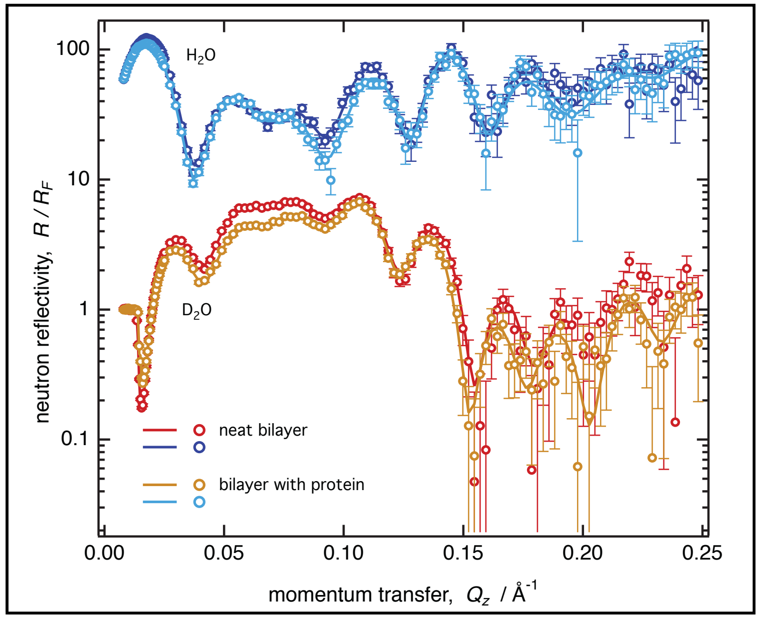
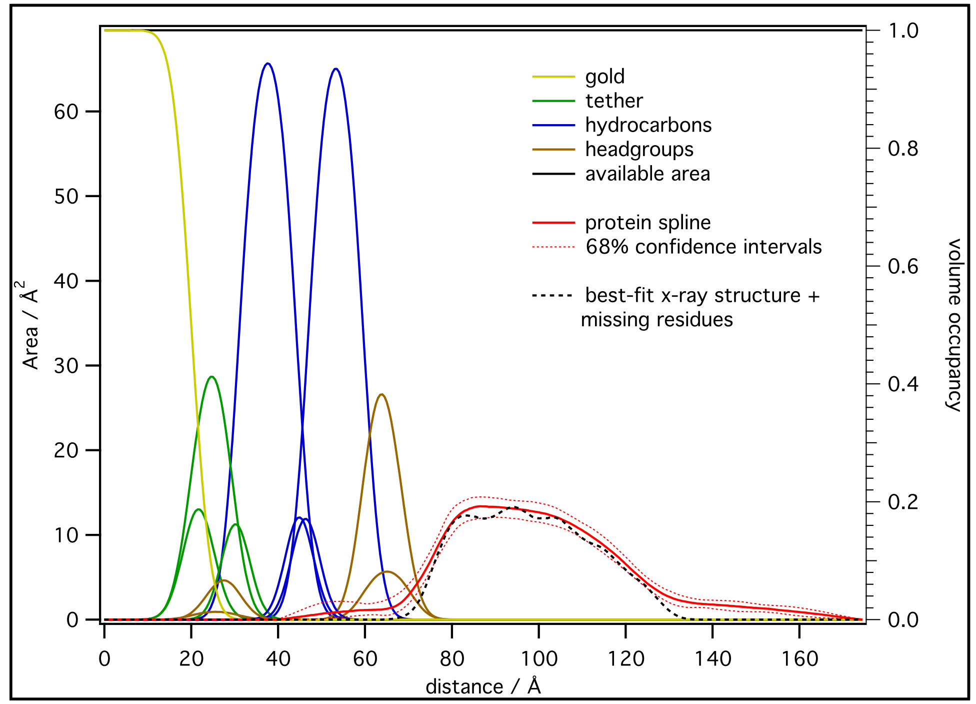
NR Data Analysis

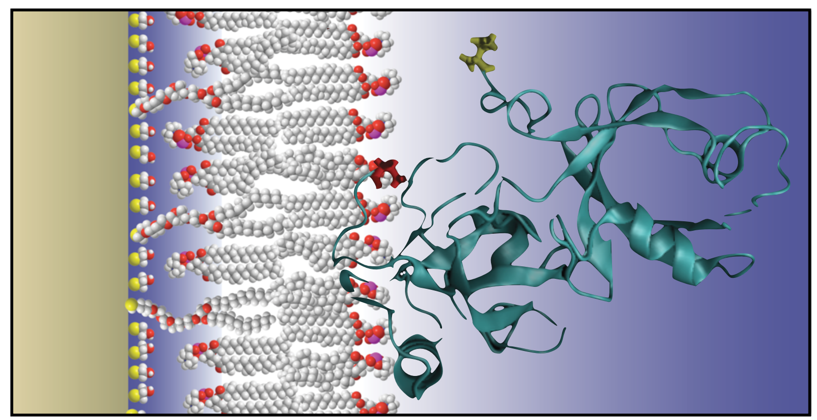
Frank Heinrich

CHRNS Center for High Resolution Neutron Scattering NIST
NIST National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce
Carnegie Mellon

Data Analysis







Goal

1. Retrieve Structural Information from collected specular reflectivity curves (simultaneously).
2. Off-specular data requires separate analysis.
3. Not unique due to loss of phase information during measurement process.
4. Inversion of Reflectivity possible only in special cases.
5. Artificial Intelligence Methods are relatively new.
6. Modeling is most common.

Result

1. 1D nSLD or structural profiles.
2. Uncertainties
3. 3D structures require complimentary information

Modeling of Specular NR

Experimental data

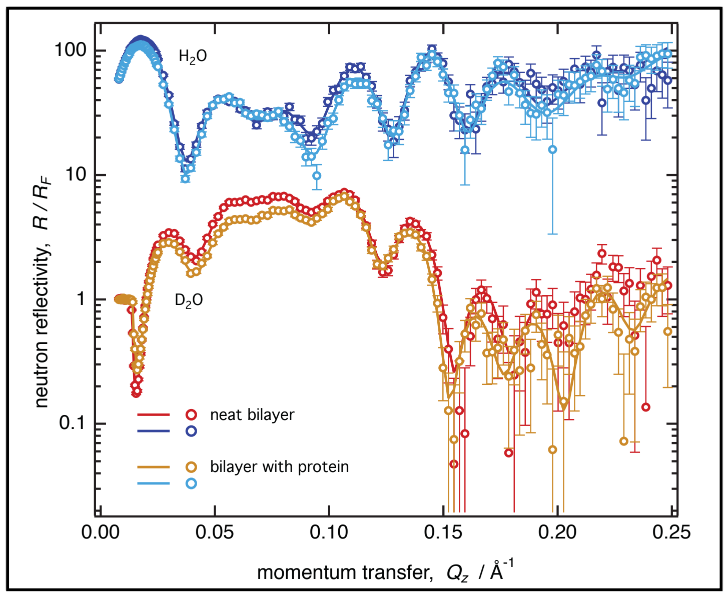
1. Sets of related data that will be simultaneously analyzed (contrasts, conditions, polarizations).
2. Experimental Design based on experience or computation.

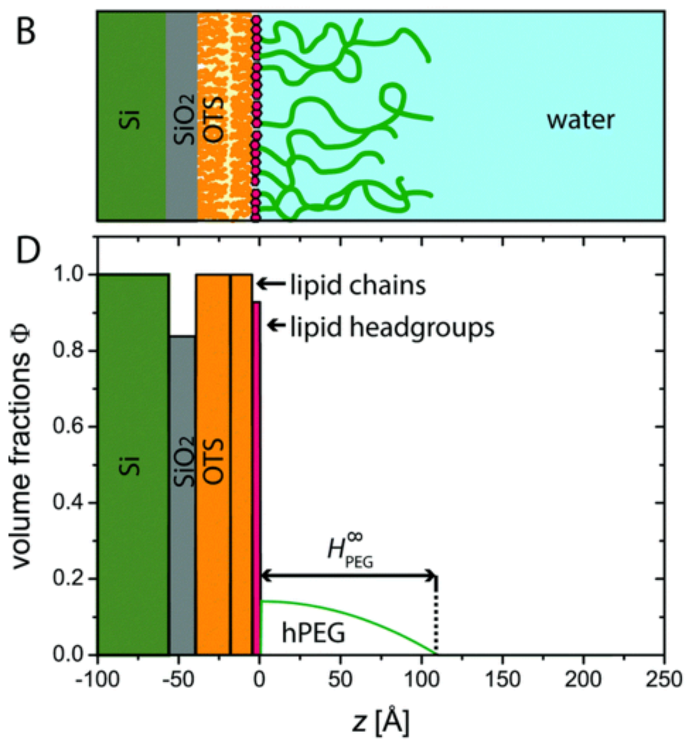
Models

1. Model: Parameterized representation of real space.
2. Choice of model: Slab model, functional layer model, composition-space model, integrative models.
3. Model parameter boundaries (fit limits, prior information)

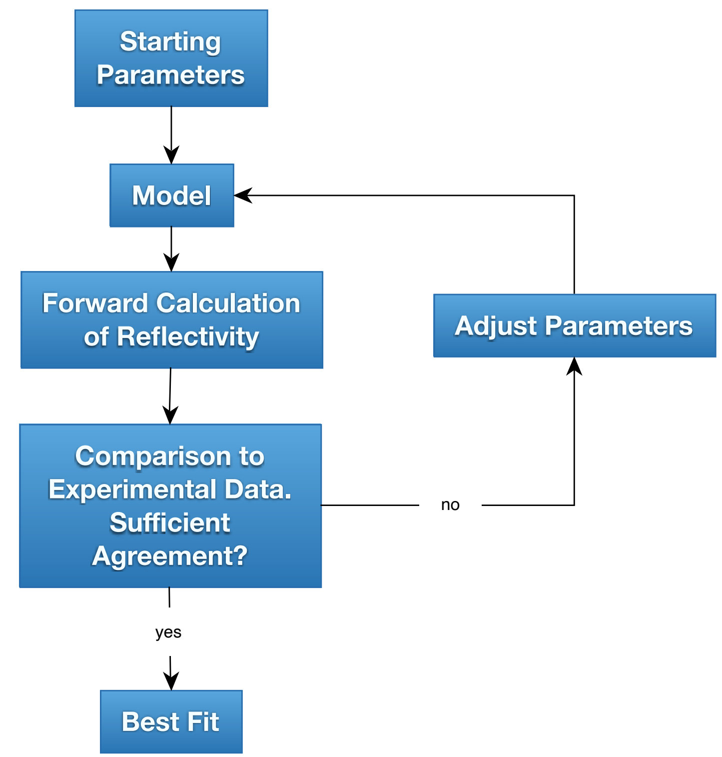
Optimizer

1. Yields best-fit or distributions of model parameters with uncertainties.
2. Global or local optimizers.
3. Software: Webfit (B. Maranville), Refl1D (NCNR Paul Kienzle), Motofit, GenX.





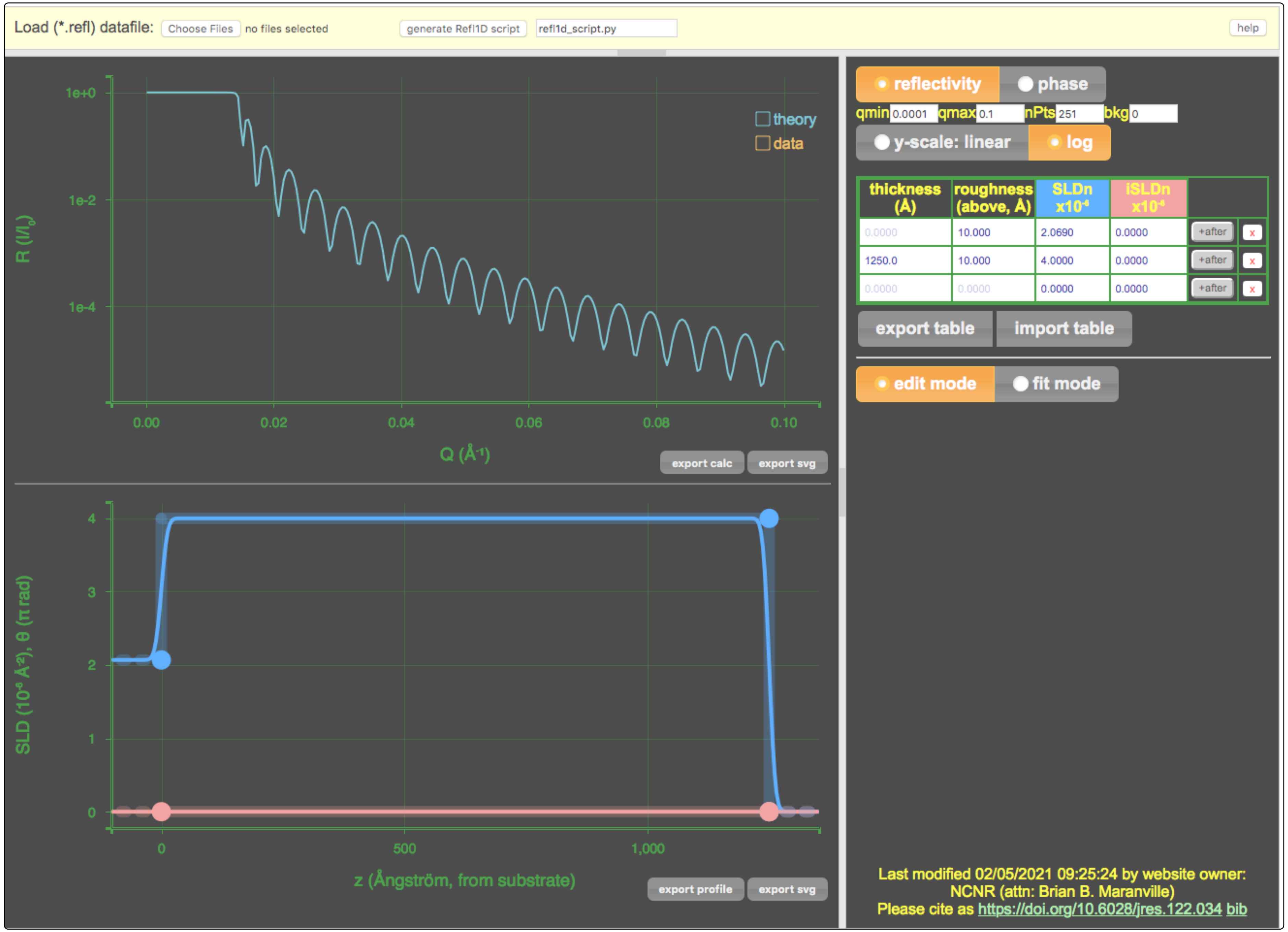
Rodriguez-Loureiro, I. et al. Soft Matter **13**, 5767-5777 (2017).



Slab Model Web Fitting

1. Reflectivity can be imported directly from Reductus.
2. Simple slab model and local optimizer allow for quick assessment of the data during the experiment.
3. Allows to export fitting script for Refl1D software.

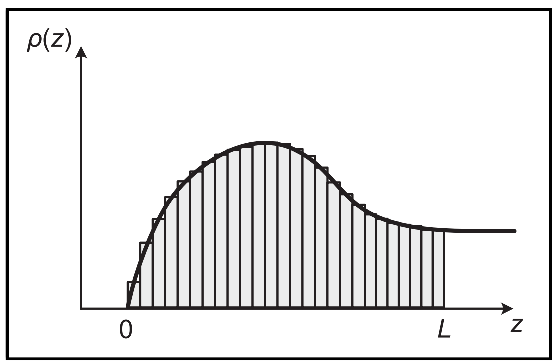
ncnr.nist.gov/instruments/magik/calculators/reflectivity-calculator.html (Dr. Brian Maranville)



Models

Slab Models

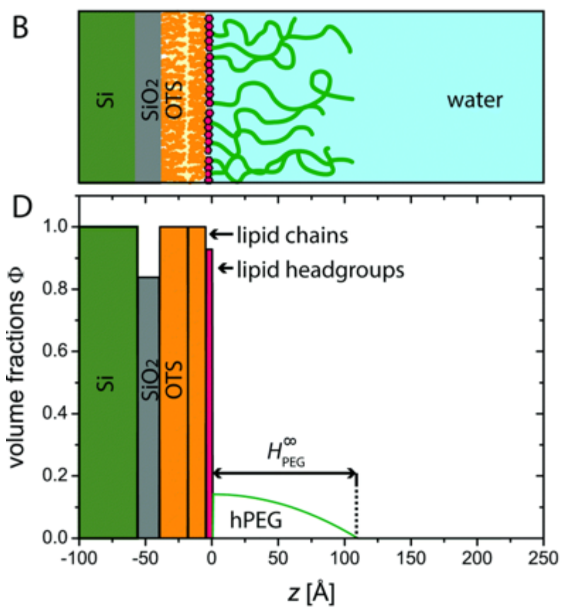
1. Slabs (layers, boxes) of thickness d and constant SLD, ρ.
2. Input for matrix or Parratt's formalism to calculate reflectivity.
3. Constraints between slabs possible
4. Basis for more complex models – micro-slicing.



Ankner, J. F. & Majkrzak, C. F. Subsurface profile refinement for neutron specular reflectivity. SPIE Proceedings **1738**, 260-269 (1992).

Functional Layer Models

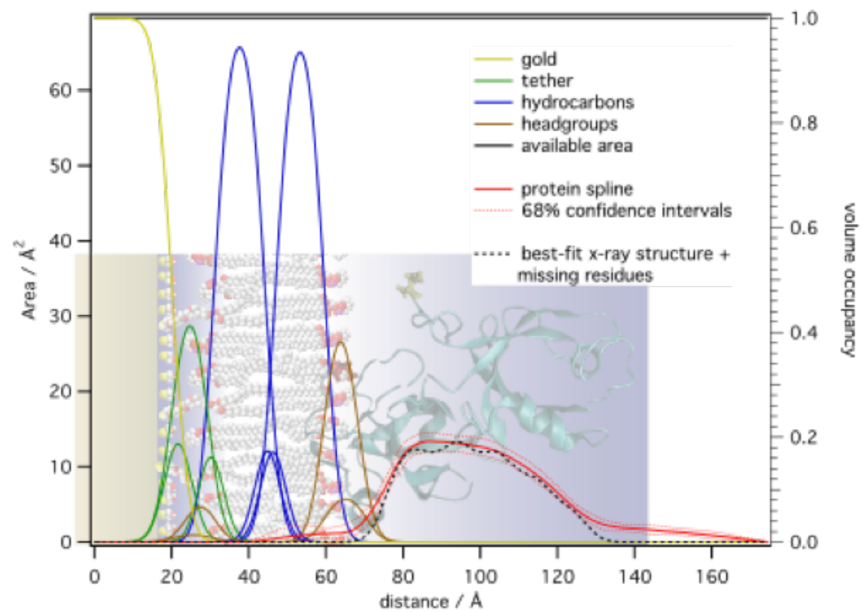
1. A canvas of micro-slices to draw models in containing any functional form, including free-form.
2. Modeling of SLD or volume profiles.
3. Constraints on components. Overlapping components.



Rodriguez-Loureiro, I. et al. Soft Matter **13**, 5767-5777 (2017).

Composition-space models

1. Describe the spatial distribution of molecular groups rather than profiles of scattering length densities.



Parameter Optimizers

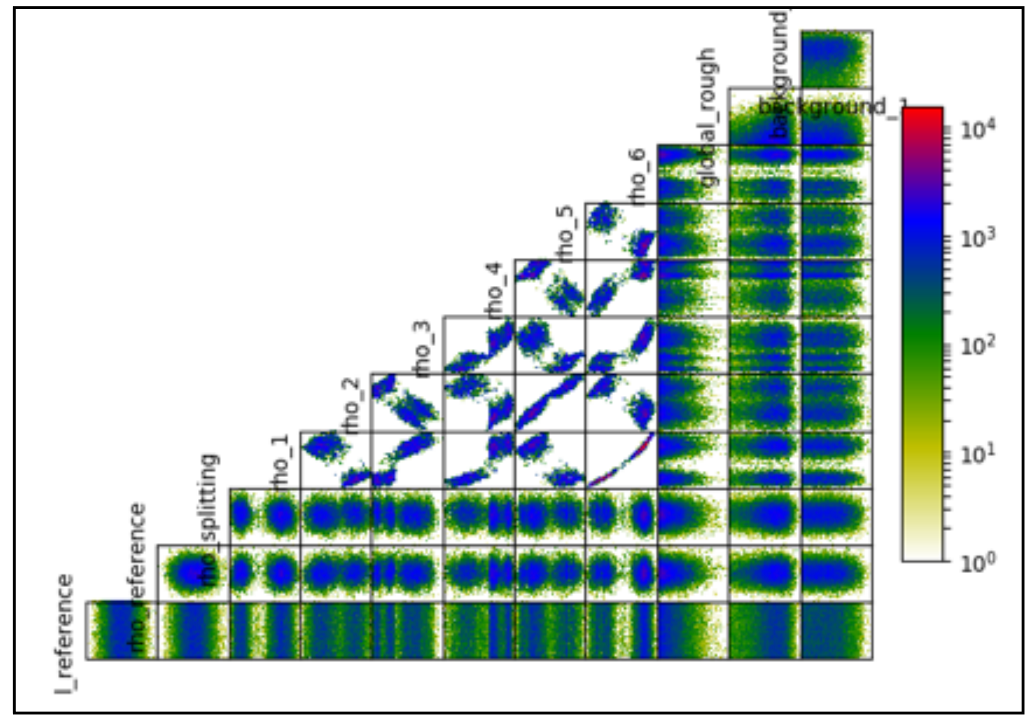
Optimizer

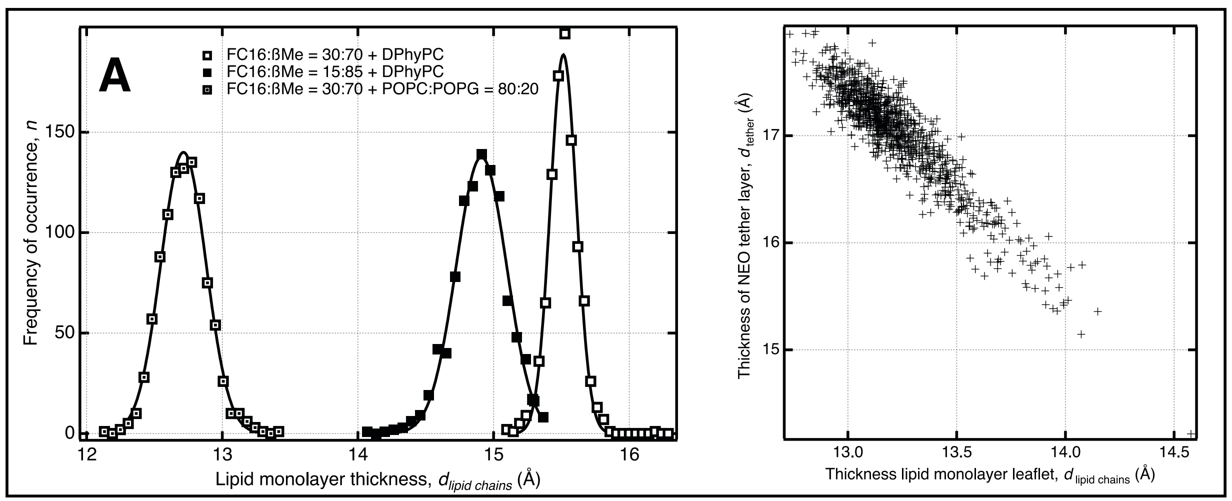
1. Local optimizers (e.g., steepest descent) find local minima close to starting parameter values. Uncertainties are also from this local environment.
2. Global optimizers are able to find global fit minima in many cases.
3. Often yield a probability distribution (posterior PDF) for confidence limits, parameter correlations, and model optimization.

Refl1D optimizers via Bumps

1. Levenberg-Marquardt
2. Nelder-Mead Simplex
3. DREAM
4. Differential Evolution
5. Quasi-Newton BFGS
6. Random Lines (experimental)
7. Particle Swarm (experimental)
8. Parallel Tempering (experimental)

2D-projection of a posterior PDF obtained with a MCMC-based global optimizer showing a non-unique solution to the modeling problem.



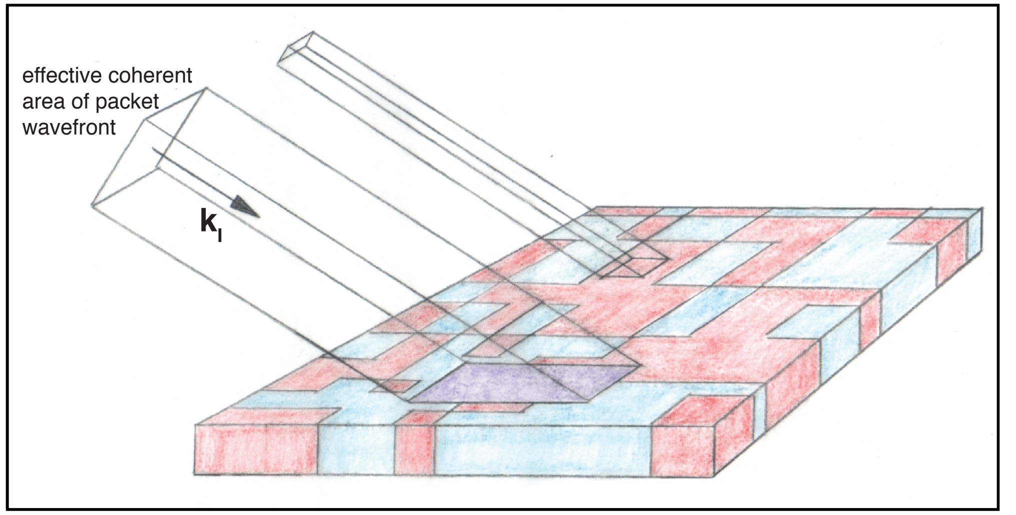


Parameter distributions and correlations obtained with a MC simulation technique.

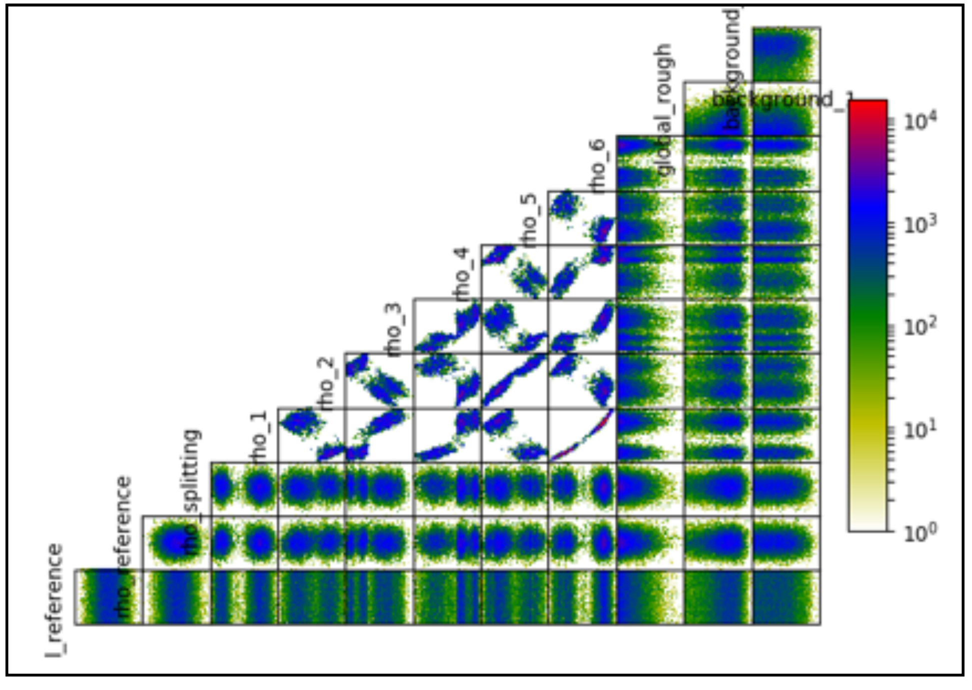
Problems

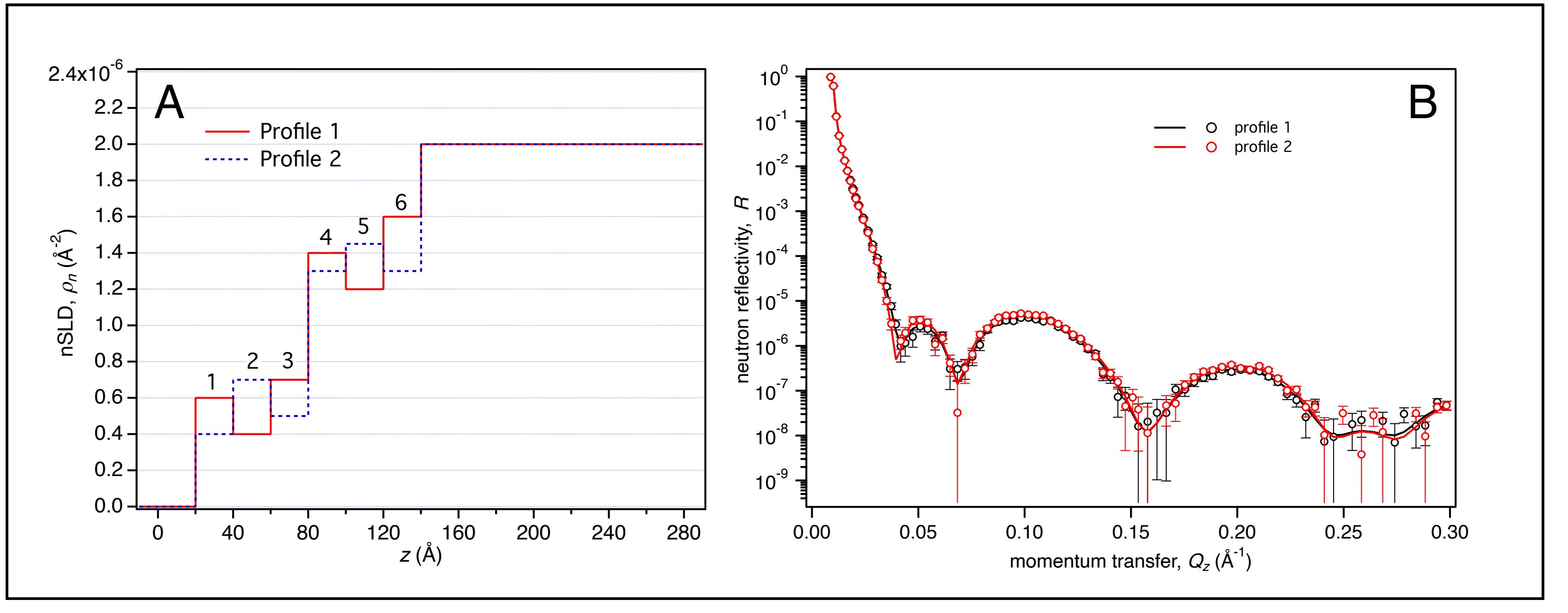
Problems

1. Non-unique solutions (contrasts, complimentary data, constraints).
2. In-plane inhomogeneities on a length scale at or larger than the coherence length of the neutron (avoid, no cure).
3. Unsuccessful optimizer.
4. Over or under-parameterization



in-plane inhomogeneities



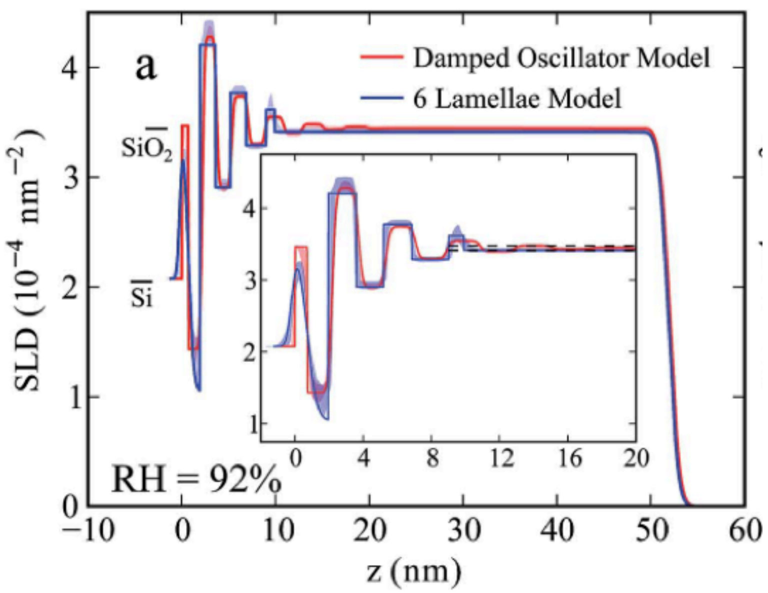


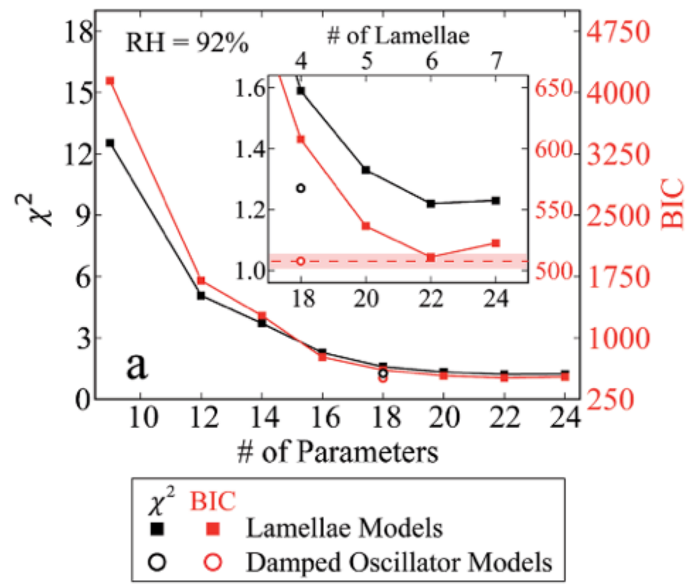
non-unique solutions / overparameterization

Model Selection

Model selection

1. Empirical Model optimization from inspecting parameter uncertainties and correlations.
2. Model type and detail optimized using the BIC.
3. Optimizing models within a broader experimental optimization using information theory.



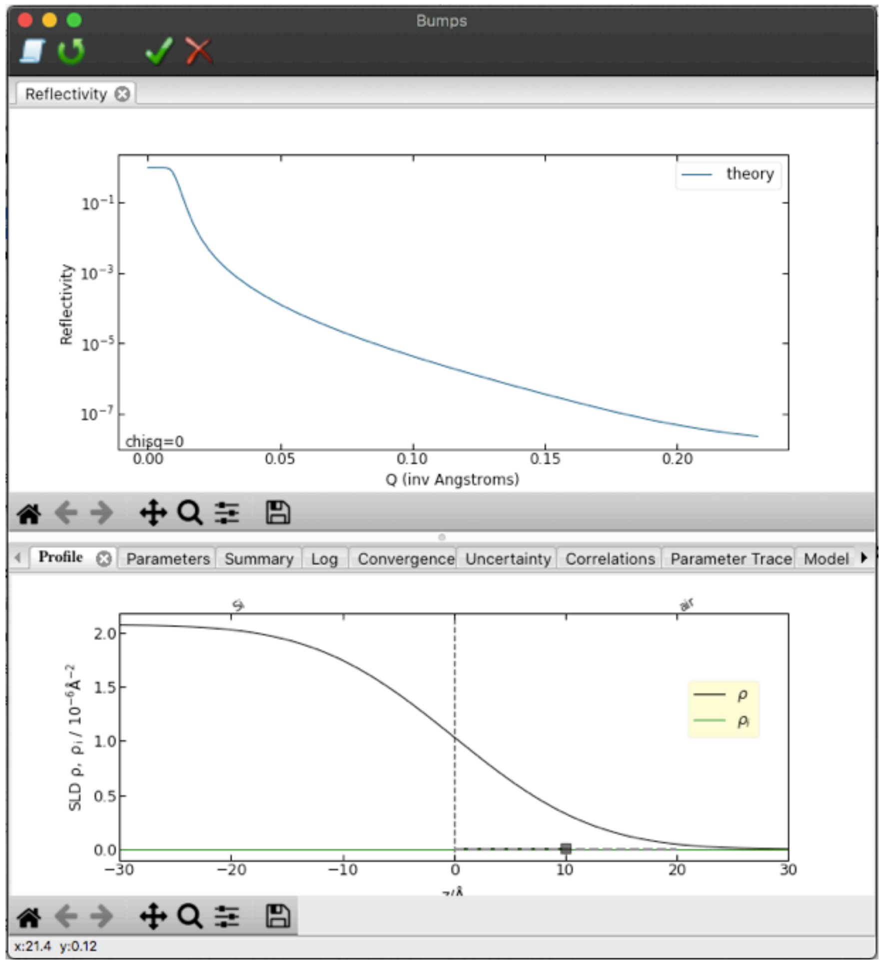


DeCaluwe, S. C., Kienzle, P. A., Bhargava, P., Baker, A. M. & Dura, J. A. Soft Matter **10**, 5763-5776 (2014).

Refl1D

Refl1D

1. Python, script-based.
2. Flexible model builder
3. Uses bumps as optimizer (library of local and global optimizers).
4. Parallelized (HPC deployment, GPU support)



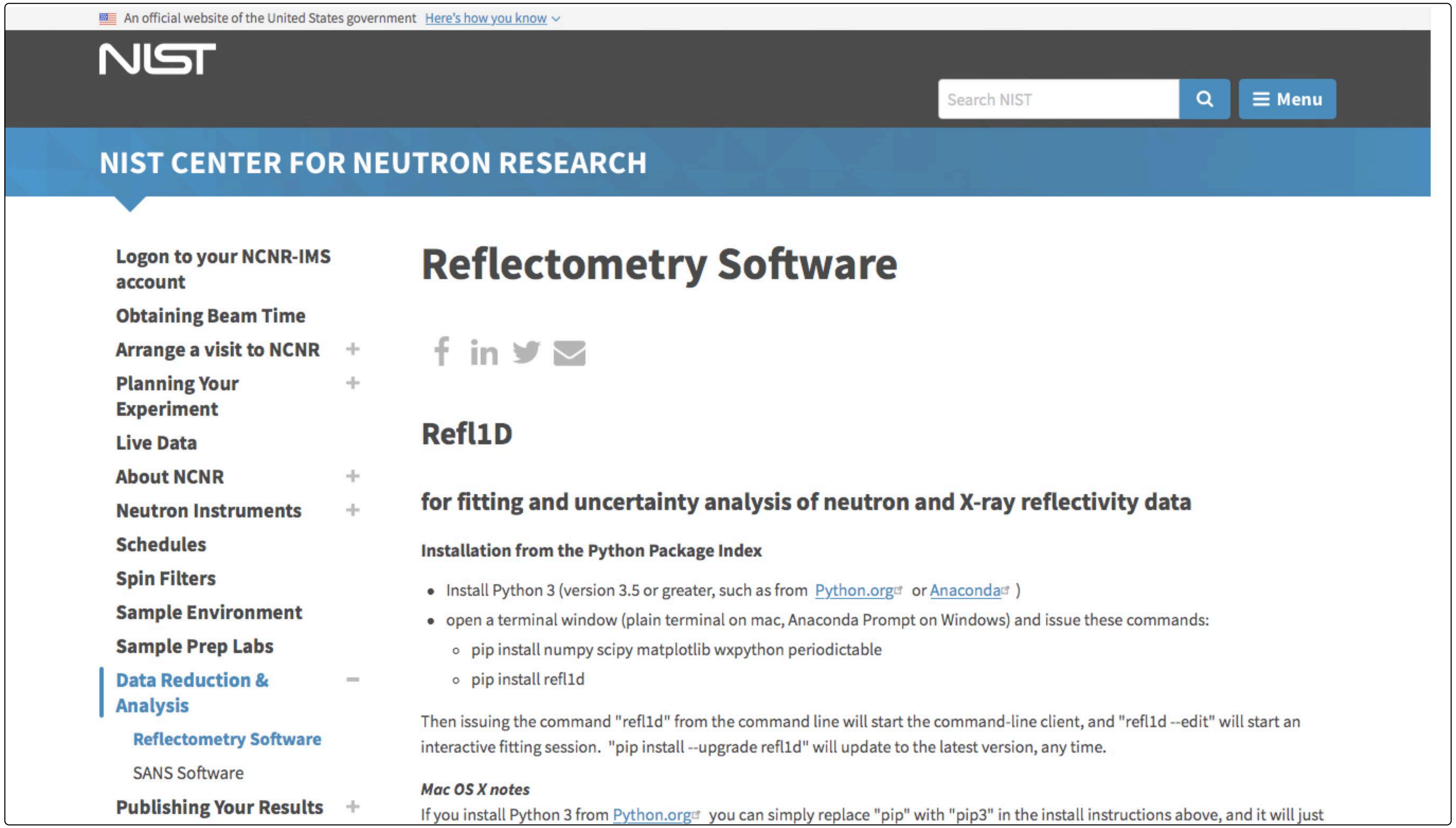
- Magnetism example
- User's Guide
  - Using Refl1D
  - Parameters
  - Data Representation
  - Materials
  - Sample Representation
  - Experiment
  - Fitting
- Reference
  - abeles - Pure python reflectivity calculator
  - anstodata - Reader for ANSTO data format
  - cheby - Freeform - Chebyshev model
  - dist - Non-uniform samples
  - errors - Plot sample profile uncertainty
  - experiment - Reflectivity fitness function
  - fitplugin - Bumps plugin definition for reflectivity models
  - flayer - Functional layers
  - freeform - Freeform - Parametric B-Spline
  - fresnel - Pure python Fresnel reflectivity calculator
  - garefl - Adaptor for garefl models
  - instrument - Reflectivity instrument definition
  - magnetism - Magnetic Models
  - material - Material
  - materialdb - Materials Database
  - model - Reflectivity Models
  - mono - Freeform - Monotonic Spline
  - names - Public API
  - ncnrdata - NCNR Data
  - polymer - Polymer models
  - probe - Instrument probe
  - profile - Model profile
  - reflectivity - Reflectivity
  - reflmodule - Low level reflectivity calculations
  - resolution - Resolution
  - snsdata - SNS Data
  - staj - Staj File
  - stajconvert - Staj File Converter
  - stitch - Overlapping reflectivity curve stitching
  - support - Environment support
  - util - Miscellaneous functions

v: latest

https://refl1d.readthedocs.io/en/latest/ (Dr. Paul Kienzle)

Tools

1. Refl1D (installation and manual)
2. Online SLD calculator
3. Web Reflectivity Calculators.



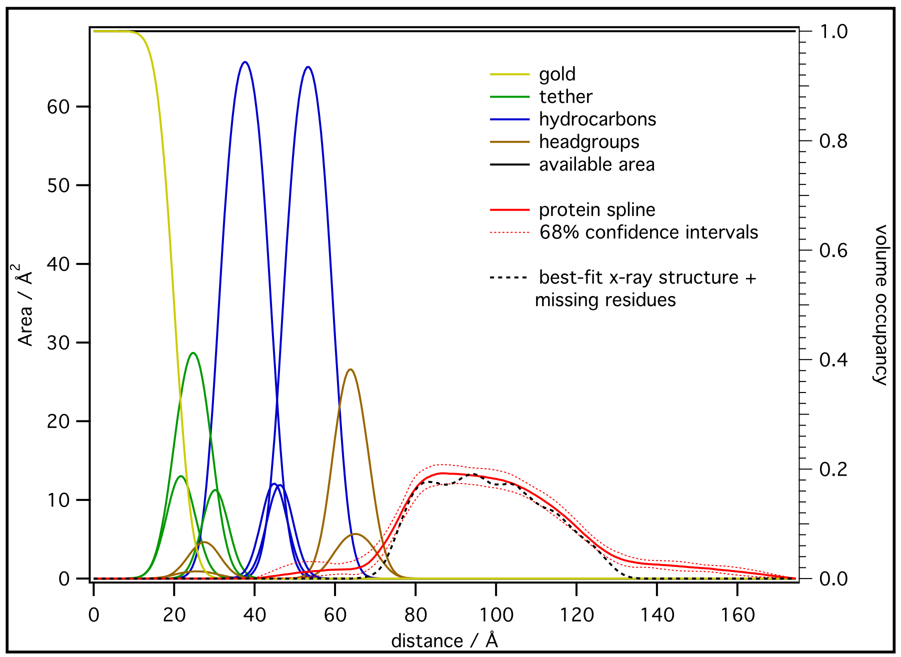
https://www.nist.gov/ncnr/data-reduction-analysis/reflectometry-software (Paul Kienzle, Brian Maranville)

Integrative Modeling

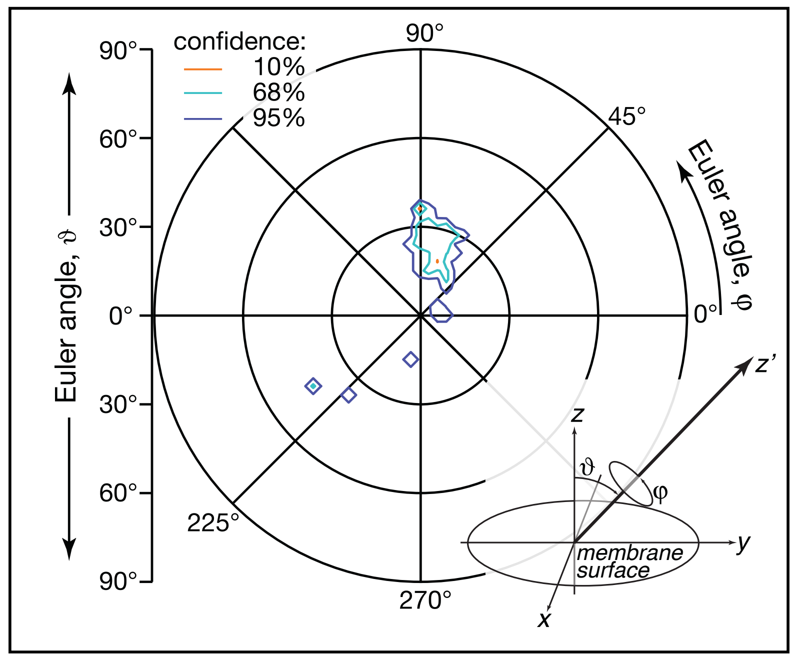
Integrative Modeling

1. Integrative modeling in biological NR retrieves the missing 2 dimensions and allow for structural insights beyond the resolution limit of the technique.
2. This is achieved by integrating external, complimentary information such as high-resolution structures of molecules in the NR data modeling.

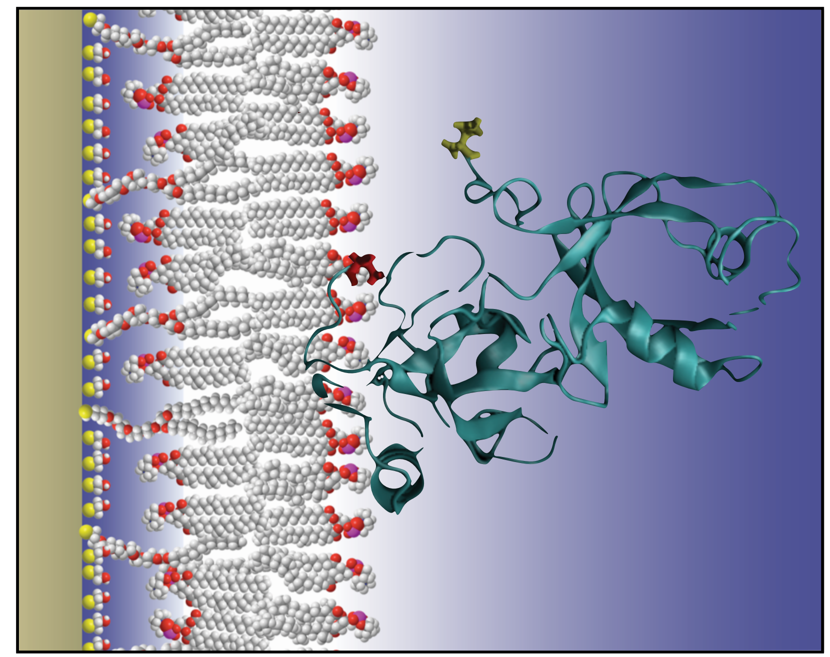
Heinrich, F. & Lösche, M. Zooming in on disordered systems: Neutron reflection studies of proteins associated with fluid membranes. Biochimica et Biophysica Acta (BBA)–Biomembranes **1838**, 2341–2349 (2014).



Composition-space model using an NMR structure



Orientation probability of the NMR structure at the membrane restraint by the NR data

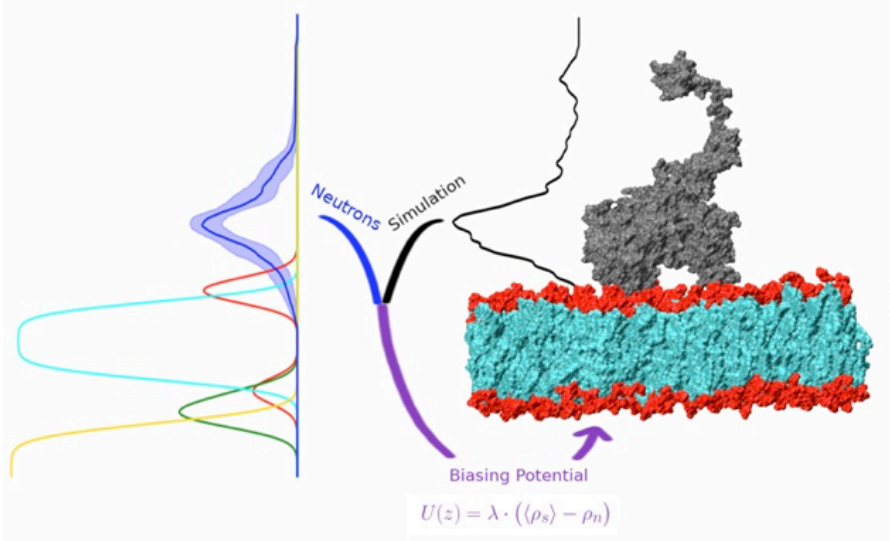


Most likely orientation of the HIV I - matrix protein determined from this analysis

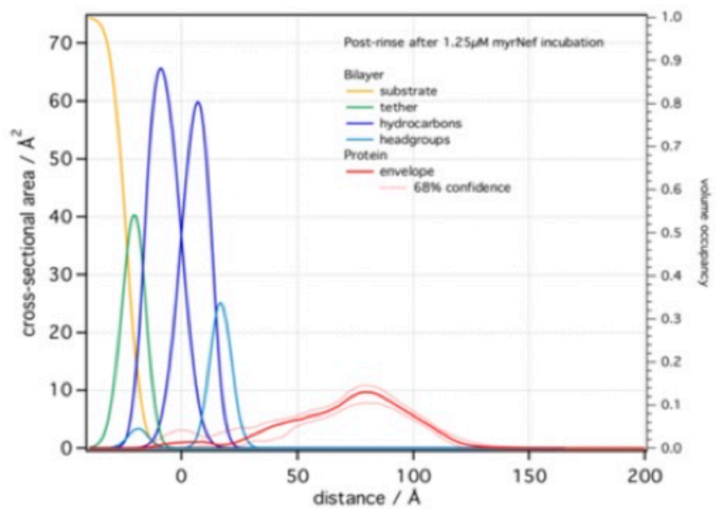
Restrained MD Simulations

Restrained MD simulations

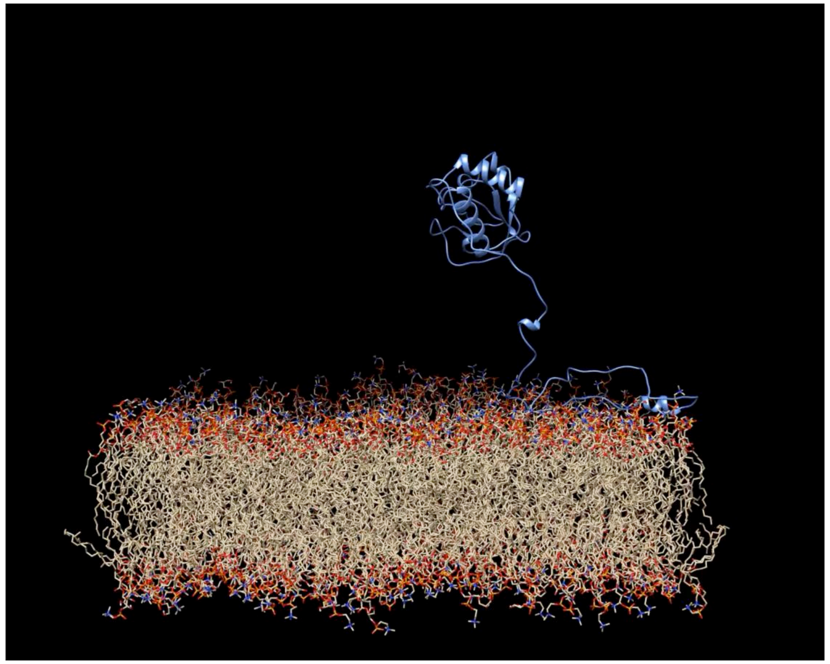
1. Free MD simulations and NR data only occasionally agree with each other.
2. Experimental data can put restraints on MD simulation using a biasing potential, leading to a reduction in conformational space sampled by the simulation and typically faster equilibration.
3. Often, the restraint MD simulation poses a compromise between independent predictions of the two techniques.

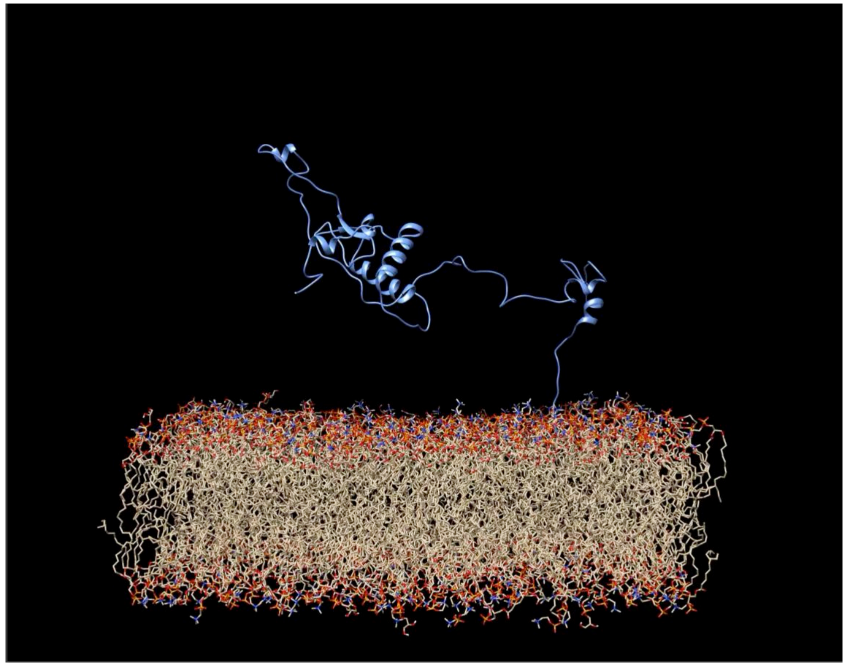


obtaining a potential from a free-form spline fit to a protein (in this Figure: HIV I - matrix)



Spline profile (potential) for HIV I - Nef used in the simulations shown below.

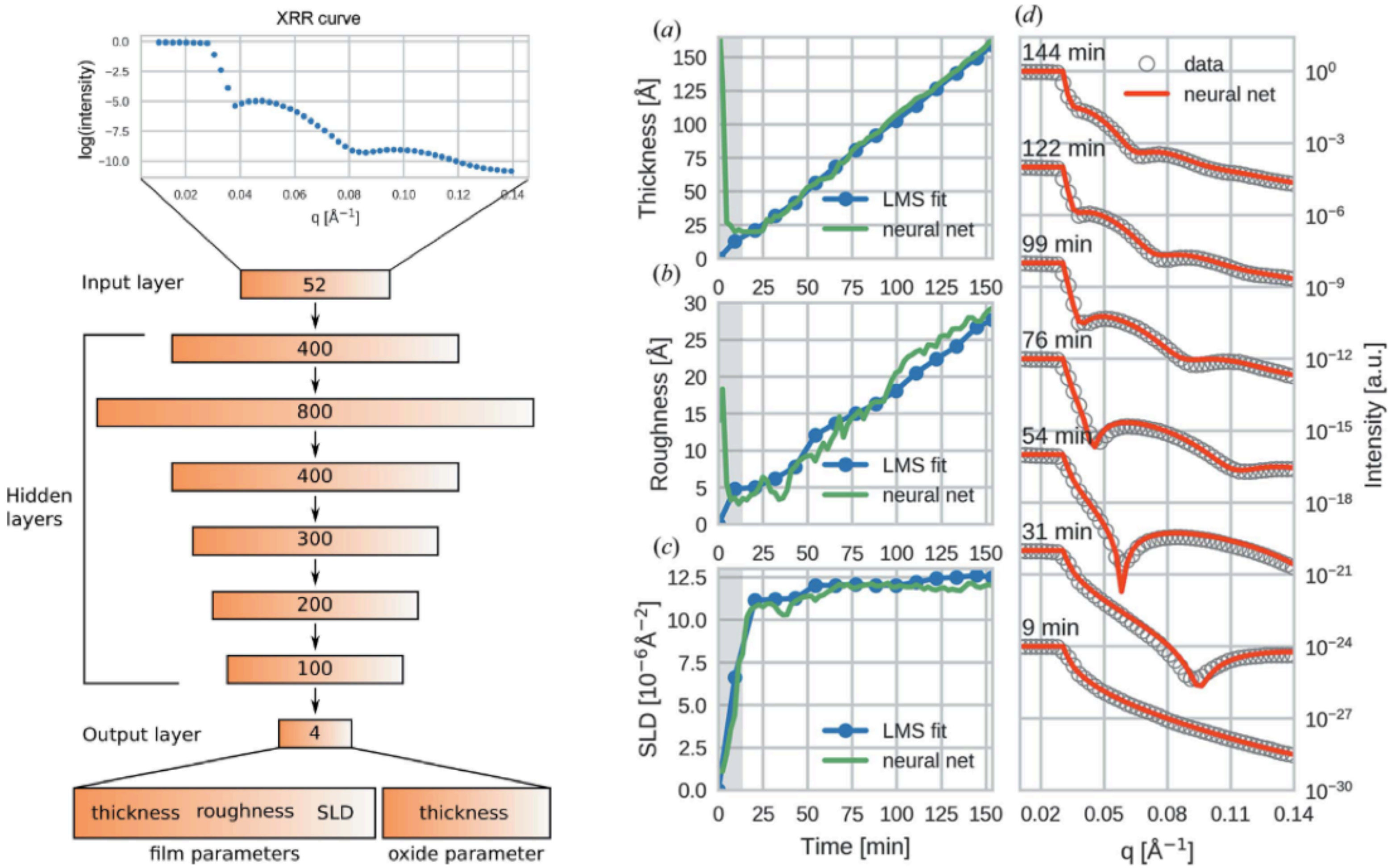




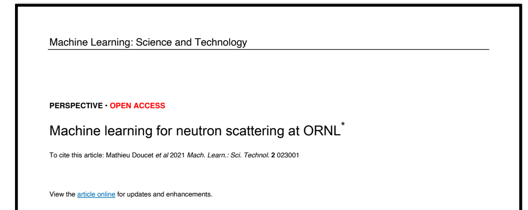
restrained (left) vs. free (right) MD simulation of HIV I - Nef

Treece, B. W., Heinrich, F., Ramanathan, A. & Lösche, M. Steering Molecular Dynamics Simulations of Membrane-Associated Proteins with Neutron Reflection Results. J Chem Theory Comput **16**, 3408-3419 (2020).

Deep Learning



Greco, A. et al. Fast fitting of reflectivity data of growing thin films using neural networks. J Appl Crystallogr **52**, 1342–1347 (2019).



Machine Learning: Science and Technology

PERSPECTIVE · OPEN ACCESS

Machine learning for neutron scattering at ORNL*

To cite this article: Mathieu Doucet *et al* 2021 *Mach. Learn.: Sci. Technol.* **2** 023001

View the article online for updates and enhancements.

Doucet, M. et al. Machine learning for neutron scattering at ORNL. Mach Learn Sci Technology **2**, 023001 (2021).

arXiv:2010.07634v3 [physics.comp-ph] 24 Jan 2021

Towards Reflectivity profile inversion through Artificial Neural Networks

Juan Manuel Carmona Loaiza[1,*] and Zamaan Raza[1]

[1]Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ). Lichtenbergstraße 1, 85748 Garching, Germany.

E-mail: j.carmona.loaiza@fz-juelich.de

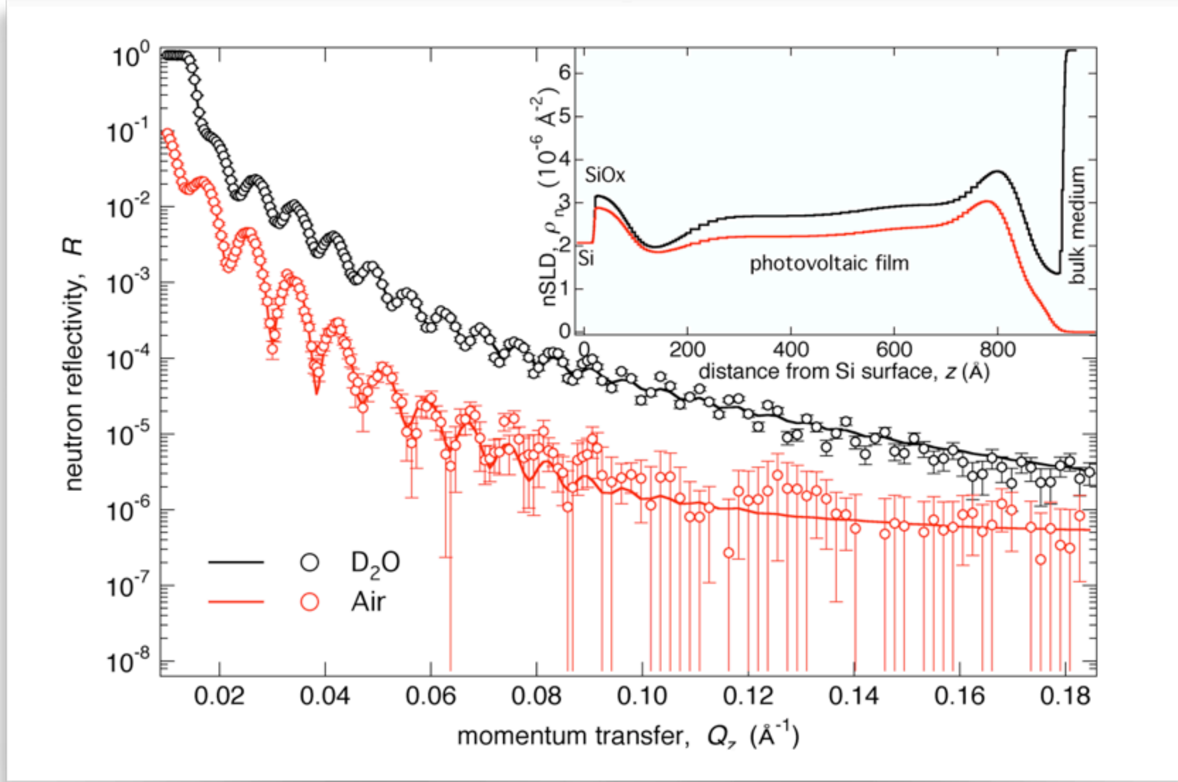
22 December 2020

Abstract. The goal of Specular Neutron and X-ray Reflectometry is to infer materials Scattering Length Density (SLD) profiles from experimental reflectivity curves. This paper focuses on investigating an original approach to the ill-posed non-invertible problem which involves the use of Artificial Neural Networks (ANN). In particular, the numerical experiments described here deal with large data sets of simulated reflectivity curves and SLD profiles, and aim to assess the applicability of Data Science and Machine Learning technology to the analysis of data generated at neutron scattering large scale facilities. It is demonstrated that, under certain circumstances, properly trained Deep Neural Networks are capable of correctly recovering plausible SLD profiles when presented with never-seen-before simulated reflectivity curves. When the necessary conditions are met, a proper implementation of the described approach would offer two main advantages over traditional fitting methods when dealing with real experiments, namely, 1. sample physical models are described under a new paradigm: detailed layer-by-layer descriptions (SLDs, thicknesses, roughnesses) are replaced by parameter free curves $\rho(z)$, allowing a-priori assumptions to be fed in terms of the sample family to which a given sample belongs (e.g. "thin film", "lamellar structure", etc.) 2. the time-to-solution is shrunk by orders of magnitude, enabling faster batch analyses for large datasets.

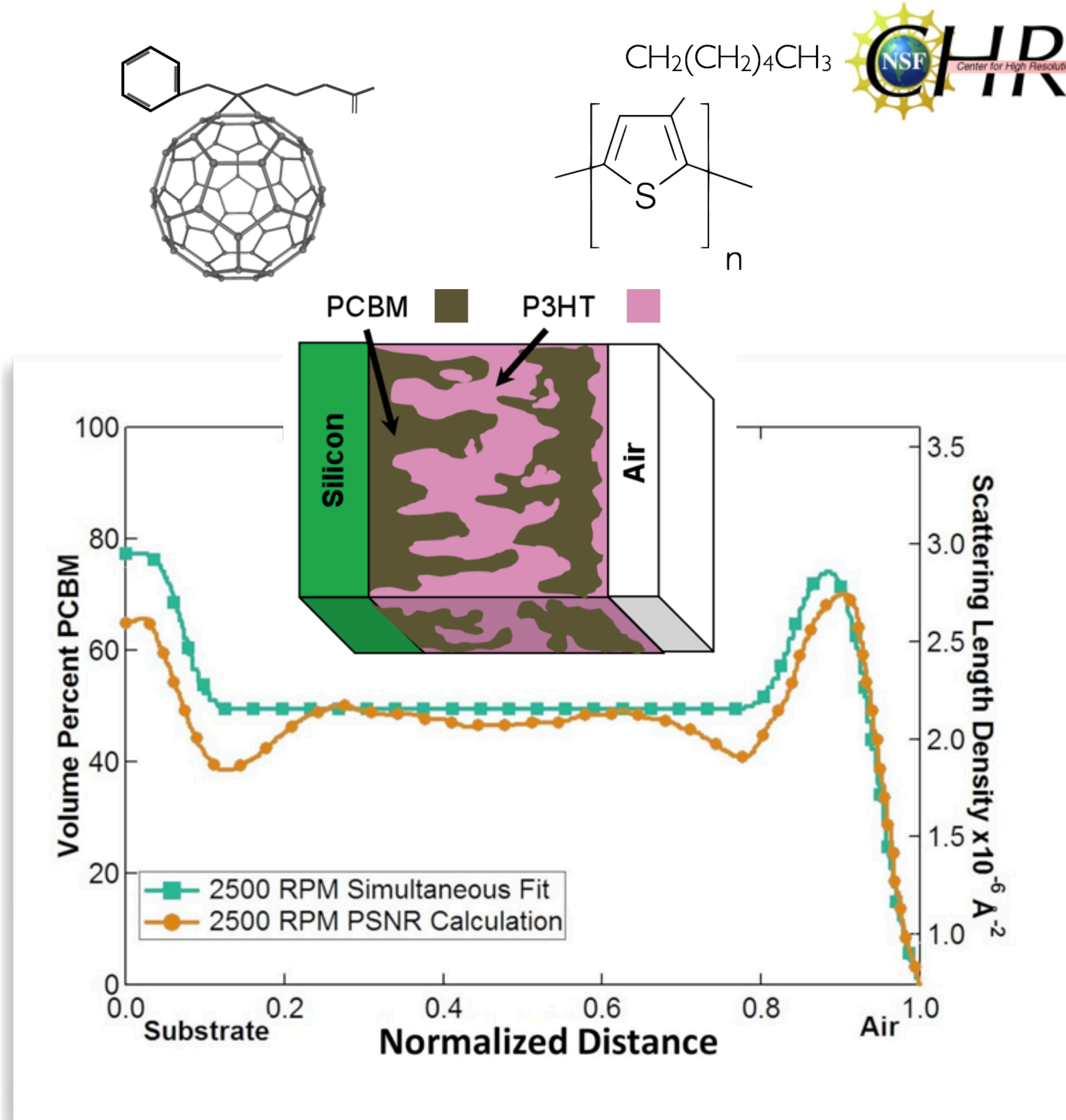
Keywords: inverse problems, neutron scattering, x-ray scattering, reflectometry, reflectivity, data science, data analysis, algorithms, artificial intelligence, machine learning, neural networks
Submitted to: *Mach. Learn. Sci. Technol.*

Carmona-Loaiza, J. M. & Raza, Z. Towards Reflectivity profile inversion through Artificial Neural Networks. Arxiv (2020).

Phase-sensitive NR



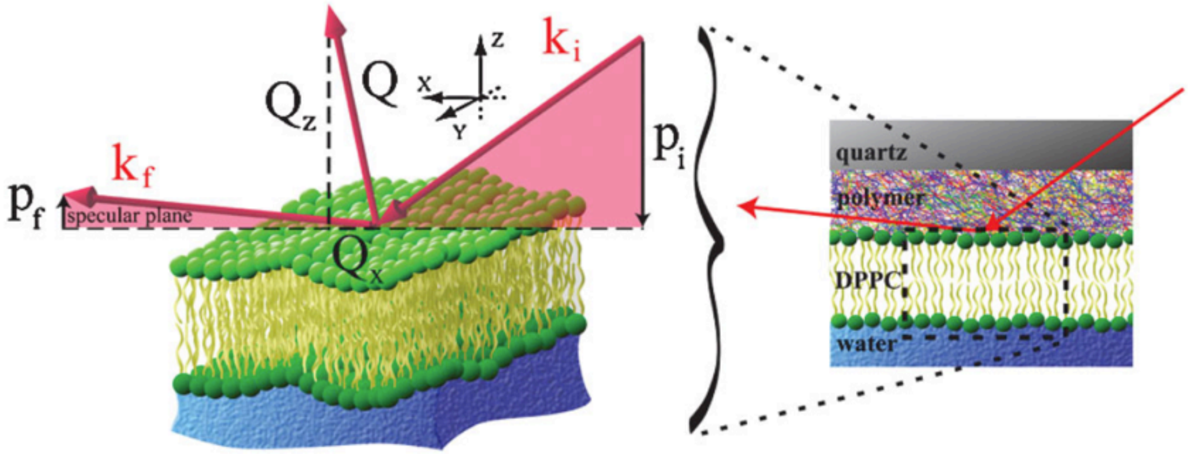
1. Inversion of Reflectivity possible under certain circumstances.
2. Requires multiple measurements of the same sample with different reference structures.
3. Model fitting has been shown to be equivalent. (Majkrzak / Berk)

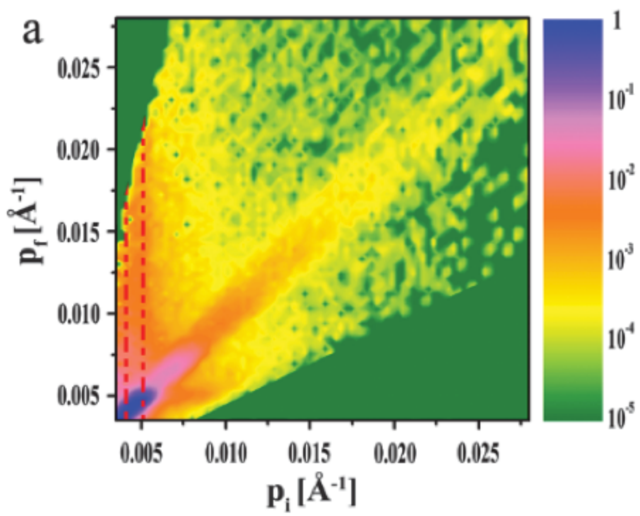


Inhomogeneous distribution of acceptors and donors in organic photovoltaics.

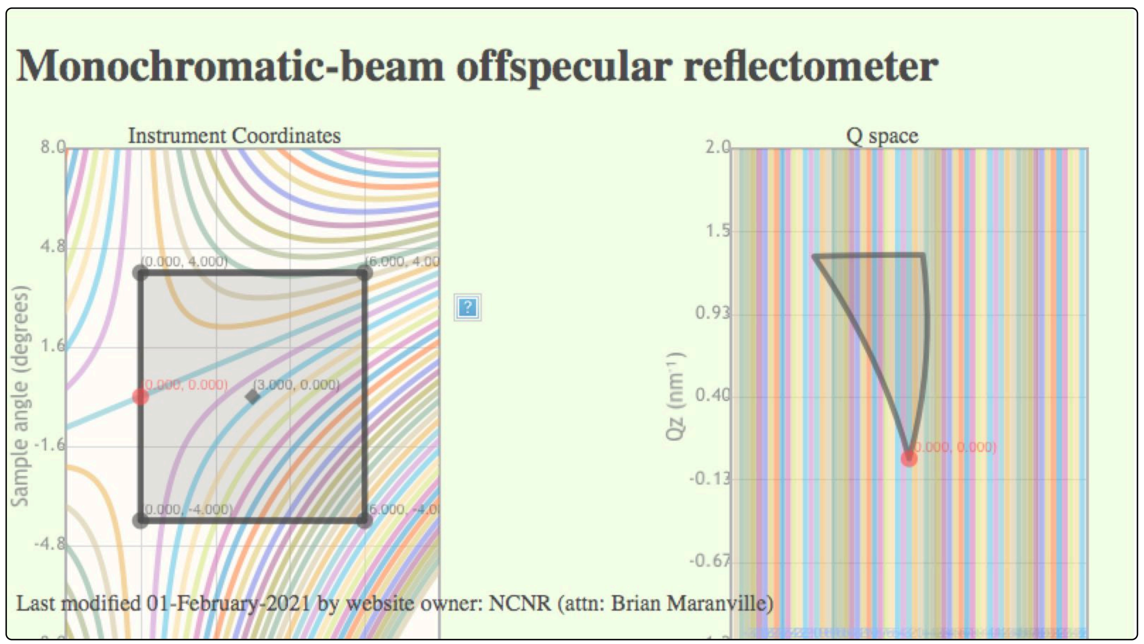
Kirby, B.J. et al. Current Opinion in Colloid & Interface Science, 17:.44 (2012)

Off-specular NR





Jablin, M. S. et al. In-plane correlations in a polymer-supported lipid membrane measured by off-specular neutron scattering. Physical Review Letters 106, 138101 (2011).



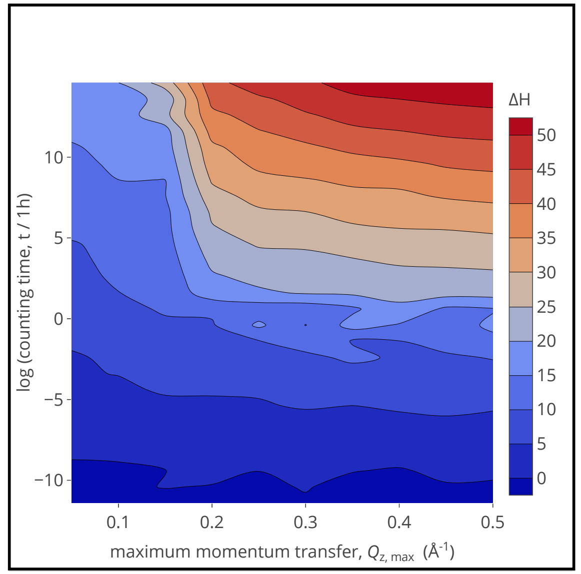
https://www.ncnr.nist.gov/instruments/magik/calculators/offspec_planner.html

1. (Periodic) in-plane structures cause off-specular scattering.
2. Analysis requires a different theory than for specular reflectivity
3. Brian Maranville.

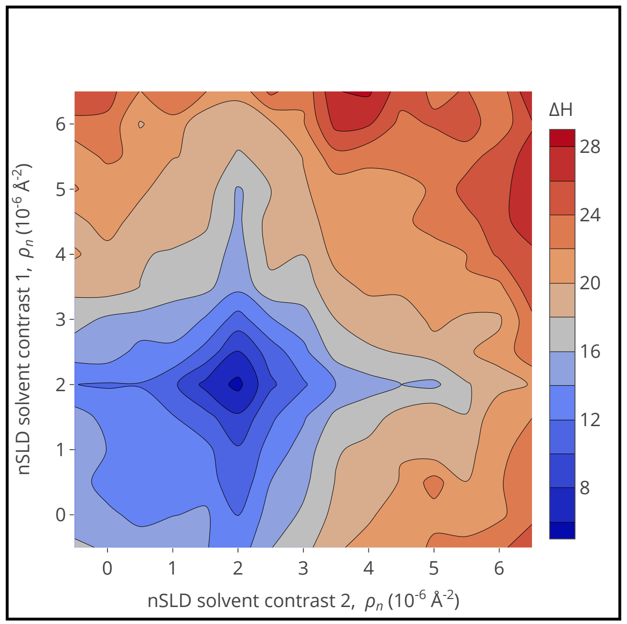
Experimental Design

Experimental Design in Neutron Reflectometry

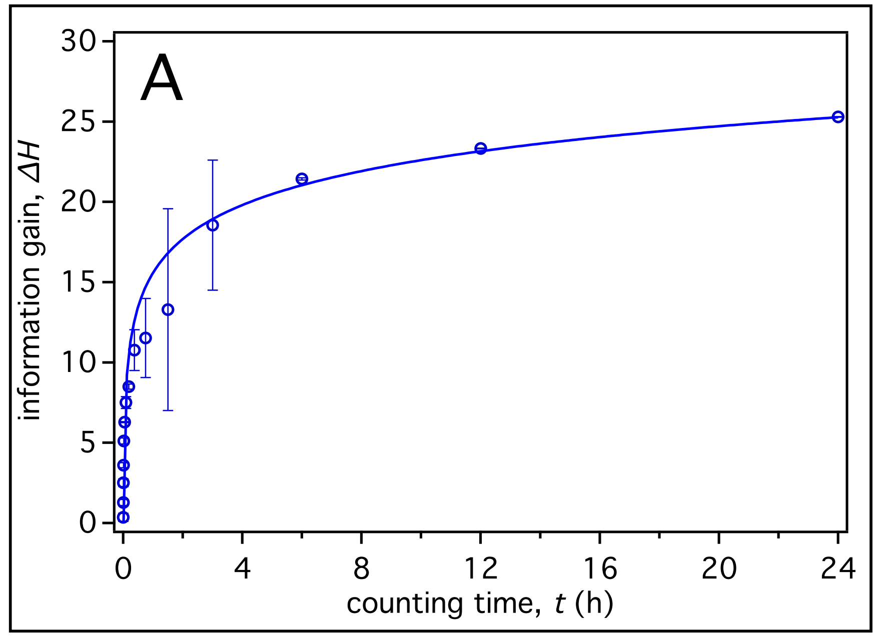
1. Optimal measurement and data modeling strategies can be found.
2. Example variables: Counting time, Q-range, SLD engineering, number and type of subsequent measurements in a series (i.e., of bulk solvent contrasts), type of substrate …



Information gain as a function of counting time and maximum momentum transfer.



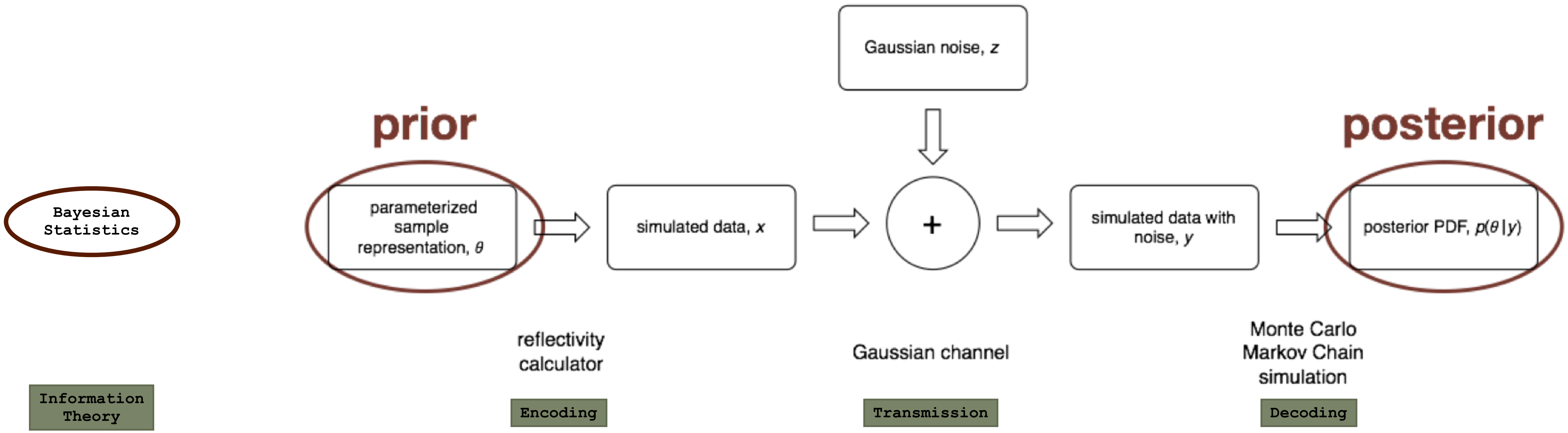
Function of information gain the deuteration of a lipid bilayer and the surrounding aqueous phase.



Information gain as a function of counting time for a test structure.

Treece, B. W. et al. Optimization of reflectometry experiments using information theory. Journal of Applied Crystallography **52**, 47–59 (2019).

Experimental Design



Experimental Design in Neutron Reflectometry

1. Optimizing the experimental design is equivalent to maximizing the information content of the measurement.
2. Calculating the information content of NR requires to think about the entire measurement process in terms of information processing and transmission, while relying on Bayesian statistics.

Treece, B. W. et al. Optimization of reflectometry experiments using information theory. Journal of Applied Crystallography **52**, 47-59 (2019).

The information gain from a measurement can be quantified as the mutual information between the prior and posterior parameter distributions, which equals their difference in entropy:

$$I(\theta;Y) = H(\theta) - H(\theta|Y)$$

The entropy is being calculated as a Shannon Entropy:

$$H(X) = -\sum_{x \in X} p(x) \log p(x)$$

Automated Measurements

The future of NR

1. A combination of information theory, Bayesian statistics, and machine learning can be used to implement an autonomous measurement system that self-optimizes for maximum information gain.