This plot shows clearly where the maximum information is available in the analysis for this particular reflectometry data.

4. Conclusions

This work outlines a general application of a Bayesian framework to the problem of model selection in neutron reflectometry. Building on the work of Sivia and co-workers in the 1990s [39], we have utilised a nested sampling method to generalise the methodology beyond the assumption of only one significant

maximum [37]. This offers the ability to go beyond the three-parameter space previously investigated to consider every possible parameter in an analysis problem. In doing so, it is possible to compare the evidence of different models, where different numbers of parameters are allowed to vary, to determine that which offers the maximum information, without over-fitting.

To show this in action, we have investigated the analysis of seven isotopic contrasts of neutron reflectometry data from a phospholipid monolayer at the air/water interface. This particular analysis had a maximum of six variable parameters and in this work, the evidence of all 63 possible permutations of these parameters was determined. This allowed for the permutation that offered the maximum evidence to be found, which under the circumstances of the chosen priors was when $d_h/V_h/d_t/V_t$ were varied. This particular combination of parameters had an evidence of $6\,872.1 \pm 0.2$, while the next best, which had more free parameters had a value of $6\,866.8 \pm 0.2$. This was interpreted in terms of the Bayes factor and it was shown that there was strong evidence for the $d_h/V_h/d_t/V_t$ model. Finally, we have compared these free parameters with those used previously to analyse the same dataset [50]. This comparison indicates that this previous analysis did not necessarily maximise the information density available to the data.

We accept that the relative simplicity of this model system allows for the use of nested sampling on a relatively short timescale, with all of the computation capable of being performed in a single day on a workstation machine. However, the methods discussed herein are relevant to any model-dependent reflectometry analysis, but we note that as the number of free parameters increases so does the dimensionality of the parameter space being sampled, leading to longer computational timescales. Therefore, we propose that for analyses with a large number of parameters the analyst starts by reducing the number of free parameters as much as is possible to ensure that the nested sampling will converge in a reasonable timescale and gradually increasing the complexity. We hope that by developing and sharing of this methodology, in future those analysing reflectometry data will consider which of the free parameters in their models may be varied and will try to determine the evidence for their particular model for other options before completing their analysis.

Reproducibility statement

Electronic Supplementary Information (ESI) available: All analysis/plotting scripts and data files allowing for a fully reproducible and automated, analysis workflow for the work presented is available at https://github.com/arm61/model_select (DOI: 10.5281/zenodo.3697795) under a CC BY-SA 4.0 license.

Author contributions

A.R.M. proposed, implemented and applied the Bayesian model comparison framework, with input from T.A., J.F.K.C. and T.S.; J.F.K.C. and T.S. sourced funding; A.R.M. wrote the manuscript, with input from all authors.

Acknowledgments

ORCID iDs

Andrew R McCluskey ⬥ https://orcid.org/0000-0003-3381-5911
Tom Arnold ⬥ https://orcid.org/0000-0001-8295-3822
Tim Snow ⬥ https://orcid.org/0000-0001-7146-6885

References

[1] McCluskey A R, Sanchez-Fernandez A, Edler K J, Parker S C, Jackson A J, Campbell R A and Arnold T 2019 Bayesian determination of the effect of a deep eutectic solvent on the structure of lipid monolayers *Phys. Chem. Chem. Phys.* **21** 6133

[2] Clifton L A, Paracini N, Hughes A V, Lakey J H, Steinke N-J, Cooper J F K, Gavutis M and Skoda M W A 2019 Self-assembled fluid phase floating membranes with tunable water interlayers *Langmuir* **35** 13735

[3] Pérez G E, Bernardo G, Gaspar H, Cooper J F K, Bastianini F, Parnell A J and Dunbar A D F 2019 Determination of the thin-film structure of Zwitterion-doped poly(3,4-ethylenedioxythiophene):Poly(styrenesulfonate): A neutron reflectivity study *ACS Appl. Mater. Interfaces* **11** 13803

[4] Majkrzak C F and Berk N F 1995 Exact determination of the phase in neutron reflectometry *Phys. Rev.* B **52** 10827

[5] Kirby B J, Kienzle P A, Maranville B B, Berk N F, Krycka J, Heinrich F and Majkrzak C F 2012 Phasesensitive specular neutron reflctometry for imaging the nanometer scale composition depth profile of thin-film materials *Curr. Opin. Colloid Interface Sci.* **17** 44

[6] de Haan V-O, van Well A A, Adenwalla S and Felcher G P 1995 Retrieval of phase information in neutron reflectometry *Phys. Rev.* B **52** 10831

[7] Majkrzak C F, Berk N F and Perez-Salas U A 2003 Phase-sensitive neutron reflectometry *Phys. Rev.* B **19** 7796

[8] Nikova E S, Salamatov Y A, Kravtsov E A and Ustinov V V 2019 Ustinov, Novel method of phase determination in neutron reflectometry using reference layer *J. Phys.: Conf. Ser* **1389** 012153

[9] Leeb H, Kasper J and Lipperheide R 1998 Determination of the phase in neutron reflectometry by polarization measurements *Phys. Lett.* A **239** 147

[10] Fiedeldey H, Lipperheide R, Leeb H and Sofianos S A 1992 A proposal for the determination of the phases in specular neutron reflection *Phys. Lett.* A **170** 347

[11] Majkrzak C F and Berk N F 1995 Exact determination of the phase in neutron reflectometry by variation of the surrounding media *Phys. Rev.* B **58** 15416

[12] Majkrzak C F *et al* 2000 First-principles determination of hybrid bilayer membrane structure by phase-sensitive neutron reflectometry *Biophys.* J. **79** 3330

[13] Koutsioubas A 2019 Model-independent recovery of interfacial structure from multi-contrast neutron reflectivity data *J. Appl. Crystallogr* **52** 538

[14] Pedersen J S and Hamley I W 1994 Analysis of neutron and x-ray reflectivity data by constrained least-squares methods *Physica* **198** 16

[15] de Haan V-O and Drijkoningen G G 1994 Genetic algorithms used in model finding and fitting for neutron reflection experiments *Physica* **198** 24

[16] Nelson A R J 2006 Co-refinement of multiple-contrast neutron/X-ray reflectivity data using MOTOFIT *J. Appl. Crystallogr.* **39** 273

[17] van der Lee A, Salah F and Harzallah B 2007 A comparison of modern data analysis methods for x-ray and neutron specular reflectivity data *J. Appl Crystallogr* **40** 820

[18] Gerelli Y 2016 Aurore: new software for neutron reflectivity data analysis *J. Appl Crystallogr* **49** 330

[19] Gerelli Y 2016 Aurore: new software for neutron refectivity data analysis corrigendum *J. Appl. Crystallogr.* **49** 712

[20] Nelson A R J and Prescott S W 2019 refnx: neutron and x-ray reflectometry analysis in Python *J. Appl. Crystallogr.* **52** 193

[21] Abelès F 1948 Sur la propagation des ondes électromagnétiques dans les mulieux sratifiés *Ann. Phys* **12** 504

[22] Parratt L G 1954 Surface studies of solids by total reflection of x-rays *Phys. Rev* **95** 359

[23] Klibanov M V and Sacks P E 1992 Phaseless inverse scattering and the phase problem in optics *J. Math. Phys.* **33** 3813

[24] Klibanov M V and Sacks P E 1994 Use of partial knowledge of the potential in the phase problem of inverse scattering *J. Comp. Phys.* **112** 273

[25] Waldie S *et al* 2018 Localization of cholesterol within supported lipid bilayers made of a natural extract of tailor-deuterated phosphatidylcholine *Langmuir* **34** 472

[26] Campbell R A, Saaka Y, Shao Y, Gerelli Y, Cubitt R, Nazaruk E, Matyszewska D and Lawrence M J 2018 Structure of surfactant and phospholipid monolayers at the air/water interface modeled from neutron reflectivity data *J. Colloid Interface Sci.* **531** 98

[27] Schalke M and Lösche M 2000 Structural models of lipid surface monolayers from x-ray and neutron reflectivity measurements *Adv. Colloid Interface Sci* **88** 243

[28] Heinrich F and Lösche M 2014 Zooming in on disordered systems: Neutron reflection studies of proteins associated with fluid membranes *BBA-Biomembranes* **1838** 2341

[29] Heinrich F 2016 Kelman Z 211–30 *Isotope Labeling of Biomolecules - Applications, series Methods in Enzymology* vol 566 editor ed (New York: Academic)

[30] Sun W-J, Suter R M, Knewtson M A, Worthington C R, Tristram-Nagle S, Zhang R and Nagle J F 1994 Order and disorder in fully hydrated unoriented bilayers of gel-phase dipalmitoylphosphatidylcholine *Phys. Rev. Lett.* **49** 4665

[31] Armen R S, Uitto O D and Feller S E 1998 Phospholipid component volumes: determination and application to bilayer structure calculations *Biophys.* J. **75** 734

[32] Hughes A V 2017 Rascal sourceforge https://sourceforge.net/projects/rscl/ note accessed: 2019-10-09

[33] Kienzle P A, Maranville B B, O'Donovan K V, Ankner J F, Berk N K and Majkrzak C F 2017 NCNR reflectometry software https://www.nist.gov/ncnr/reflectometry-software note accessed: 2019-09-15

[34] Bayes T 1763 An essay towards solving a problem in the doctrine of chances *Phil. Trans. Roy. Soc.* **53** 370

[35] Lee S and Ditko S 1962 Introducing Spiderman *Amazing Fantasy (Marvel Comics)* **15** 11

[36] Mayer J, Khairy K and Howard J 2010 Drawing an elephant with four complex parameters *Am. J. Phys* **78** 648

[37] Sivia D S, Hamilton W A and Smith G S 1991 Analysis of neutron reflectivity data: maximum entropy, Bayesian spectral analysis and speckle holography *Physica* **173** 121

[38] Geoghagan M, Jones R A L, Sivia D S, Penfold J and Clough A S 1996 Experimental study of surface segregation and wetting in films of a partially miscible polymer blend *Thin Solid Films* **53** 825

[39] Sivia D S and Webster J R P 1998 The Bayesian approach to reflectivity data *Physica* **248** 327

[40] Hughes A V, Patel D S, Widmalm G, Klauda J B, Clifton L A and Im W 2019 Physical properties of bacterial outer membrane models: neutron reliectometry & molecular simulation *Biophys. J.* **116** 1095

[41] Pullen N and Morris R J 2014 Bayesian model comparison and parameter inference in systems biology using nested sampling *PLOS ONE* **9** e88419

[42] Skilling J 2006 Nested sampling *Conf. Proc.* **735** 395

[43] Speagle J S 2020 dynesty: A Dynamic Nested Sampling Package for Estimating Bayesian Posteriors and Evidences *MNRAS* **493** 3132–58

[44] Sivia D S and Skilling J 2005 *Data Analysis: A Bayesian Tutorial* 2nd ed (Oxford, UK: Oxford University Press)

[45] Nelson A, Prescott S, Hannum A, McCluskey A R and Gresham I 2019 refnx/refnx v017

[46] Kass R E and Raftery A E 1995 Bayes Factors *J. Am. Stat. Assoc* **90** 773

[47] Cornish N J and Littenberg T B 2007 Tests of Bayesian model selection techniques for gravitational wave astronomy *Phys. Rev. D* **76** 083006

[48] Ensign D L and Pande V S 2010 Bayesian detection of intensity changes in single molecule and molecular dynamics trajectories *J. Phys. Chem. B* **114** 280

[49] Hollinshead C M, Harvey R D, Barlow D J, Webster J R P, Hughes A V, Weston A and Lawrence M J 2009 Effects of surface pressure on the structure of distearoylphosphatidylcholine monolayers formed at the air/water interface *Langmuir* **25** 4070

[50] McCluskey A R, Grant J, Smith A J, Rawle J L, Barlow D J, Lawrence M J, Parker S C, Edler K J and Phys J 2019 Assessing molecular simulation for the analysis of lipid monolayer reflectometry *J. Phys. Comm* **3** 075001

[51] Névot L and Croce P 1980 Caractérisation des surfaces par réflexion rasante de rayons x. application à l'étude du polissage de quelques verres silicates *Rev. Phys. Appl. (Paris)* **15** 761

[52] Kučerka N, Kiselev M A and Balgavý P 2004 Determination of bilayer thickness and lipid surface area in unilamellar dimyristoylphosphatidylcholine vesicles from small-angle neutron scattering curves: a comparison of evaluation methods *Eur. Biophys. J.* **33** 328

[53] Balgavý P, Kučerka N, Gordeliy V I and Cherezov V G 2001 Evaluation of small-angle neutron scattering curves of unilamellar phosphatidylcholine liposomes using a multishell model of bilayer neutron scattering length density *Acta. Phys. Slovaca* **51** 53

[54] Tanford C 1980 *The Hydrophobic Effect:Formation of Micelles and Biological Membranes* 2nd edn (New York, USA: Wiley)

[55] Braslau A, Deutsch M, Pershan P S, Weiss A H, Als-Nielsen J and Bohr J 1985 Surface roughness of water measured by x-ray reflectivity *Phys. Rev. Lett.* **54** 114

[56] Sinha S K, Sirota E B, Garoff S and Stanley H B 1988 X-ray and neutron scattering from rough surfaces *Phys. Rev. B* **38** 2297

[57] Tikhonov A M, Mitrinovic D M, Li M, Huang Z and Schlossman M L 2000 An x-ray reflectivity study of the water-docosane interface *J. Phys. Chem. B* **104** 6336

[58] Li Z X, Thirtle P N, Thomas R K, Penfold J, Webster J R P and Rennie A R 1998 Some improvements and extensions of the application of specular neutron reflection to the study of interfaces *Physica* **248** 171

[59] DeWolf C, Leporatti S, Kirsch C, Klinger R and Brezesinski G 1999 Phase separation in phosphatidylinositol/phosphatidylcholine mixed monolayers *Chem. Phys. Lipids* **97** 129

[60] Brezesinski G, Müller H J, Toca-Herrera J L and Krustev R 2001 X-ray diffraction and foam film investigations of PC head group interaction in water/ethanol mixtures *Chem. Phys. Lipids* **110** 183

[61] Struth B, Rieutord F, Konovalov O, Brezesinski G, Grübel G and Terech P 2001 Organization of two-dimensional phospholipid monolayers on a gel-forming substrate *Phys. Rev. Lett.* **88** 025502

[62] Sivia D S, David W I F, Knight K S and Gull S F 1993 An introduction to Bayesian model selection *Physica* **66** 234