Autonomous screening of complex phase spaces using Bayesian optimization for SAXS measurements

Khaled Younes[a,*], Michael Poli[b], Priyanka Muhunthan[a], Ivan Rajkovic[c], Stefano Ermon[b], Thomas M Weiss[c] and Matthias Ihme[a]

[a]Department of Mechanical Engineering, Stanford University, Stanford, CA 94305, United States
[b]Department of Computer Science, Stanford University, Stanford, CA 94305, United States
[c]SLAC National Accelerator Laboratory, Stanford Synchrotron Radiation Lightsource, Menlo Park, CA 94025, United States

ARTICLE INFO

Keywords:

Bayesian optimization

SAXS

online screening

phase space sampling

supercritical fluids

ABSTRACT

The advent of modern, ultrafast X-ray experiments has enabled scientists to probe physical phenomena at an ever smaller scale. However, this has come at a cost of excessive data generation, to the point where current storage and hardware capabilities are routinely surpassed. As such, handling the data efficiently and selectively storing only the information of most relevance is crucial. In this paper, we propose to use Bayesian optimization as a method to alleviate this problem. We apply the method to locate global features in Small Angle X-ray Scattering spectra obtained from conducting experiments with supercritical CO_2. By evaluating the algorithm on more than 250 data points, we show that the implementation is versatile, robust, and computationally efficient, often converging with just a few iterations and with a minimal error penalty. This paves the way for creating fully autonomous experiments, where data science algorithms such as the one presented herein operate hand-in-hand with the expert user to maximize scientific discovery and minimize the associated experimental cost.

1. Introduction

Small Angle X-ray Scattering (SAXS) is an experimental technique that has widespread use in structural biology and biophysics [1]. Primarily, due to the non-intrusive nature of the X-ray source, commonly generated by synchrotron radiation, and the small angles studied, SAXS is capable of characterizing density differences on the order of the nanoscale [2]. In turn, this makes the technique well-suited for studying protein structures and ensembles in solutions of biological samples [3, 4]. Moreover, owing to its simplicity, SAXS has been used to reveal molecular cluster formations and structural inhomogeneities in supercritical fluids [5, 6, 7, 8, 9].

In the context of supercritical fluids (SCFs), which are ubiquitous in nature [10, 11] and in engineering applications [12, 13, 14, 15], SAXS elucidates structural information through two main parameters: the correlation length, which gives a measure of the cluster size, and the density fluctuations, which provide a measure of the intra-molecular cluster variation. Both of these quantities are directly obtained from the SAXS intensity spectra, after background subtraction and following the Ornstein-Zernike theory [16].

From a thermodynamic perspective, the correlation length and density fluctuations are also of fundamental importance. In particular, the peaks in correlation lengths, when traversed in a pressure-temperature state space, trace a ridge that forms the Widom line [6, 8, 17]—a line that demarcates liquid-like from gas-like behavior in the supercritical phase diagram [18, 19], while the density fluctuations can be shown to proportionally relate to the isothermal compressibility [16, 20]—a measurable, macroscopic thermodynamic property. Thus, SAXS not only offers structural insight into the cluster inhomogeneities in supercritical fluids, but it additionlly provides a direct link between the microstructure and the macroscopic behavior of supercritical samples.

However, conducting SAXS experiments, at a given temperature and pressure condition, to acquire a statistically-converged correlation length and density fluctuation parameter is time consuming. On average, if conducted reasonably

*Corresponding author
kyounes@stanford.edu (K. Younes)
ORCID(s): 0000-0001-9562-0963 (K. Younes)

far from the critical point of the fluid, a single measurement takes approximately 5 minutes of beam time; closer to the critical point, minute instabilities disturb the system and a converged spectrum is harder to attain, causing the measurement time to rise exponentially. Hence, this makes an expansive investigation, spanning conditions of engineering relevance, extremely prohibitive [21]. Furthermore, since in most cases emphasis is placed on locating global features in the correlation lengths and density fluctuations, performing a manual search to locate those is not fruitful. The problem is compounded by the fact that access to modern synchrotron facilities suitable for performing such experiments is often oversubscribed and competitive, with a wait time of a few months up to a year.

In this paper, we propose to use Bayesian optimization to address the task of efficiently locating the extremum in the correlation lengths and density fluctuations. The approach is applied directly on the SAXS intensity curves and serves as a proof of concept to enable real-time and rapid screening of the thermodynamic state space.

The remainder of the paper is structured as follows: §2 presents the algorithmic implementation and provides details on the experimental setup and procedure, §3 highlights some key findings and benchmarks the results against a traditional search method, and, finally, conclusions are drawn in §4.

2. Methods

2.1. Bayesian Optimization

Bayesian optimization (BO) broadly belongs to a class of global optimization methods that is frequently applied to expensive, real-world problems [22]. The premise behind the algorithm is to leverage a probabilistic model, as opposed to a deterministic one, to represent the black-box function, $f(\mathbf{x})$ with $\mathbf{x}$ being the parameter space, that one is trying to optimize. Then, by incorporating measurement noise and any prior knowledge on the problem, Bayes' theorem is used to construct an acquisition function, $\alpha(\mathbf{x})$, that serves as a guide to inform the next experimental or numerical query of the unknown $f(\mathbf{x})$. Iteratively updating the model with new measurements, either until a budget constraint (e.g., a total number of experiments or user allocation time) is met, or a convergence criteria is attained, subsequently allows the probabilistic model to reconstruct an approximate depiction of the actual, black-box function $f(\mathbf{x})$ [22, 23]. Simply put, the goal of BO is to devise a gradient-free sampling strategy that locates the global maximum of the property of interest efficiently. A summary of the algorithm is given below.

Algorithm: Bayesian Optimization [22]

while convergence not reached and budget not exhausted
  Step 1: sample $\mathbf{x}_i$ from $f(\mathbf{x})$ such that acquisition function is optimal,
    $\mathbf{x}_i \rightarrow \text{argmax } \alpha(\mathbf{x})$
  Step 2: query $f(\mathbf{x})$ at newly sampled $\mathbf{x}_i$
  Step 3: update data set D with $\mathbf{x}_i$, $f(\mathbf{x}_i)$
  Step 4: update probabilistic model
end **while**

To this end, two underlying factors are necessary for an effective application of BO: (1) $f(\mathbf{x})$ has no analytical form, and/or (2) $f(\mathbf{x})$ is expensive to evaluate. The presence of these factors is crucial to justify the selection of BO as an extremum-seeking method. Indeed, this is the case in SAXS, for the experiment itself is expensive to conduct and the evolution of the structural parameters in the state space is not known *a priori*. In fact, the Bayesian framework has found applications in SAXS experiments before, to infer conformational ensemble structures in protein solutions [24, 25, 26, 27]. In this work, we use BO as a feature locator for SAXS data obtained from supercritical fluid samples. More specifically, while prior works have utilized Bayes' theorem to construct model fits that best resemble the SAXS spectra at a *single* parametric condition, here we use BO to sample and locate peaks, in the least number of queries possible, for a *wide range* of parametric conditions. As a result, the framework is formulated with an optimization loop to achieve this target.

We utilize a Gaussian process (GP) as our probabilisic model [28]. The advantage of using a GP is that, being a collection of infinitely-many random variables each composed of a multivariate normal distribution, derived quantities of interest, such as the mean values $\mu(\mathbf{x})$, are defined analytically. Furthermore, GPs embed prior knowledge on the problem using the covariance function, $K(\mathbf{x}, \mathbf{x})$, and provide prediction uncertainties explicitly. Thus, the user is able to intuitively follow the algorithmic execution and readily interpret the output of the acquisition function. For more details on Bayesian optimization and Gaussian processes, the reader is referred to Refs. [22] and [28], respectively. A detailed application of BO to an X-ray Free Electron Laser (XFEL) is also given in [29].

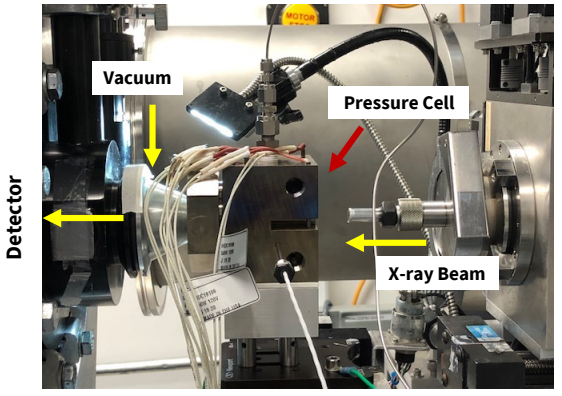

Figure 1: Schematic diagram showing the pressure cell as mounted on the beamline. The X-ray beam passes through the diamond window (from right to left) and scatters onto the detector.

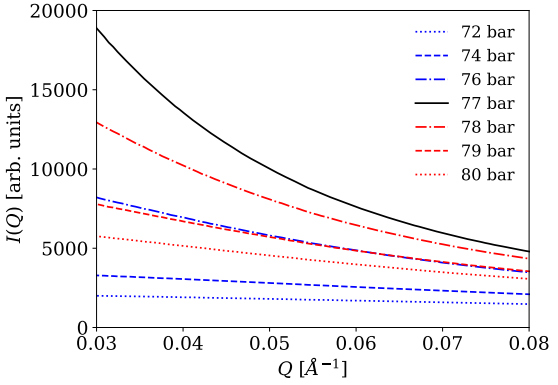

Figure 2: SAXS intensity curves *vs.* Q for supercritical CO_2 at $T = 33°C$. Each curve is obtained after the pressure reaches equilibrium, with 60 frames and an exposure time of 5 s per frame.

2.2. Experimental Setup and Procedure

A pressure cell is designed and built to contain the supercritical fluid and maintain the required temperature and pressure conditions to probe the supercritical phase space throughout the experiment. The pressure cell is manufactured from Titanium with an o-ring seal to provide corrosion resistance and ensure leakage-free operation. It houses two single-crystal diamond windows for optical access and six cartridge heaters connected to a cryo-con PID controller to accurately control the temperature; the pressure is controlled externally via a Teledyne Isco 100DM syringe pump. The cell has been successfully tested and is rated for temperatures up to 300°C and pressures up to 500 bar. A pictorial illustration of the cell, as mounted on the beamline is shown in Fig. 1. The design is similar to that presented by Fulton *et al.* [30], and a full characterization of its performance and inner-workings will be detailed elsewhere.

Due to its relevance in geological carbon sequestration [15], the sample studied is pure (99.999% purity) CO_2 (critical point: $T_c = 30.98°C$, $P_c = 73.8$ bar). The experiments were conducted at the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 4-2 within the SLAC National Accelerator Laboratory facility. The X-ray energy was 15 keV, and the data collection proceeded as follows. After mounting the cell on the beamline and aligning the beam, the sample is heated to the desired temperature above the critical point, i.e., $T > T_c$. Following temperature stabilization, which roughly took 30 minutes, the cell was then pressurized to the desired pressure in the supercritical state with $P > P_c$. The system is deemed to equilibriate when deviations in the temperature and pressure lie within $\Delta T = \pm 0.1\,°C$ and $\Delta P = \pm 0.1$ bar for a single measurement period, after which the shutter is opened and the X-ray beam is shone through onto the sample. 50 frames were collected for a given thermodynamic (T, P) condition, each with a 5 s exposure time; this is done in accordance with previously established procedures [5, 6, 17].

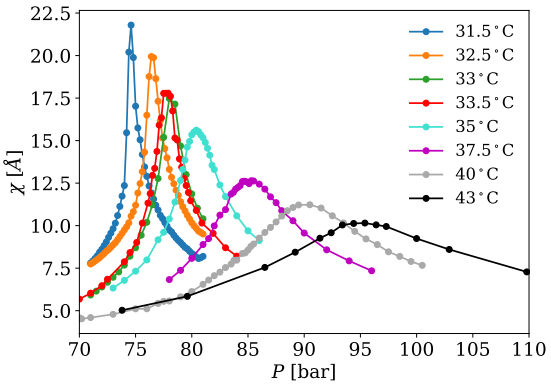

Figure 3: Correlation lengths, χ, as a function of pressure at different isotherms, calculated after performing the Ornstein-Zernike analysis on the SAXS intensity curves. Symbols show the experimentally measured points, whereas the curves resemble a cubic interpolation through the data.

The scattering intensity is obtained within a momentum transfer, Q, range between 0.03−0.08 Å^{-1}. Sample raw intensity signals for various pressures at $T = 33$°C are given in Fig. 2. To obtain the correlation length and density fluctuations from the SAXS spectra, the Ornstein-Zernike analysis is employed [16]. In particular, the correlation length is given by:

$$I(Q) = \frac{I(0)}{1 + \chi^2 Q^2}, \tag{1}$$

where I is the intensity and χ is the correlation length. Eq. (1) is inverted and a straight line is fitted to the data. The y-intercept of the linear fit, in combination with the slope, then allows for a direct solution of χ^2. Physically, the correlation length corresponds to the decay in the density-density autocorrelation function [20].

The density fluctuation is obtained from the zero-|Q| intensity according to:

$$\rho' = \frac{1}{Z^2}\frac{I(0)}{N}, \tag{2}$$

where Z is the total number of electrons and N is the number of CO_2 molecules in the scattering volume. It can also be shown that $\rho' = N k_b \kappa_T T$ [20], where N is the number of moles, k_b is the Boltzmann constant, and κ_T is the isothermal compressibility.

3. Results

A total of 271 operating points, each corresponding to a unique thermodynamic (T, P) condition probed in the supercritical phase space, spanning temperatures of $T = 31.5$−43°C and pressures of $P = 70$−110 bar were collected using SAXS on CO_2. The results for the correlation lengths are shown in Fig. 3 as a function of pressure, i.e., at different isotherms. Since it was shown by Nishikawa *et al.* [5, 6, 7, 8, 17] that for a pure supercritical fluid, such as the one studied herein, peaks in correlation lengths align with the peaks in density fluctuations, Bayesian optimization is applied only on the correlation length curves: $f(\mathbf{x}) \rightarrow \chi(P)$.

To do so, the Matérn 5/2 kernel is chosen as the covariance function $K(\mathbf{x}, \mathbf{x})$ [31], and the Monte-Carlo based expected improvement acquisition function is selected to guide the experimental screening [32]. Other combinations of covariance and acquisition functions were explored, but they did not yield improved performance (results not shown). To assess convergence, the algorithm tracks both the total number of experimental queries performed and the relative difference in pressure between two successive queries; it terminates when either (a) the total number of queries reaches 20 (experimental budget constraint) or (b) the difference between two successive queries is $< 1\%$. The former criterion guarantees that the algorithm does not run endlessly, potentially exhausting valuable experimental resources, while the latter is set to avoid excessive repetitive sampling within the same neighborhood of points. As will be shown below, the experimental budget constraint is almost never reached, and the convergence will be solely based on the relative

Table 1
Summary of all cases studied with and without Bayesian optimization. The error is defined as follows: $100\times|P_{BO}-P_{exp}|/P_{exp}$. The relative tolerance threshold for convergence is set to 10^{-2}, and the last column gives the total number of BO queries before that is met.

Temp. (°C)	Exp. Peak (bar)	BO Peak (bar)	Error (%)	Queries
31.5	74.58	74.57	0.01	7
32.5	76.49	74.50	0.02	6
33	78.19	78.22	0.04	5
33.5	77.57	77.58	0.01	10
35	80.42	80.43	0.01	5
37.5	85.38	85.1	0.33	6
40	89.98	90.04	0.07	6
43	95.22	95.12	0.1	6

position of two successive queries in the pressure domain. Measurement noise is incorporated explicitly into the GP through the standard deviation of the correlation length. In general, this value does not exceed 15% for all sampled points. The entire BO implementation is programmed using BoTorch in Python v. 3.9 [33], and the data and code are available upon request.

An illustrative step-by-step example showcasing the operation of the algorithm is given in Figs. 4 and 5 for the $T = 32.5$°C isotherm. A few observations regarding the figures are worth mentioning. First, the algorithm takes as input an initial distribution of sampled points. In turn, this can be chosen either randomly or uniformly in a number of ways. For this work, we choose the initial sample to consist of the lower and upper bounds of the pressure range of the given isotherm. Since the input parameter space is normalized by the maximum pressure value in the implementation, this ascertains that the algorithm does not visit physically unfeasible pressure conditions. In practice, queries can be penalized according to their associated cost [33]. However, this step was found to be unnecessary for the analysis conducted here. Second, it can be seen that, at least initially, the confidence interval of the GP is relatively large (first two iterations of the BO algorithm in Fig. 4). Yet, the algorithm automatically improves the predictions with each new iteration. In essence, this is the premise behind employing the expected improvement acquisition function, where the goal is to maximize the certainty in the model outcome. Third, after sampling a few points in the high pressure region, $P > 75$ bar, the algorithm detects that the low pressure region is still unexplored. To combat this, it balances between exploration and exploitation by querying the low $P-$region of the domain. From a computational viewpoint, one can specify exactly how much exploration *vs*. exploitation is done by the algorithm through setting the exploration constant in the acquisition function formulation. In this instance, the default `qExpectedImprovement` formulation in BoTorch was used with $\xi = 0$ as the exploration constant [33]. Fourth, by the fifth iteration (Fig. 5), the convergence to the predicted maximum is within 1%, and the uncertainty in the vicinity of the peak is small. Despite that, the algorithm performs an additional iteration (Fig. 5, bottom) to ensure that the global maximum, and not a local one, was indeed visited; this is further verified by noting that the expected improvement acquisition function (right column in Fig. 5) takes the shape of a delta function, indicating that the algorithm has high certainty near the maximum region. Finally, in spite of having much larger uncertainty in the low pressure region of the domain, BO successfully arrives at the maximum of the correlation length with $< 0.05\%$ error. This demonstrates the versatility of the method and highlights the robustness of using a GP as the probabilistic model.

The algorithm was also applied for the remainder of the isotherms at $T = 31.5, 33, 33.5, 35, 37.5, 40$, and 43°C. The findings and error metric, with respect to the experimentally measured maxima, are summarized in Table 1 and overlaid on the phase diagram of CO_2 in Fig. 6. In addition to that, the performance is benchmarked against the bisection method [34] in Fig. 7. (The bisection method, also known as the interval-halving method, locates the maximum of the correlation length in threefold. First, it bisects a given pressure interval in half; for consistency, the initial interval is set to be the same as the input to the BO loop with 2 starting points—the lowest and highest pressure conditions studied at the isotherm. Second, it computes the correlation length for the three resulting pressure values. Third, it selects the 2-point interval with the highest correlation length values and successively bisects it until convergence is attained, i.e., until the pressure values in the interval do not change by more than 1%). As can be seen, the BO algorithm consistently, with the exception of the $T = 33.5$°C case, achieves faster convergence when compared to the bisection search method, with savings ranging from 30−50% and with minimal error incurred. On average, this translates to ∼400 minutes of spared user beam time.

4. Conclusions

SAXS experiments offer a unique perspective on supercritical fluids through delineating structural parameters, namely, the correlation length and density flucutations, that reveal the molecular cluster formations and cluster inhomogeneities, respectively. In turn, conducting scattering experiments for a wide range of pressure and temperature conditions spanning the supercritical phase space, which is relevant for a growing number of engineering applications, is practically infeasible. Primarily, this is due to the prohibitive cost and resources associated with traversing such a complex space. Yet, in most cases, detecting global features in the correlation lengths and density fluctuations is sufficient. Instead of performing a manual grid or a bisection search, in this work, we employ Bayesian optimization to achieve this purpose. Being a gradient-free search tool, BO devises a sampling strategy, based on a probabilistic model, that efficiently locates the maxima in the correlation length structural parameter. The advantage of such an approach is that it remains robust to measurement noise and uncertainties, which are directly incorporated into the implementation. Furthermore, the implementation is computationally cheap and can be run in real-time with minimal overhead. It was shown that, on average, the algorithm locates the extremum of the correlation length in 3−6 fewer iterations than the bisection search method, saving valuable user beam time. The work demonstrated the viability of the method on SAXS spectra, but the approach is generalizable and can be applied to several other large-scale X-ray experiments, including but not limited to, X-ray Photon Correlation Spectroscopy (XPCS) and X-ray Absorption Spectroscopy (XAS).

Acknowledgements

The authors would like to acknowledge funding from the U.S. Department of Energy, Office of Science through DOE (BES) Awards DE-SC0021129 and DE-SC0022222. K. Younes also acknowledges partial financial support from the Natural Sciences and Engineering Research Council of Canada through the NSERC PGS-D award.

References

[1] L. Boldon, F. Laliberte, L. Liu, Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application, Nano Reviews 6 (1) (2015) 25661. doi:10.3402/nano.v6.25661.

[2] T. Li, A. J. Senesi, B. Lee, Small Angle X-ray Scattering for Nanoparticle Research, Chemical Reviews 116 (18) (2016) 11128–11180. doi:10.1021/acs.chemrev.5b00690.

[3] A. G. Kikhney, D. I. Svergun, A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins, FEBS Letters 589 (2015) 2570–2577. doi:10.1016/j.febslet.2015.08.027.

[4] C. Prior, O. R. Davies, D. Bruce, E. Pohl, Obtaining Tertiary Protein Structures by the ab Initio Interpretation of Small Angle X-ray Scattering Data, Journal of Chemical Theory and Computation 16 (3) (2020) 1985–2001. doi:10.1021/acs.jctc.9b01010.

[5] K. Nishikawa, I. Tanaka, Correlation lengths and density fluctuations in supercritical states of carbon dioxide, Chemical Physics Letters 244 (1995) 149–152. doi:10.1016/0009-2614(95)00818-0.

[6] K. Nishikawa, I. Tanaka, Y. Amemiya, Small-Angle X-ray Scattering Study of Supercritical Carbon Dioxide, J. Phys. Chem. 100 (1996) 418–421. doi:10.1021/jp951803p.

[7] K. Nishikawa, T. Morita, Small-Angle X-ray-Scattering Study of Supercritical Trifluoromethane, J. Phys. Chem. B 101 (1997) 1413–1418. doi:10.1021/jp963075r.

[8] K. Nishikawa, T. Morita, Inhomogeneity of molecular distribution in supercritical fluids, Chemical Physics Letters 316 (2000) 238–242. doi:10.1016/S0009-2614(99)01241-5.

[9] P. Muhunthan, K. Younes, D. Sokaras, T. Matsui, I. Rajkovic, T. M. Weiss, M. Ihme, Small Angle X-ray Scattering of Supercritical CO2: Comparison with EoS, in prep. (2023).

[10] W. Martin, J. Baross, D. Kelley, M. J. Russell, Hydrothermal vents and the origin of life, Nature Reviews Microbiology 6 (11) (2008) 805–814. doi:10.1038/nrmicro1991.

[11] D. Bolmatov, V. V. Brazhkin, K. Trachenko, Thermodynamic behaviour of supercritical matter, Nature Communications 4 (1) (2013) 2331. doi:10.1038/ncomms3331.

[12] P. E. Savage, Organic Chemical Reactions in Supercritical Water, Chemical Reviews 99 (2) (1999) 603–622. doi:10.1021/cr9700989.

[13] J. Bellan, Supercritical (and subcritical) fluid behavior and modeling: drops, streams, shear and mixing layers, jets and sprays, Progress in Energy and Combustion Science 26 (4-6) (2000) 329–366. doi:10.1016/S0360-1285(00)00008-3.

[14] G. Brunner, Applications of Supercritical Fluids, Annual Review of Chemical and Biomolecular Engineering 1 (1) (2010) 321–342. doi:10.1146/annurev-chembioeng-073009-101311.

[15] H. De Coninck, S. M. Benson, Carbon dioxide capture and storage: Issues and prospects, Annual Review of Environment and Resources 39 (2014) 243–270. doi:10.1146/annurev-environ-032112-095222.

[16] L. S. Ornstein, F. Zernike, Accidental deviations of density and opalescence at the critical point of a single substance, Proceedings of the Royal Netherlands Academy of Arts and Sciences 17 (1914) 793–806.

[17] K. Nishikawa, A. A. Arai, T. Morita, Density fluctuation of supercritical fluids obtained from small-angle X-ray scattering experiment and thermodynamic calculation, Journal of Supercritical Fluids 30 (3) (2004) 249–257. doi:10.1016/j.supflu.2003.09.003.

[18] P. F. McMillan, H. E. Stanley, Fluid phases: Going supercritical, Nature Physics 6 (7) (2010) 479–480. doi:10.1038/nphys1711.
[19] G. G. Simeoni, T. Bryk, F. A. Gorelli, M. Krisch, G. Ruocco, M. Santoro, T. Scopigno, The Widom line as the crossover between liquid-like and gas-like behaviour in supercritical fluids, Nature Physics 6 (7) (2010) 503–507. doi:10.1038/nphys1683.
[20] H. E. Stanley, Introduction to Phase Transitions and Critical Phenomena, Oxford University Press, New York, 1971.
[21] N. Spycher, K. Pruess, CO2-H2O mixtures in the geological sequestration of CO2. II. Partitioning in chloride brines at 12–100°C and up to 600 bar, Geochimica et Cosmochimica Acta 69 (13) (2005) 3309–3320. doi:10.1016/j.gca.2005.01.015.
[22] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, N. de Freitas, Taking the Human Out of the Loop: A Review of Bayesian Optimization, Proceedings of the IEEE 104 (1) (2016) 148–175. doi:10.1109/JPROC.2015.2494218.
[23] J. Mockus, Bayesian Approach to Global Optimization: Theory and Applications, Springer Netherlands, 2012.
[24] L. D. Antonov, S. Olsson, W. Boomsma, T. Hamelryck, Bayesian inference of protein ensembles from SAXS data, Physical Chemistry Chemical Physics 18 (8) (2016) 5832–5838. doi:10.1039/C5CP04886A.
[25] S. Bowerman, A. S. Rana, A. Rice, G. H. Pham, E. R. Strieter, J. Wereszczynski, Determining Atomistic SAXS Models of Tri-Ubiquitin Chains from Bayesian Analysis of Accelerated Molecular Dynamics Simulations, Journal of Chemical Theory and Computation 13 (6) (2017) 2418–2429. doi:10.1021/acs.jctc.7b00059.
[26] W. Potrzebowski, J. Trewhella, I. Andre, Bayesian inference of protein conformational ensembles from limited structural data, PLOS Computational Biology 14 (12) (2018) e1006641. doi:10.1371/journal.pcbi.1006641.
[27] F. Pesce, K. Lindorff-Larsen, Refining conformational ensembles of flexible proteins against small-angle x-ray scattering data, Biophysical Journal 120 (22) (2021) 5124–5135. doi:10.1016/j.bpj.2021.10.003.
[28] C. E. Rasmussen, C. K. I. Williams, Gaussian Processes for Machine Learning , MIT Press, Cambridge, 2006.
[29] J. Duris, D. Kennedy, A. Hanuka, J. Shtalenkova, A. Edelen, P. Baxevanis, A. Egger, T. Cope, M. McIntire, S. Ermon, D. Ratner, Bayesian Optimization of a Free-Electron Laser, Physical Review Letters 124 (12) (3 2020). doi:10.1103/PhysRevLett.124.124801.
[30] J. L. Fulton, Y. Chen, S. M. Heald, M. Balasubramanian, High-pressure, high-temperature x-ray absorption fine structure transmission cell for the study of aqueous ions with low absorption-edge energies, Review of Scientific Instruments 75 (12) (2004) 5228–5231. doi:10.1063/1.1813131.
[31] M. G. Genton, Classes of Kernels for Machine Learning: A Statistics Perspective, Journal of Machine Learning Research 2 (2001) 299–312. doi:10.5555/944790.944815.
[32] J. T. Wilson, R. Moriconi, F. Hutter, M. P. Deisenroth, The reparameterization trick for acquisition functions, ArXiv (2017). doi:10.48550/arXiv.1712.00424.
[33] M. Balandat, B. Karrer, D. R. Jiang, S. Daulton, B. Letham, A. G. Wilson, E. Bakshy, BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization, ArXiv (2020). doi:10.48550/arXiv.1910.06403.
[34] R. L. Burden, J. D. Faires, Numerical Analysis, 3rd Edition, Prindle, Weber, Schmidt, Boston, 1985.

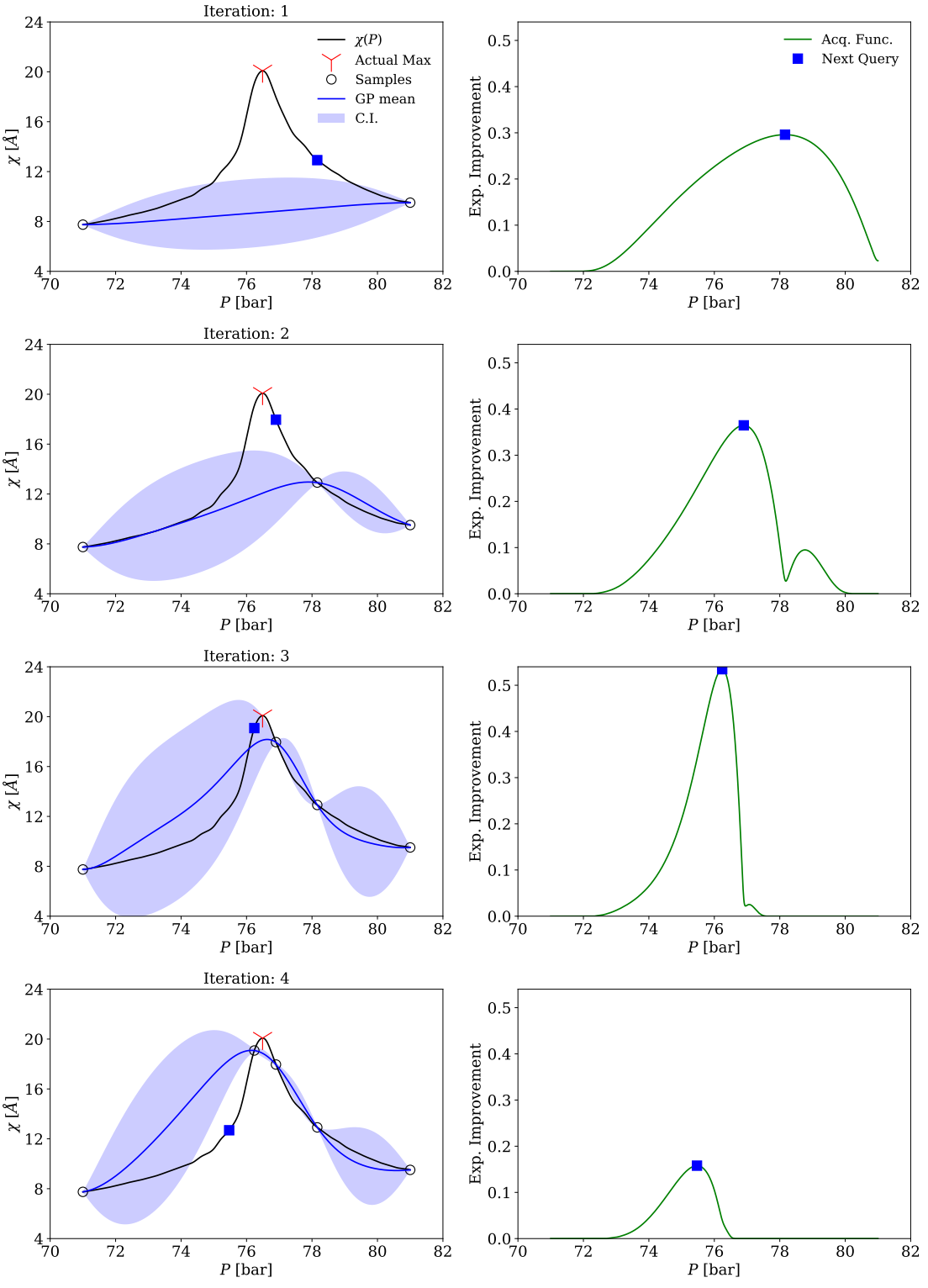

Figure 4: Demonstration of Bayesian optimization with the first 4 iterations for the $T = 32.5°C$ isotherm (orange curve in Fig. 3). Note that the queries performed by the algorithm are given by the open circles and that the algorithm has no access to the interpolated black curve and the actual location of the maximum (red cross), which are shown for illustrative purposes only. The corresponding expected improvement acquisition function is shown in green, on the right, and the next experimental query is chosen based on its global maximum; this serves to minimize the uncertainty in the Gaussian process mean prediction (blue curve) and shrink the 95% confidence interval (shaded blue region).

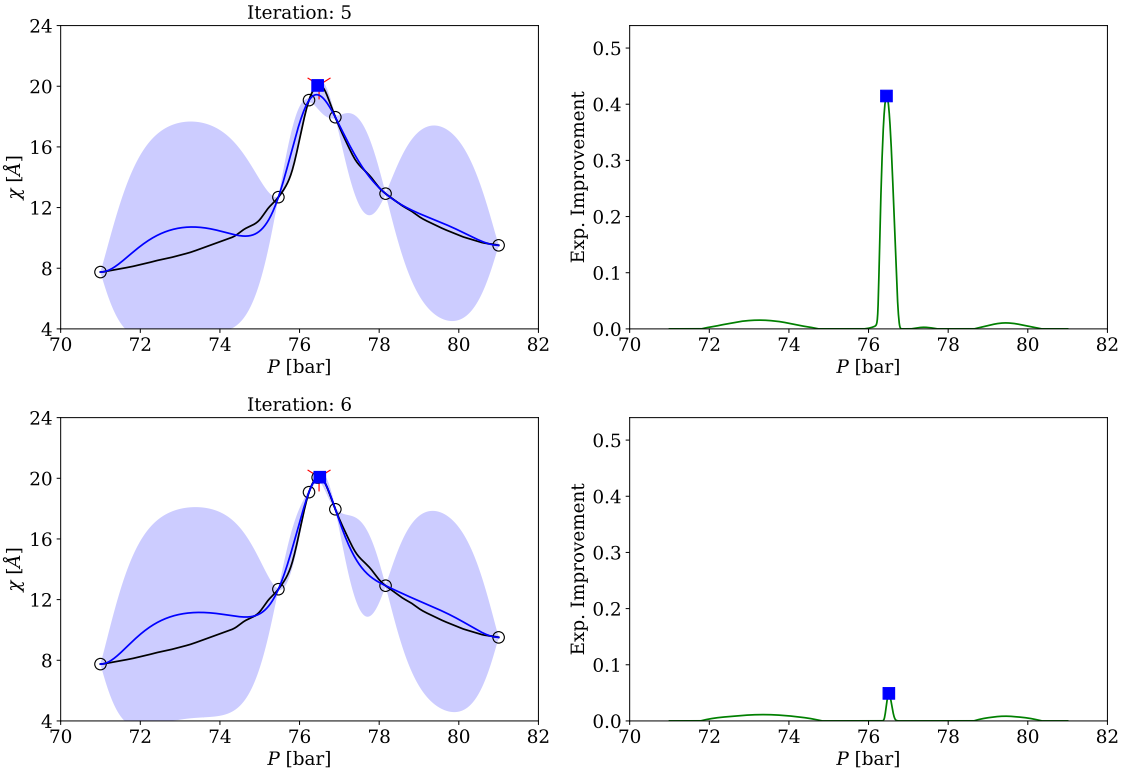

Figure 5: The last 2 iterations of BO as applied on the $T = 32.5°C$ isotherm (orange curve in Fig. 3). These iterations are performed by the algorithm to ensure that the global maximum, and not a local maximum, is attained. The labels and legend are identical to that in Fig. 4.

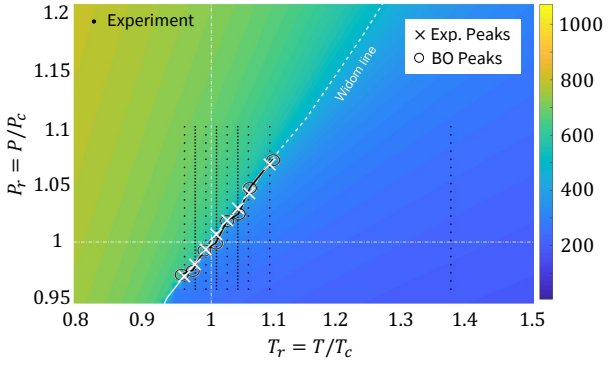

Figure 6: The phase diagram of CO_2 overlaid with the experimentally conducted conditions (black dots), experimentally measured peaks in the correlation length (white crosses), and the predicted peaks by the BO algorithm (open circles). The error remains below 1% for all cases. Contours show the density in kg/m^3.

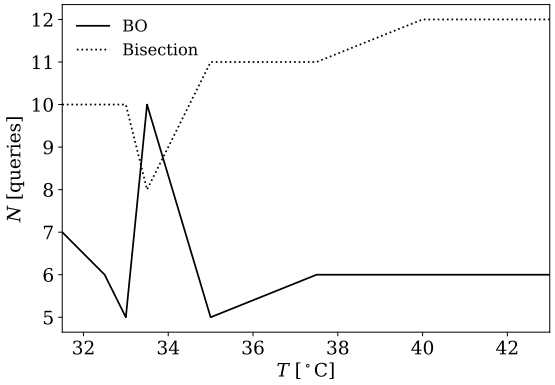

Figure 7: Number of queries to reach convergence at each isotherm presented in Table 1 for the Bayesian optimization algorithm (solid line) and the bisection search method (dotted line) [34].