

Adsorbate Organization Characterized by Sublevelset Persistent Homology

Nitesh Kumar^{a*} and Aurora E Clark^{a,b†}

^a Department of Chemistry, Washington State University, Pullman, WA, 99163

^b Department of Chemistry, University of Utah, Salt Lake City, UT, 84112

(Dated: January 17, 2023)

Interfacial adsorbate organization influences a variety physicochemical properties and reactivity. Surfaces that are rough, defect laden, or have large fluctuations (as in soft matter interfaces), can lead to complex adsorbate structures. This is amplified if adsorbate-adsorbate interactions lead to self-assembly. Although image analysis algorithms somewhat common for the study of solid interfaces (from microscopy for example), images are often not readily available for adsorbates at soft matter surfaces and the complexity of adsorbate organization necessitates the development of new characterization approaches. Here we propose the use of adsorbate “density” images from molecular dynamics simulations of liquid/vapor and liquid/liquid interfaces. Topological data analysis is employed to characterize surface active amphiphile self-assembly under non-reactive and reactive conditions. We develop a chemical interpretation of sublevelset persistent homology barcode representations of the density images, in addition to descriptors that clearly between different reactive and non-reactive organizational regimes. The complexity of amphiphile self-assembly at highly dynamic liquid/liquid interfaces represents a worst-case scenario for adsorbate characterization and as such the methodology developed is completely generalizable to a wide variety of surface image data, whether from experiment or computer simulation.

I. Introduction

Despite their fundamental importance, liquid/liquid interfaces are notoriously challenging to characterize both experimentally and within the analysis of modeling and simulation data. Although macroscopic properties like interfacial tension and surface potential are relatively easy to determine, several factors make a molecular-scale understanding of interfacial organization a challenge.[1–3] The small lengthscale of the interfacial region (extending only a few molecular widths) limits the spectroscopic and scattering methods that provide chemical insight; the most prevalent are Vibrational Sum Frequency Generation (VSFG) or Second Harmonic Generation (SHG) and photo-emission, while X-ray scattering and reflectivity methods have also been developed.[4–6] Atomistic molecular dynamics (MD) simulations have played an incredibly important role in providing molecular-level details of the interfacial structure. For both simulation and experiment, the concentration gradients of solutes approaching the interface and their adsorption behavior, as well as their impact on interfacial properties such as reactivity and transport, have been a major area of study.[7–10]

To define the interface more precisely, we introduce the *Gibbs dividing surface* (GDS) and the *instantaneous surface*. The ensemble average position of the interface may be represented by the Gibbs dividing surface, which is the position of an ideal, flat, interfacial plane where the density of one of the cosolvents (often water) is half the density in the bulk.[11] The typical coordinate system employed has the interfacial plane in the xy dimension and the perpendicular direction to the z axis. In

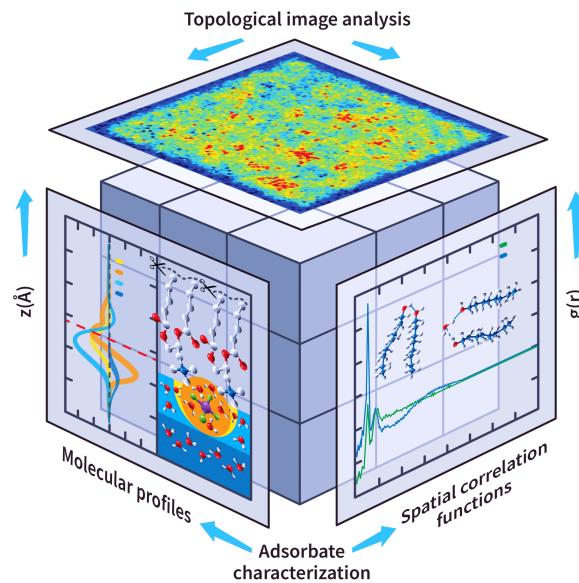


FIG. 1. An overview of the analysis involved in the characterization of molecular organization and/or speciation based upon the spatial correlation function, molecular profiles, and the persistent homology image analysis from adsorbate densities developed within this work.

general, the interface is, however, not flat and instead there has been significant effort toward understanding the characteristics of the instantaneous surface that represents the molecular-scale film where two immiscible phases meet. The Willard-Chandler (WC) surface is a coarse-grained representation based upon the isodensity surface of the interfacial solvent molecules (typically water).[12] Complementary is the Identification of

* nitesh.kumar@wsu.edu

† aurora.clark@utah.edu

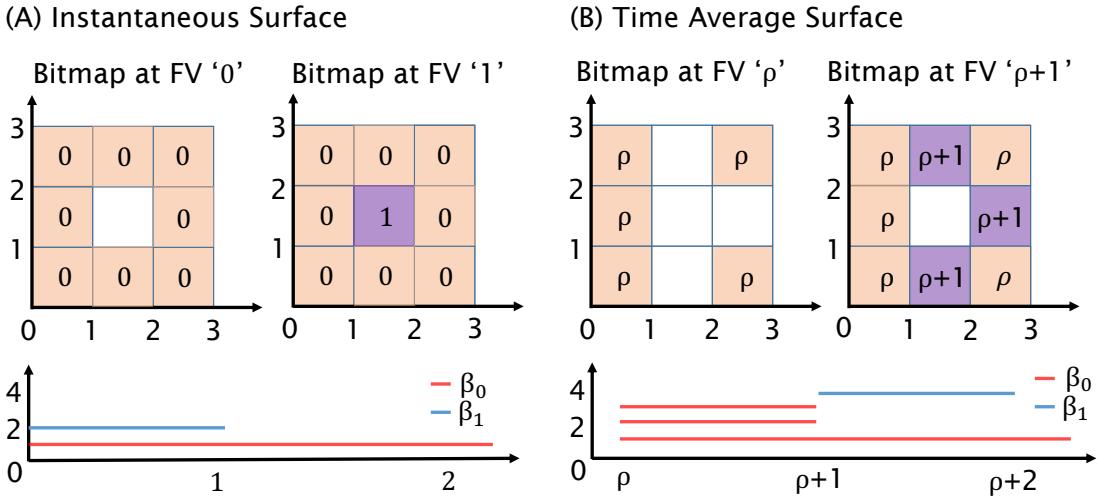


FIG. 2. The topology characterization of the (A) instantaneous and (B) time-averaged density surfaces using the sublevelset persistent homology. The example shows the generation of bitmaps for two filtration values (FV, referred to as sublevels), with $FV=\rho$ and $\rho+1$ (0 and 1 in case of the instantaneous surface). The barcodes show the persistence of 0- and 1-dimensional homology features (β_0 and β_1) as a function of filtration value (ρ).

Truly Interfacial Molecules algorithm (ITIM), which uses probe spheres perpendicular to the instantaneous surface to identify the solvent molecules of the immiscible solvents that are in direct contact with one another.[13] Along the GDS or the instantaneous surface, several complementary analyses may be performed. Analysis of MD simulations often focuses upon the 1-d distribution of molecules (or molecular density) perpendicular to the interfacial plane. Molecular density distributions of molecules present in the instantaneous surface can be analyzed to determine the interfacial width (IW); this is derived from the full width at half maximum of the mean density distribution and can be compared directly to values obtained using X-ray reflectivity.[14, 15] Using algorithms like ITIM, the identity of the molecules present at the surface makes it possible to create spatial correlation functions of the interfacial organization. For example, 2-dimensional radial distribution functions (RDFs) can be created based on the projection of the surface molecules into an xy plane. This approach is useful for understanding adsorbate-adsorbate interactions.[16–18] Subensemble analysis can also be performed, revealing the breadth of local molecular environments that contribute to interfacial reactions or responses to stimuli. For example, recent work has used ITIM to isolate different orientations of H₂O at the interface that result from different solvation environments about adsorbed ions.[19, 20] This was then used to rationalize the ensemble average changes to the orientation of water measured by VSFG as you approach the interface. Based upon this discussion we classify the analysis of MD data into two representative groups, those based upon spatial correlation functions and those based upon molecular distributions as illustrated in Figure 1.

Although these analyses provide significant insight, there remains much benefit to expand the computational tools that describe interfacial organization more comprehensively – to quantitatively incorporate concepts of order/disorder of adsorbate configurations and to characterize organization at different length scales. Toward this end, we have borrowed well-established concepts from image processing to develop a robust approach for characterizing molecular organization at the liquid interface (or any other) and introduce descriptors that trend with well-known interfacial properties. In this work, we consider the surface image obtained by projecting atoms or molecules on the instantaneous surface onto a xy plane (a 2-dim manifold embedded in 3-dim space). Within the topological data analysis (TDA) of manifolds, sublevelset persistent homology (PH) is a valuable tool and has been employed in many studies to characterize image data.[21–25] It is particularly adept at identifying minor differences or anomalies within identical images.[26–28] Most image analyses have focused on *time-independent* data, however in computational chemistry temporal PH descriptors could dramatically enhance our understanding of adsorption/desorption phenomena at surfaces and reactivity.

The image of adsorbates on the surface may be created in multiple ways. We consider two cases in which the adsorption at any instant in time is evaluated and the second case using a short-time average of adsorbate positions (to create a density image). In both cases, a grid in xy is first created and then the projection of the adsorbate from the 3-dimensional instantaneous liquid-liquid interface onto the xy grid generates a bitmap, where each cell of the bitmap represents a point whose value is the number or density of adsorbates in that location.

An example of the resulting image is shown at the top of Figure 1. The features in the image are encoded using sublevelset persistent homology, which tracks the number and dimensionality of connected components within the image as a function of a filtration value (FV, also known as a sublevel) that spans the set of pixel intensities. As shown in Figure 2, the lowest dimension information, called zero-dimensional topological features (labelled β_0) are connected components that are isolated pixels or those that are connected because they have an intensity value less than or equal to filtration value. A cycle or hole in the image is considered 1-dimensional information and labelled β_1 and is indicative of a density barrier or maximum that has yet to be overcome in the filtration and is surrounded by a low-lying basin. The persistence of these topological features is tracked across all sublevels by monitoring the first appearance (or *birth*) of the feature at a specific FV and its disappearance (or *death*) caused by merging with other components at a higher-valued filtration. We utilize a compact barcode representation of this information, though alternative representations could include persistence diagrams and persistence landscapes.[29] Analysis of the distribution of topological features can be used to create metrics or descriptors of the image - as has been done in image of liquid crystal nanocomposites[26] or immune cells.[30] Additionally, the vectorized barcode representation is easily implemented in machine learning frameworks for the identification of specific image features.[31]

The current work demonstrates the broad applicability of sublevel set persistence to understand images that convey adsorbate organization and concentration at an interface. PH descriptors quantify organizational features in a manner consistent with chemical intuition and also incorporate important new information regarding the scale of topological features and their temporal evolution. In turn, this provides added insight into the impact of long-range correlations upon interfacial behavior and reactivity.

II. Data Set

Two biphasic chemical systems and one model interface were investigated to validate the image analysis workflow and the chemical relevance of various persistent homology descriptors. These consist of a surfactant amphiphile, tributylphosphate (TBP), adsorbed to either (1) a water/vapor interface or (2) a $\text{LiNO}_3\text{(aq)}$ interface with hexane. TBP is a well-known amphiphile in the separations of complex aqueous mixtures where its adsorption behavior and transport properties of solutes have been well-studied.[32–34] At the water/vapor interface all TBP in the system will be adsorbed to the interface. At the $\text{LiNO}_3\text{(aq)}$ /hexane interface the concentration of adsorbed TBP can be tuned by $\text{LiNO}_3\text{(aq)}$ concentration to alter both the average interfacial organization as well as the dynamics of adsorption and desorption. The equilibrium organization of TBP at the $\text{LiNO}_3\text{(aq)}$ / hexane in-

terface, analyzed through the sublevel set PH, was compared to a model image constructed based on a density analogous to TBP randomly placed at the interface. This enabled detailed comparisons of the image of the adsorbed TBP density distribution on the surface in an ideal stochastic image vs. that from the real equilibrated system.

A. Molecular Dynamics Simulations

Chemical Systems The simulation boxes consisted of two liquid interfaces with either the vapor or hexane phases as shown in Figure S1 in the Supplementary Information and with exact molecular compositions presented in Table S1. Four different TBP adsorbed water/vapor systems were generated with initial configurations that randomly placed 72, 96, 120, and 144 TBP molecules at each interface. Three different $\text{LiNO}_3\text{(aq)}$ /TBP/hexane systems were generated by placing 238 TBP molecules in the organic phase. The aqueous phase electrolyte concentration varied from 1, 3, and 5 M $\text{LiNO}_3\text{(aq)}$. The chosen TBP concentration led to a 50% v/v fraction (i.e., 1.5 M TBP) in the organic phase, which is consistent with active extraction conditions of TBP transporting H_2O and LiNO_3 into the organic phase.[18]

Equilibration Protocol The equilibration protocol has been benchmarked in Ref. 18. Briefly, all-atom classical molecular dynamics simulations were performed in the GROMACS software package.[35] A leap-frog algorithm with a time step of 2 fs was used to integrate molecular motions. Each water/TBP/vapor system was simulated for 40 ns in an NVT ensemble, and the last 20 ns were used for analysis. The $\text{LiNO}_3\text{(aq)}$ /TBP/hexane systems were equilibrated in NPT and NVT for a minimum of 20 ns using the Nose-Hoover thermostat and Parrinello Rahman barostat.[36, 37] Production runs were performed for more than 150 ns in the NVT ensemble. The electronic continuum correlation (ECC) optimized Li^+ and NO_3^- force fields were employed in combination with TIP3P water as benchmarked in Ref. 38 and the optimized TBP force fields taken from Ye et al.[39] Long-range electrostatics were handled using Particle-mesh Ewald (PME) summation.[40] A 16 Å cutoff was used for both short-ranged and van der Waals atomic interactions. Hydrogen-containing bonds were constrained using LINCS.[41] The calculation of collision diameter and well depth cross-terms were evaluated using the Lorentz-Berthelot combination rules.[42]

B. Surface Construction and Bitmap Creation

Within the simulation trajectory, TBP molecules adsorbed to the instantaneous surface were identified using the ITIM algorithm[14] in a manner consistent with Ref. 18. Density distributions were obtained by projecting the atomic coordinates of the P=O O-atoms onto the $60 \times 60 \text{ \AA}$ xy plane using a 100×100 grid. The bitmap representation of TBP sorption at a single point in time is a distribution of pixels with values of 0 or 1 (absence or

presence of TBP). As indicated in Figure 2, at a filtration value of 0 the pixels having a value of 0 are connected. A single TBP adsorbed represents a cycle ($\rho = 1$) at FV = 0 because it is a pixel intensity maximum that has yet to be overcome at that filtration level. The temporal fluctuation of this image type was analyzed using a variety of metrics (*vide infra*). Alternatively, adsorbate images were created representing a *density distribution* of adsorbed TBP over a short-timescale. Here, we analyzed the time average of bitmaps over 100 snapshots, separated in time by 100 ps, which creates a surface density whose maximum is $30 \times 10^{-4} \text{ \AA}^{-2}$ for each pixel. This time separation is within the average lifetime of the TBP sorbed to the surface and thus these short time-averaged densities should capture the average short-time distribution of TBP on the surface. As shown in Figure S2, similar PH bar code distributions (Betti curves) are observed for the zero and 1-dimensional features irrespective of whether a 10 ps spacing or 100 ps spacing of snapshots was employed. A total of 10 time-averaged surfaces were generated to obtain the average and standard deviations of the persistent homology descriptors. The impact of grid size upon topological features was examined, as described in the Supplementary Information (Figures S3 - S5).

C. Model Stochastic TBP Adsorbate Images

Using the time averaged concentration of the adsorbed TBP at the $\text{LiNO}_3\text{(aq)}/\text{hexane}$ interface (Table I), images were constructed to represent stochastic distributions of adsorbed TBP. The stochastic image was constructed using Mersenne Twister, a pseudo-random number generator employed in the NumPy Python[43] package on 100×100 grids with an upper-density limit set to the maximum TBP density observed from the short time-average images from the equilibrated simulation trajectories. The mean of 10 surfaces obtained from the pseudo-random number generator was used for sublevelset PH analysis. The sum of the densities in the grid representation for both the equilibrated and stochastic distributions is the average number of TBP adsorbed to the interface, as compared in Table I.

TABLE I. Ensemble averaged number of adsorbed TBP per interface in $\text{LiNO}_3\text{(aq)}/\text{hexane}$ as compared to that obtained from the stochastic reconstruction.

$[\text{LiNO}_3]$ M	Equilibrated System	Stochastic Distribution
1	79.3	72.3
3	49.7	47.6
5	30.2	35.0

III. Sublevelset Persistent Homology Analysis Protocol

The computational workflow is illustrated in Figure S6, beginning with the ingestion of the molecular dynamics coordinates of an interface, the creation of the instantaneous surface, the projection of the adsorbed molecular density on the xy plane, and terminating with the calcu-

lation of the sublevel persistent homology features (calculated using the GUDHI software package[44]). The scripts developed in this work, alongside model data, are provided in the GitLab repository at Ref. 45.

A. Sublevelset PH Descriptors

The sublevelset PH descriptors are categorized into two types: 1) those descriptors that encompass the behavior of the sublevelset PH features over the entire FV range and are represented as a distribution, and 2) those descriptors that analyze the distribution to yield a single number. The range of filtration values sampled depends upon whether the instantaneous or averaged adsorbate images were examined. In the former, the pixel intensities are either 0 or 1, while in the latter the two-dimensional density distribution $\rho(x, y)$ employs a $\text{FV}(\rho)$ that is sampled every $\Delta\rho = 1.0 \times 10^{-4} \text{ \AA}^{-2}$. Any of these descriptors may be studied as a function of time to yield information about the dynamic evolution of the adsorbate organization.

Betti Curves and Excess Betti Curves. The output of the sublevelset PH (number of n-dimensional bars, birth and death times) is used to compute the *Betti curve*; this is the total number of β_0 or β_1 bars as a function of FV.[46–49] In this case, it makes the most sense to examine Betti curves of the short-time average time of TBP adsorption, which creates a density image. The difference between the Betti curves obtained from a two-dimensional surface representing the equilibrium density distribution of TBP ($\beta_n(\rho)$) and the analogous stochastic distributions ($\beta'_n(x)$) was used to obtain excess Betti curves:

$$\tilde{\beta}_n(\rho) = \beta_n(\rho) - \beta'_n(\rho). \quad (1)$$

The excess Betti curves indicate the change to organizational features relative to a random TBP density distribution. In the real simulation data TBP is known to form $\text{TBP}(\text{H}_2\text{O})\text{TBP}$ dimers that further assemble into “protrusion”, self-assembled macrostructures that are responsible for solute transport.[18, 33]

Betti Index. The cumulative number of bars over the sublevel set of the Betti curves provides insight into how quickly the components merge as a function of $\text{FV}(\rho)$. For example, if at the lowest $\text{FV}(\rho)$ there are 200 β_0 components that are all within a narrow range of ρ , then those components will merge (or die) very quickly and the cumulative number of bars will be close to the value of the lowest ρ sublevel. However, if there are large variations in ρ that inhibit the merging of the components (indicating “patches” of TBP density or heterogeneity of the surface adsorption) then the sum of all bars across all ρ will be much higher. To obtain the the cumulative β_n , the values of the Betti-curve was first interpolated using basis-spline function[50] as implemented in

Scipy[51] and distributed onto 10^3 bins and integrated as:

$$h(\beta_n) = \int_0^{\rho_{\max}} \beta_n dx. \quad (2)$$

The value of $h_i(\beta_n)$ at ρ_{\max} is termed the Betti index. A high Betti index represents a high density of bars within the persistent barcode that persist for a long range of ρ values. Further, the excess Betti index is defined as

$$\tilde{h}(\beta_n) = h(\beta_n) - h'(\beta_n), \quad (3)$$

where $h(\beta_n)$ and $h'(\beta_n)$ represent the Betti index for time-averaged and the analogous stochastic distributions, respectively.

Persistent Lifetimes and Entropies. The lifetime of the bar is the bar length in ρ is $l_i = \rho_b - \rho_d$, where ρ_b and ρ_d are the FV values at birth and death. A descriptor of the distribution of lifetimes is the persistent entropy, which has been previously employed to characterize heterogeneities within hydrogen bonding networks of aqueous electrolytes[52]:

$$PE = \sum_i -\frac{l_i}{L} \log \frac{l_i}{L}, \quad (4)$$

where L is the cumulative lifetime of all bars.[53, 54] A high value of PE reflects large variations in birth and death times.[55] Persistent entropies were calculated for both 0- and 1-dimensional homology features and compared for both the equilibrated and stochastic distributions. The *excess persistent entropy* is computed as the difference between the entropies of the equilibrated and the stochastic surfaces.

Nonideality Index of the Instantaneous Surface. It is desirable to have an index that directly reflects TBP...TBP interactions, and as such, non-ideality of the surface. At any instant in time a single TBP adsorbed at the instantaneous surface is manifested as a cycle within the sub-levelset persistent homology barcode if adjacent empty pixels are connected (Figure 2). Provided there are no interactions between TBPs (meaning there are adjacent pixels with a value of 1), then the number of cycles is equivalent to the number of adsorbed TBP. In this case, we state that the homogeneous, noninteracting adsorbate surface is “ideal”. We define the non-ideality index I as:

$$I = \frac{\text{number of observed cycles}}{\text{average number of adsorbates}} \quad (5)$$

The temporal fluctuations of I , in particular, should yield significant insight into the dynamics of adsorbate organi-

zation. This is highly relevant to solute transport across the oil/water phase boundary, given that TBP forms self-assembled protrusion architectures that are key to the transport mechanism.

IV. Results and Discussion

Figure 3 presents representative adsorbate density images of TBP at the water/vapor and electrolyte/hexane interface. Although there are changes to the xy interface dimensions, the homogeneous nature of the density distribution at the water/vapor interface (Figure 3A) is very similar to a density image with a stochastic distribution (Figure 3B), while clear patches of TBP appear in the electrolyte/hexane interface (Figure 3C). Yet to date, very few quantitative measures exist to differentiate organizational features across lengthscale and to correlate organizational features to macroscopic interfacial properties. Further, distinct organizational patterns may be related to reactive processes at the interface, including transport.

A. Analysis of Persistence Barcodes

Sublevelset persistent homology of the images of adsorbate density provides a means by which to understand the distribution and variations in density maxima and minima in the image. This is to be differentiated from direct correlations with spatial distributions, of which there are several tools that already exist for computational chemists, as described in the Introduction. To develop chemical intuition regarding the information contained in the PH barcode information, we first analyze the differences in the Betti curves across different dimensions for the TBP laden $\text{LiNO}_3^{(aq)}$ /hexane interface. Significant variations in TBP organization and reactive transport properties have previously been reported as a function of different aqueous solution conditions.[18, 33] At an aqueous 1M LiNO_3 concentration c.a. ~ 80 TBP are adsorbed to the interface - yet their transport H_2O or LiNO_3 into the organic phase is very slow on the timescale of the MD simulations. However, under 5 M LiNO_3 the TBP concentration is reduced at the interface to only ~ 30 adsorbates yet the rate of solute transport through surface TBP protrusion assemblies is significantly faster.

Let us first consider the comparison of the adsorbate density images under non-transporting conditions (80 TBP) against an image representing a stochastic TBP distribution at the same concentration (Figure 4). The Betti curves of each image are presented as a function of the ρ filtration. For both images, the β_0 curves start very close to zero, which indicates that in each pixel there is a small probability of a TBP being adsorbed. In the stochastic distribution the number of β_0 components at the maximum in the Betti curve is much larger than in the image created from the equilibrated MD simulation, this means the density is spread out across many more pixels and with more random density values for each pixel. Given this, in the stochastic image barcode each pixel is less likely to be merged with another pixel at low filtration

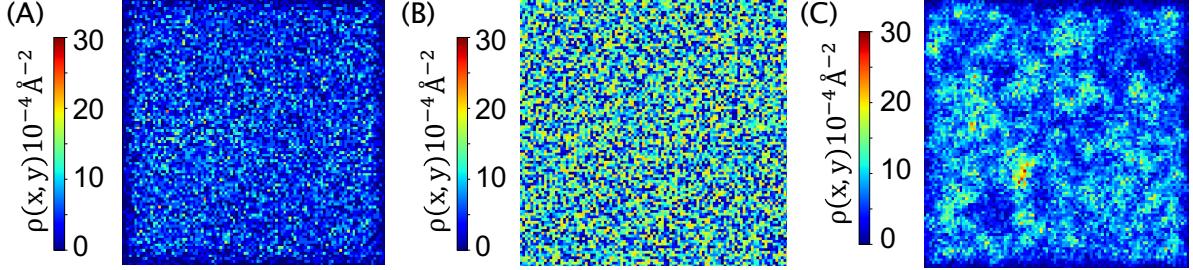


FIG. 3. Representative density surfaces analyzed in this study. (A) Adsorption density surface of ~ 72 TBP adsorbed at the water/vapor interface ($x = y = 60 \text{ \AA}$). (B) Stochastic distribution model surface representing ~ 72 TBP on a $x = y = 52 \text{ \AA}$ surface, (C), and an average of ~ 79 TBP adsorbed at $\text{LiNO}_3(\text{aq})/\text{hexane}$ interface ($x = y = 52 \text{ \AA}$).

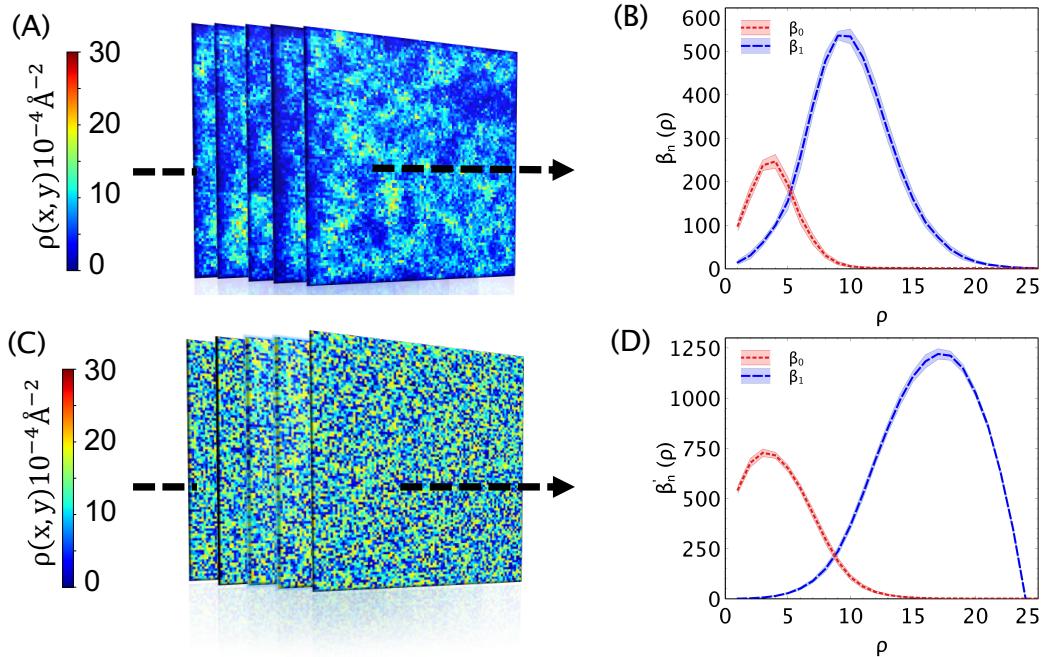


FIG. 4. (A) Time-averaged TBP adsorption surfaces obtained from equilibrated molecular dynamics simulations with ~ 80 TBP adsorbed at $\text{LiNO}_3(\text{aq})/\text{hexane}$ interface, and their respective (B) 0- and 1-d Betti curves. Analogous (C) stochastic adsorption surfaces with their respective (D) 0- and 1-d Betti curves. Error bars represent the standard deviation. A total of 10 surfaces are used to compute the standard deviation.

values and the density maxima are revealed at higher ρ filtrations. In the barcode of the equilibrated MD image, the TBP density is not randomly distributed and is instead influenced by TBP...TBP interactions. Patches of TBP density exist throughout the image, which cause groupings of merged pixels that have a ρ intensity \leq the filtration, and as such a smaller number of connected components are observed at low ρ . These merged β_0 components form cycles, β_1 , that show the distribution of density maxima surrounded by basins of the low-lying connected adsorbate density. Thus, there is a growth of β_1 distribution a lower ρ than in the stochastic image. This rationale is illustrated within Figure 5, which visually depicts the distribution of pixels as a function of the

surface organization in surface with no self-assembled TBP as would be expected in a stochastic distribution versus the image from the equilibrated simulation.

To better understand the maxima and minima in density within the images, we now consider the distribution of birth and death times for the β_0 and β_1 bars within the associated Betti curves. Figure 6 presents this data for the average density image of the non-transporting $n = 80$ and transporting $n = 30$ TBP $\text{LiNO}_3(\text{aq})/\text{hexane}$ interfaces. Every β_0 bar that is born represents a minima on the density surface that dies upon merger of the pixels that $\rho \leq$ the FV. If the density minima are connected through a maximum along a single dimension x or y in the grid, then when the FV overcomes the barrier the pix-

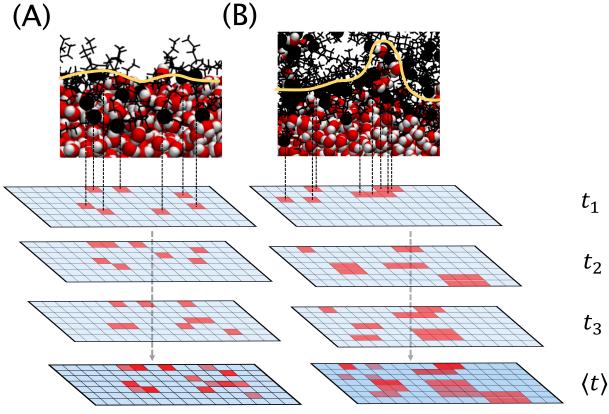


FIG. 5. An illustration of the merging of pixels at each filtration level of the time-averaged adsorption surfaces under non-transporting and solute transporting systems.

els merge without the formation of any cycles or β_1 components. However, if a density maxima is surrounded by a radial basin that surrounds it - then merging of the β_0 minima to form that basin results in the birth of a cycle or β_1 bar in the Betti curve. One might anticipate that a radial basin in density is somewhat uneven with its own small barriers in density that need to be overcome in the filtration before the basin forms. If this is the case, then there will be a significant number of deaths of β_0 bars before the basin is fully formed and the β_1 bar is born. If there are many inter-connected basins of density, then the death of a small number of β_0 bars may lead to large numbers of β_1 bars being born. Given this, we observe in the $n = 80$ TBP image that the β_1 bars are born nearly linearly with the death of β_0 bars between ρ values of 0 - 7, after which cycles grow at a much more rapid rate than the death of β_0 features. This indicates that there are many inter-connected basins on the density surface, but that there are many low-lying rough features on the density surface that must be overcome in the FV of ρ , up to a density value of ≈ 7 before the basins connect. These basins surround density maxima that are represented by the β_1 bars in the Betti curve. Interestingly, just as the number of β_1 bars are being born is rapidly increasing, there too is an increase in the death of β_1 bars - meaning that the density maxima are being overcome. Note that broad range of the distribution of the β_1 death times. This indicates that there is a large variation in the density maxima values on the density surface.

Comparing this behavior to the density image of $n = 30$ TBP at the $\text{LiNO}_3\text{(aq)}/\text{hexane}$ interface several important differentiating characteristics emerge. First, the very low TBP concentration on the surface means that the β_1 bars are born at the same time that the β_0 bars. This means there are many small patches of TBP density that may or may not form cycles but are surrounded by a surface that has many pixels of $\rho = 0$ (a huge basin of den-

sity). As the filtration in ρ is increased, the low TBP concentration immediately leads to the merger of the β_0 features because the intervening density maxima are easily overcome - yet all of these mergers lead to the formation of β_1 bars as indicated by the complete overlap of the β_0 death and β_1 birth distributions and the very large number of β_1 bars that are born for a relatively small number of β_0 bar deaths. Interestingly, one might anticipate that at such a low TBP adsorbate density that the TBPs would be rather homogeneously distributed on the surface and not interact with one another. However if this were the case then the deaths of the β_1 density maxima would occur rapidly and over a very short density range. Instead, it is clearly evident from the death distribution of the β_1 bars that there are large variations in the density maxima that are as pronounced as in the $n = 80$ surface condition. This has significant bearing upon the organization of the TBP in the $n = 30$ surface and their ability to transport solutes, as revealed by the PH descriptors developed below.

B. Trends of PH Descriptors with Experimental Observables

It is important that descriptors of the PH barcodes have good correlation with experimental observables. The water/vapor system is used as an example with an increasing number of adsorbed TBP; this system does not involve mass transport, and the variations in the descriptors will reflect only changes to the TBP surface coverage. The two-dimensional density distributions of TBP on the water/vapor surface with adsorbate concentrations from $n = 72$ to 144 TBP per interface and the resulting Betti curves and persistent entropy are presented in Figures S7 - S8. As anticipated, the 0-d Betti curves reflect an increase in the total number of persistent features representing the increase in the TBP density at the surface. The 1-d Betti curves do not show a significant increase in the number of cycles but illustrate an increase in the complexity of the TBP organization at the surface by larger persistence of bars and the shift of the maxima to a greater filtration value.

As a starting point, we first consider the correlations of the Betti index and persistent entropy with trends in macroscopic (experimentally measurable) interfacial properties that depend upon adsorbate concentration, such as interfacial tension. The respective persistent entropies $\text{PE}(\beta_n)$ and Betti-index $h(\beta_n)$ also increase as a result of variations in the adsorbate density/concentration. More importantly, as shown in Figure 7, these sublevelset PH descriptors exhibit strong inverse correlations with the macroscopic predicted interfacial tension obtained from the simulation data using the pressure tensor method.[56] The strength of the negative correlation between interfacial tension and PE (as represented by the slope m fit by linear regression) is quantified to be $> 10^3$ times than h , showing a greater sensitivity of PE over the Betti-index.

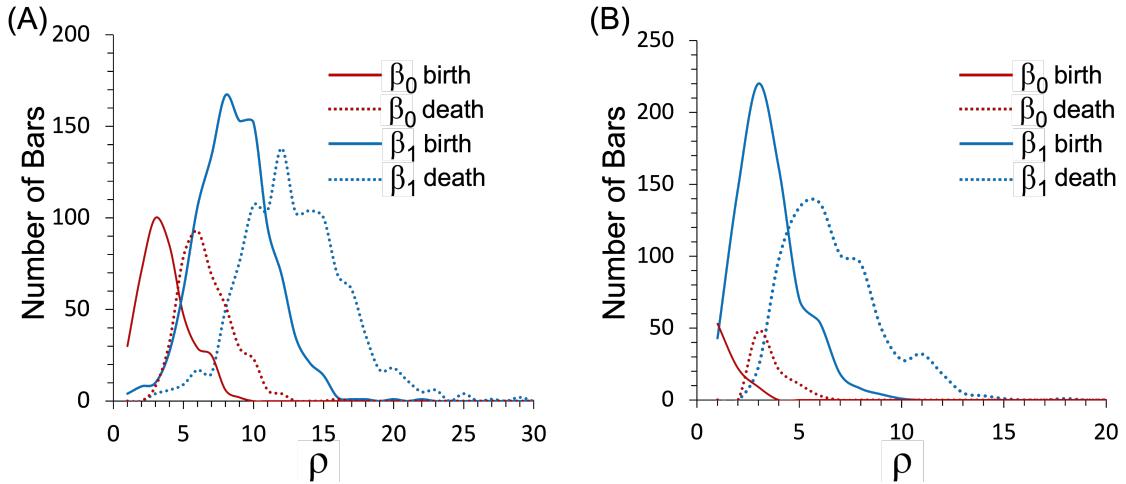


FIG. 6. (A) Distribution of birth and death times for β_0 and β_1 bars within the analogous Betti curves for the density image of the $n = 80$ TBP laden $\text{LiNO}_3\text{(aq)}$ /hexane interface (non-transporting conditions). (B) Distribution of birth and death times for β_0 and β_1 bars within the analogous Betti curves for the density image of the $n = 30$ TBP laden $\text{LiNO}_3\text{(aq)}$ /hexane interface (transporting conditions).

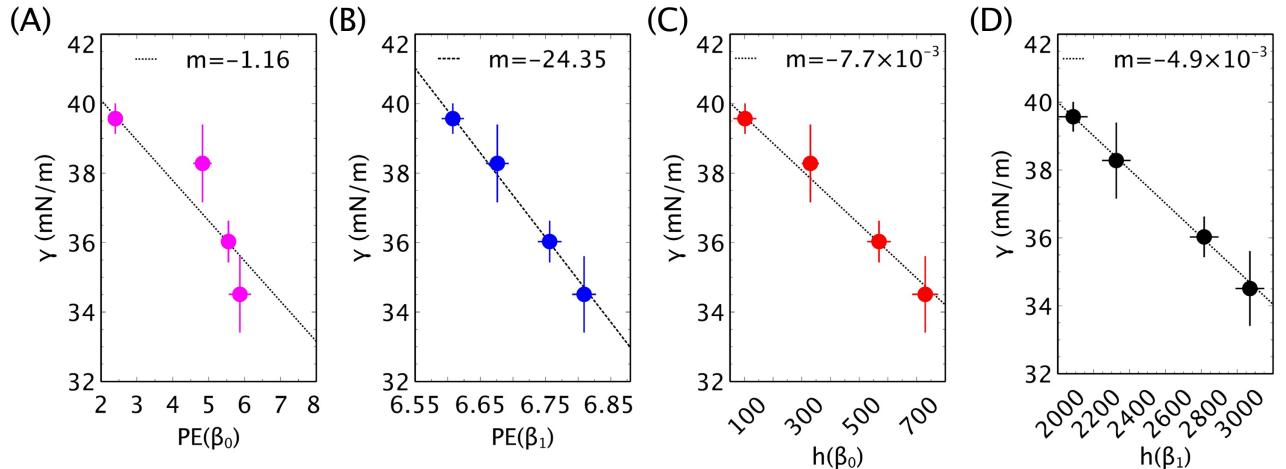


FIG. 7. The correlation between the interfacial tension (γ) with (A) 0- and (B) 1-d persistent entropy and (C) 0- and (D) 1-d Betti-index for the systems with 72, 96, 120 and 144 TBP per interface at the water/vapor surface. The data is fitted to a linear regression model to obtain the slopes provided in the inset.

C. Descriptors of Adsorbate Organization

Unlike the relatively homogeneous distribution of the TBP monolayer at the water/vapor interface, the adsorption at a non-polar liquid interface involves TBP...TBP interactions that may form self-assembled species that are manifested as the density maxima in the density images as described in Section IV A. Although we have demonstrated that the Betti index and PE are well-correlated with the adsorbate concentration, understanding how PH descriptors can connect the variation in the density maxima and minima of the image to chemical reactivity is essential. The PH descriptor behavior is examined

for images as a function of TBP concentration where we delineate the non-transporting $n = 80$ TBP adsorbate $\text{LiNO}_3\text{(aq)}$ /hexane density image from the transporting $n = 30$ image. Although this change to the surface transport activity is anticipated to be related to the surface TBP heterogeneity, this is not readily apparent from visual inspection of the density images (Figure S9) due to the large change to adsorbate TBP concentration. One might intuit that the adsorbate organization heterogeneity should be less in the 80 TBP adsorbed conditions (non-transporting) relative to the 30 adsorbed TBP conditions (transporting). Given the concentration sensitivity of the PH barcodes, we employ excess PH descriptors

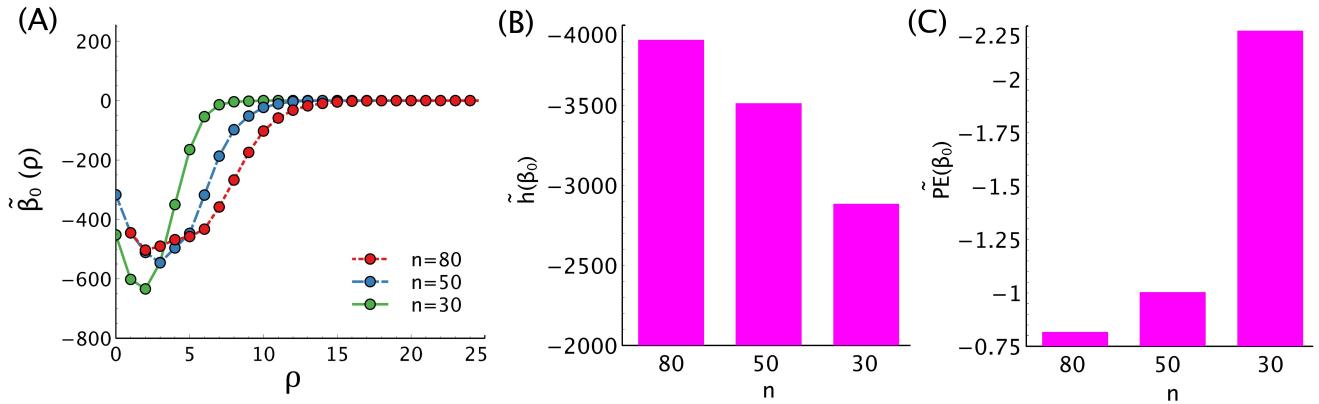


FIG. 8. (A) Excess Betti curve of 0-dimensional components, $\beta_0(\rho) - \beta'_0(\rho) = \tilde{\beta}_0(\rho)$. (B) Excess Betti-index of 0-dimensional components, $h(\beta_0) - h'(\beta_0) = \tilde{h}_0(\rho)$. (C) Excess persistent entropy $PE(\beta_0) - PE'(\beta_0) = \tilde{PE}(\beta_0(\rho))$.

based upon the difference of the PH descriptor of the images from the equilibrated MD simulation versus an image that represents a stochastic ideal distribution.

The excess Betti curve (denoted by $\tilde{\beta}_0(\rho)$) is presented in Figure 8A as a function of TBP adsorbate concentration on the $\text{LiNO}_3\text{(aq)}/\text{hexane}$ interface. Here, the excess $\tilde{\beta}_0(\rho)$ is negative across the entire filtration range. This means that at any given filtration ρ value there are more connected components in the image representing a stochastic TBP distribution than in the density image of the equilibrated liquid/liquid interface (as would be implied from Figure 4). The convergence of $\tilde{\beta}_0(\rho)$ toward zero is correlated to the pixel intensities in each image (meaning the value of ρ). Thus, the values of the ρ filtration at which $\tilde{\beta}_0(\rho)$ approaches zero are directly related to TBP concentration. Examining the concentration dependence of $\tilde{\beta}_0(\rho)$, there is a clear differentiation between the surface density images of under transporting conditions (30 TBP adsorbates) versus non-translating (80 TBP adsorbates). The $\tilde{\beta}_0(\rho)$ for the translating surface is much more negative than in the translating case, meaning that the variations in surface density (heterogeneity of the adsorption) is significantly enhanced under translating conditions. In the case of the $\tilde{\beta}_1(\rho)$ curves (Figure S10), at low ρ filtration, the values of $\tilde{\beta}_1$ are positive because the density images from the equilibrated MD simulation have TBP-TBP interactions that cause the formation of cycles β_1 at the expense of the number of β_0 components. The maximum values of $\tilde{\beta}_1$ are 370, 360, and 329 for 30, 50, and 80 adsorbed TBP, respectively. As the ρ filtration value is increased a negative excess is observed that is consistent with the fact that in the stochastic image the cycles form at larger ρ values due to the homogeneous distribution of on the surface which is manifested in the density image.

In summary, the excess $\tilde{\beta}_n(\rho)$ curves (including the values of $\tilde{\beta}_n$ at the peak maxima and/or minima) indicate an ability to differentiate organizational differences

in TBD adsorbate that are manifested in the density images and correlate well adsorbate reactivity (i.e., self-assembled transporting architectures). By comparing the β_n values between the density images of the equilibrated MD simulation and an analogous stochastic adsorbate distribution, the adsorbate concentration dependence is largely removed and the significant variations adsorbate organization that lead to large variations in surface density can be compared across different chemical systems.

Integrating the Betti and excess Betti curves overall ρ filtration values (Figure S11) yields the Betti index and excess Betti index. The cumulative number of bars within the adsorption surface is highly dependent on lifetimes of the components, i.e., how quickly components merge as a function of the ρ filtration. Given that the convergence of the excess Betti curves in Figure 8A is directly related to the maximum pixel intensity (max ρ value), the integration causes the excess $\tilde{\beta}_n$ to reflect solely variations in concentration of the adsorbate on the surface (Figure 8B). Much more insight into the variations of adsorbate organization is obtained when examining the distribution of β_0 and β_1 lifetimes as a function of TBP concentration and in comparing those distributions for the images from the equilibrated MD data and analogous stochastic adsorbate images. This is obtained from the persistent entropies (PE) and excess persistent entropies (\tilde{PE}). Changes to the excess persistent entropies are attributed to variations in the bar lifetimes in the images from the equilibrated MD surfaces. As observed in Figure 8C, the excess PE for the translating system is dramatically enhanced relative to the non-translating system. This indicates significant variations in density maxima and minima on the surface under translating conditions and reflects large variations in self-assembled translating protrusions therein. This agrees well with the analysis of surface protrusions in this system using geometric measure theory.[57]

D. Dynamic Evolution of Adsorbate Organization at the Instantaneous Surface

To study the dynamic behavior of the TBP organization at the surface, one must analyze the persistent homology features of images created at each step in time. Recall that in this instance, all pixels in the density image without an adsorbed TBP are connected components at a filtration value of $\rho = 0$, and the presence of a TBP creates a cycle (β_1 bar) at $\rho = 1$ in the density image. For an ideal case, where each TBP is isolated and does not interact to form any TBP self-assembled surface species (ideal conditions), the total number of TBP will be equivalent to the total number of cycles. However, as TBPs interact and form self-assembled surface protrusions, the number of cycles will vary significantly from the number of adsorbed TBP. This is visually depicted in Figure 5 and forms the basis of the definition of the non-ideality index $I(\beta_1)$ in Equation 5, where a value of 1 indicates ideality and the smaller the value becomes, the less ideal the TBP interactions. As shown in Figure S15 the $I(\beta_1)$ values for all TBP laden water/vapor interfaces have values quite close to 1, indicating their inability to form self-assembled species that can lead to solute transport. However, in the $\text{LiNO}_3\text{(aq)}/\text{hexane}$ interface

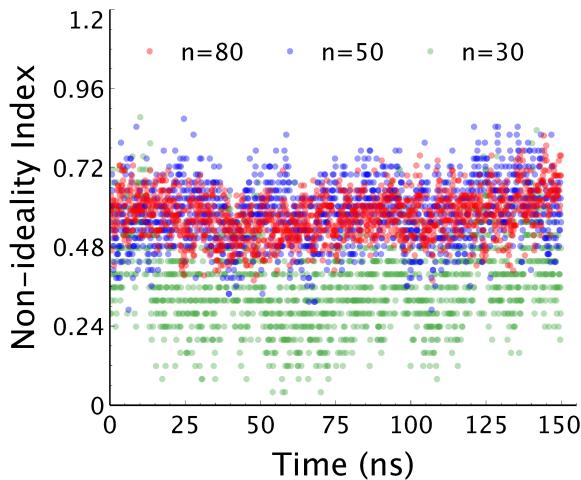


FIG. 9. Non-ideality index $I(\beta_1)$ as a function of time for the instantaneous adsorption of n TBP at the $\text{LiNO}_3\text{(aq)}/\text{hexane}$ interface.

face the enhanced surface roughness coupled to transport phenomena dramatically reduce the average $I(\beta_1)$ value, irrespective of the TBP concentration. The highest (most ideal) $I(\beta_1)$ occurs in the non-transporting $n = 80$ TBP surface which has a value of 0.578, indicating significant TBP-TBP interactions (Figure 9). Yet this is still much larger than the non-ideality index for the transporting $n = 30$ TBP system. This indicates significant self-assembly of the TBP's available on the surface despite a much lower TBP surface density. Alternatively

put, and in agreement with the observations made regarding the birth and death time distributions of β_0 and β_1 , the non-transporting $n = 80$ TBP system has a base TBP adsorbate density that is non-interacting but with many groupings of self-assembled TBP. In contrast, for the $n = 30$ TBP system a much larger fraction of the adsorbed TBP form self-assemblies. Perhaps more interesting from this analysis is the magnitude of fluctuations in $I(\beta_1)$ as a function of time. The standard deviation in the $I(\beta_1)$ for the non-transporting $n = 80$ case is nearly half that of the solute transporting $n = 30$ system, which is associated with a high rate of adsorption/desorption in the solute transporting system as well as surface reorganization of the adsorbates upon solute transport.

V. Summary

Adsorbate density images from molecular dynamics simulations of liquid interfaces have been proposed as a means to understand interfacial organization and reactivity. This data representation allows us to use sublevelset persistent homology barcodes to encode variations in the adsorbate density within the image. Although chemists often think about adsorbate organization in the context of spatial distribution, the density maxima and minima clearly reflect heterogeneity in the surface organization and “hot spots” of chemical reactivity caused by adsorbate...adsorbate interactions. Persistent homology descriptors like persistent entropy and the newly developed non-ideality index describe variations in the topology of the adsorbate density and are highly sensitive to both concentration and surface organization. They are shown to trend well with experimentally measurable quantities like interfacial tension. We explicitly consider the amphiphile TBP at the $\text{LiNO}_3\text{(aq)}/\text{hexane}$ interface under solution conditions where the amphiphiles self-assemble yet do not readily transport solutes (on the timescale of the simulation), or where they self-assemble and have fast kinetics of solute transport. The PH descriptors easily differentiate between these two surface behaviors both from ensemble average data as well as temporal analysis of the fluctuations in the descriptors. The increasing accessibility of static and time-resolved surface images of adsorbates (through microscopy and simulation) and the computational expedience and sensitivity of sublevelset persistent homology indicates that the descriptors and workflow proposed in this work will be of significant benefit to the interfacial chemistry community.

Acknowledgments

The authors acknowledge the Department of Energy, Basic Energy Sciences Separations program (DE-SC0001815) for funding. This research used resources from the Center for Institutional Research Computing at Washington State University.

-
- [1] O. Bjorneholm, M. H. Hansen, A. Hodgson, L.-M. Liu, D. T. Limmer, A. Michaelides, P. Pedevilla, J. Rossmeisl, H. Shen, G. Tocci, *et al.*, Water at interfaces, *Chemical reviews* **116**, 7698 (2016).
- [2] W. Kunz, *Specific ion effects* (World Scientific, 2010).
- [3] S. W. Devlin, I. Benjamin, and R. J. Saykally, On the mechanisms of ion adsorption to aqueous interfaces: air-water vs. oil-water, *Proceedings of the National Academy of Sciences* **119**, e2210857119 (2022).
- [4] Y. Shen, A few selected applications of surface nonlinear optical spectroscopy, *Proceedings of the National Academy of Sciences* **93**, 12104 (1996).
- [5] P. S. Pershan and M. Schlossman, *Liquid surfaces and interfaces: synchrotron x-ray methods* (Cambridge University Press, 2012).
- [6] U. I. Premadasa, Y.-Z. Ma, R. L. Sacci, V. Bocharova, N. A. Thiele, and B. Doughty, Understanding self-assembly and the stabilization of liquid/liquid interfaces: The importance of ligand tail branching and oil-phase solvation, *Journal of Colloid and Interface Science* **609**, 807 (2022).
- [7] J. E. Sutton, S. Roy, A. U. Chowdhury, L. Wu, A. K. Wan-hala, N. De Silva, S. Jansone-Popova, B. P. Hay, M. C. Cheshire, T. L. Windus, *et al.*, Molecular recognition at mineral interfaces: implications for the beneficiation of rare earth ores, *ACS applied materials & interfaces* **12**, 16327 (2020).
- [8] M. Ahmed, M. Blum, E. J. Crumlin, P. L. Geissler, T. Head-Gordon, D. T. Limmer, K. K. Mandadapu, R. J. Saykally, and K. R. Wilson, Molecular properties and chemical transformations near interfaces, *The Journal of Physical Chemistry B* **125**, 9037 (2021).
- [9] S. Yoo, B. Qiao, T. Douglas, W. Bu, M. Olvera de la Cruz, and P. Dutta, Specific ion effects in lanthanide-amphiphile structures at the air–water interface and their implications for selective separation, *ACS Applied Materials & Interfaces* **14**, 7504 (2022).
- [10] I. Benjamin, Chemical reactions and solvation at liquid interfaces: A microscopic perspective, *Chemical reviews* **96**, 1449 (1996).
- [11] J. S. Rowlinson and B. Widom, *Molecular theory of capillarity* (Courier Corporation, 2013).
- [12] A. P. Willard and D. Chandler, Instantaneous liquid interfaces, *The Journal of Physical Chemistry B* **114**, 1954 (2010).
- [13] M. Segá, G. Hantal, B. Fábián, and P. Jedlovszky, Pytim: A python package for the interfacial analysis of molecular simulations (2018).
- [14] M. Segá, S. S. Kantorovich, P. Jedlovszky, and M. Jorge, The generalized identification of truly interfacial molecules (itim) algorithm for nonplanar interfaces, *The Journal of chemical physics* **138**, 044110 (2013).
- [15] D. M. Mitrinovic, Z. Zhang, S. M. Williams, Z. Huang, and M. L. Schlossman, X-ray reflectivity study of the water–hexane interface, *The Journal of Physical Chemistry B* **103**, 1779 (1999).
- [16] N. Kumar, M. J. Servis, Z. Liu, and A. E. Clark, Competitive interactions at electrolyte/octanol interfaces: a molecular perspective, *The Journal of Physical Chemistry C* **124**, 10924 (2020).
- [17] Z. Liu and A. E. Clark, An octanol hinge opens the door to water transport, *Chemical science* **12**, 2294 (2021).
- [18] N. Kumar and A. E. Clark, Unexpected inverse correlations and cooperativity in ion-pair phase transfer, *Chemical science* **12**, 13930 (2021).
- [19] W. Rock, B. Qiao, T. Zhou, A. E. Clark, and A. Uysal, Heavy anionic complex creates a unique water structure at a soft charged interface, *The Journal of Physical Chemistry C* **122**, 29228 (2018).
- [20] S. Nayak, R. R. Kumal, Z. Liu, B. Qiao, A. E. Clark, and A. Uysal, Origins of clustering of metalate-extractant complexes in liquid–liquid extraction, *ACS Applied Materials & Interfaces* **13**, 24194 (2021).
- [21] J. Townsend, C. P. Micucci, J. H. Hymel, V. Maroulas, and K. D. Vogiatzis, Representation of molecular structures with persistent homology for machine learning applications in chemistry, *Nature communications* **11**, 1 (2020).
- [22] Y. Shimizu, T. Kurokawa, H. Arai, and H. Washizu, Higher-order structure of polymer melt described by persistent homology, *Scientific reports* **11**, 1 (2021).
- [23] Y. Lee, S. D. Barthel, P. Dłotko, S. M. Moosavi, K. Hess, and B. Smit, Quantifying similarity of pore-geometry in nanoporous materials, *Nature communications* **8**, 1 (2017).
- [24] H. Adams and M. Moy, Topology applied to machine learning: From global to local, *Frontiers in Artificial Intelligence* **4**, 54 (2021).
- [25] P. Dłotko and T. Wanner, Topological microstructure analysis using persistence landscapes, *Physica D: Nonlinear Phenomena* **334**, 60 (2016).
- [26] I. Membrillo Solis, T. Orlova, K. Bednarska, P. Lesiak, T. R. Woliński, G. D'Alessandro, J. Brodzki, and M. Kaczmarek, Tracking the time evolution of soft matter systems via topological structural heterogeneity, *Communications Materials* **3**, 1 (2022).
- [27] S. A. Jassim, Persistent homology features and multiple topologies for image analysis, in *Mobile Multimedia/Image Processing, Security, and Applications 2020*, Vol. 11399 (SPIE, 2020) pp. 53–64.
- [28] A. Hofmann, M. Krufczik, D. W. Heermann, and M. Hausmann, Using persistent homology as a new approach for super-resolution localization microscopy data analysis and classification of $\gamma\text{h}2\text{ax}$ foci/clusters, *International journal of molecular sciences* **19**, 2263 (2018).
- [29] R. Ghrist, Barcodes: the persistent topology of data, *Bulletin of the American Mathematical Society* **45**, 61 (2008).
- [30] O. Vipond, J. A. Bull, P. S. Macklin, U. Tillmann, C. W. Pugh, H. M. Byrne, and H. A. Harrington, Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors, *Proceedings of the National Academy of Sciences* **118**, e2102166118 (2021).
- [31] A. Smith, S. Runde, A. Chew, A. Kelkar, U. Maheshwari, R. Van Lehn, and V. Zavala, Topological analysis of molecular dynamics simulations using the euler characteristic, (2022).
- [32] M. Baaden, M. Burgard, and G. Wipff, Tbp at the water–oil interface: the effect of tbp concentration and water acidity investigated by molecular dynamics simulations, *The Journal of Physical Chemistry B* **105**, 11131 (2001).
- [33] M. J. Servis and A. E. Clark, Surfactant-enhanced heterogeneity of the aqueous interface drives water extraction into organic solvents, *Physical Chemistry Chemical*

- Physics **21**, 2866 (2019).
- [34] X. Ye, S. Cui, V. F. de Almeida, B. P. Hay, and B. Khomami, Uranyl nitrate complex extraction into tbp/dodecane organic solutions: a molecular dynamics study, Physical Chemistry Chemical Physics **12**, 15406 (2010).
- [35] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX **1**, 19 (2015).
- [36] W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Physical review A **31**, 1695 (1985).
- [37] M. Parrinello and A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, Journal of Applied physics **52**, 7182 (1981).
- [38] N. Kumar, M. J. Servis, and A. E. Clark, Uranyl speciation in the presence of specific ion gradients at the electrolyte/organic interface, Solvent Extraction and Ion Exchange **40**, 165 (2022).
- [39] X. Ye, S. Cui, V. F. de Almeida, and B. Khomami, Molecular simulation of water extraction into a tri-n-butylphosphate/n-dodecane solution, The Journal of Physical Chemistry B **117**, 14835 (2013).
- [40] T. Darden, D. York, and L. Pedersen, Particle mesh ewald: An n log (n) method for ewald sums in large systems, The Journal of chemical physics **98**, 10089 (1993).
- [41] B. Hess, H. Bekker, H. J. Berendsen, and J. G. Fraaije, Lincs: a linear constraint solver for molecular simulations, Journal of computational chemistry **18**, 1463 (1997).
- [42] T. A. Halgren, The representation of van der waals (vdw) interactions in molecular mechanics force fields: potential form, combination rules, and vdw parameters, Journal of the American Chemical Society **114**, 7827 (1992).
- [43] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, Array programming with NumPy, Nature **585**, 357 (2020).
- [44] C. Maria, J.-D. Boissonnat, M. Glisse, and M. Yvinec, The gudhi library: Simplicial complexes and persistent homology, in *Mathematical Software – ICMS 2014*, edited by H. Hong and C. Yap (Springer Berlin Heidelberg, Berlin, Heidelberg, 2014) pp. 167–174.
- [45] N. Kumar and A. E. Clark, Adsorbate-ph (2023), <https://gitlab.com/aurora-clark-public/adsorbate-ph>.
- [46] H. Wagner, C. Chen, and E. Vuçini, Efficient computation of persistent homology for cubical data, in *Topological methods in data analysis and visualization II* (Springer, 2012) pp. 91–106.
- [47] C. Chen and M. Kerber, Persistent homology computation with a twist, in *Proceedings 27th European Workshop on Computational Geometry*, Vol. 11 (2011) pp. 197–200.
- [48] A. Cerri, B. D. Fabio, M. Ferri, P. Frosini, and C. Landi, Betti numbers in multidimensional persistent homology are stable functions, Mathematical Methods in the Applied Sciences **36**, 1543 (2013).
- [49] D. V. Anand, Z. Meng, K. Xia, and Y. Mu, Weighted persistent homology for osmolyte molecular aggregation and hydrogen-bonding network analysis, Scientific reports **10**, 1 (2020).
- [50] C. De Boor and C. De Boor, *A practical guide to splines*, Vol. 27 (springer-verlag New York, 1978).
- [51] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods **17**, 261 (2020).
- [52] K. Xia, Persistent homology analysis of ion aggregations and hydrogen-bonding networks, Physical Chemistry Chemical Physics **20**, 13448 (2018).
- [53] H. Chintakunta, T. Gentimis, R. Gonzalez-Diaz, M.-J. Jimenez, and H. Krim, An entropy-based persistence barcode, Pattern Recognition **48**, 391 (2015).
- [54] E. Merelli, M. Rucco, P. Sloot, and L. Tesei, Topological characterization of complex systems: Using persistent entropy, Entropy **17**, 6872 (2015).
- [55] M. Rucco, F. Castiglione, E. Merelli, and M. Pettini, Characterisation of the idiotypic immune network through persistent entropy, in *Proceedings of ECCS 2014* (Springer, 2016) pp. 117–128.
- [56] J. Alejandre, D. J. Tildesley, and G. A. Chapela, Molecular dynamics simulation of the orthobaric densities and surface tension of water, The Journal of chemical physics **102**, 4574 (1995).
- [57] E. Alvarado, Z. Liu, M. J. Servis, B. Krishnamoorthy, and A. E. Clark, A geometric measure theory approach to identify complex structural features on soft matter surfaces, Journal of Chemical Theory and Computation (2020).
- [58] F. Sedlmeier, Y. von Hansen, L. Mengyu, D. Horinek, and R. R. Netz, Water dynamics at interfaces and solutes: disentangling free energy and diffusivity contributions, Journal of Statistical Physics **145**, 240 (2011).
- [59] H. Masoomy, B. Askari, S. Tajik, A. K. Rizi, and G. R. Jafari, Topological analysis of interaction patterns in cancer-specific gene regulatory network: persistent homology approach, Scientific Reports **11**, 1 (2021).
- [60] A. Ulman, Formation and structure of self-assembled monolayers, Chemical reviews **96**, 1533 (1996).
- [61] M. Kimura, I. Obayashi, Y. Takeichi, R. Murao, and Y. Hiraoka, Non-empirical identification of trigger sites in heterogeneous processes using persistent homology, Scientific reports **8**, 1 (2018).
- [62] H. Edelsbrunner, D. Letscher, and A. Zomorodian, Topological persistence and simplification, in *Proceedings 41st annual symposium on foundations of computer science* (IEEE, 2000) pp. 454–463.
- [63] S. Berg, D. Kutra, T. Kroeger, C. N. Straehle, B. X. Kausler, C. Haubold, M. Schiegg, J. Ales, T. Beier, M. Rudy, et al., Ilastik: interactive machine learning for (bio) image analysis, Nature Methods **16**, 1226 (2019).
- [64] J. T. Nardini, B. J. Stoltz, K. B. Flores, H. A. Harrington, and H. M. Byrne, Topological data analysis distinguishes parameter regimes in the anderson-chaplain model of angiogenesis, PLOS Computational Biology **17**, 1 (2021).
- [65] P. Jungwirth and D. J. Tobias, Ions at the air/water interface, The Journal of Physical Chemistry B **106**, 6361 (2002).

- [66] Z. Liang, W. Bu, K. J. Schweighofer, D. J. Walwark, J. S. Harvey, G. R. Hanlon, D. Amoanu, C. Erol, I. Benjamin, and M. L. Schlossman, Nanoscale view of assisted ion transport across the liquid–liquid interface, *Proceedings of the National Academy of Sciences* **116**, 18227 (2019).
- [67] K. Xia, X. Feng, Y. Tong, and G. W. Wei, Persistent homology for the quantitative prediction of fullerene stability, *Journal of computational chemistry* **36**, 408 (2015).
- [68] M. Li, M. H. Frank, V. Coneva, W. Mio, D. H. Chitwood, and C. N. Topp, The persistent homology mathematical framework provides enhanced genotype-to-phenotype associations for plant morphology, *Plant physiology* **177**, 1382 (2018).
- [69] Z. Meng, D. V. Anand, Y. Lu, J. Wu, and K. Xia, Weighted persistent homology for biomolecular data analysis, *Scientific reports* **10**, 1 (2020).
- [70] Y. Onodera, S. Kohara, P. S. Salmon, A. Hirata, N. Nishiyama, S. Kitani, A. Zeidler, M. Shiga, A. Masuno, H. Inoue, *et al.*, Structure and properties of densified silica glass: characterizing the order within disorder, *NPG Asia Materials* **12**, 1 (2020).
- [71] J.-P. Serre, Homologie singulière des espaces fibrés, *Annals of Mathematics* , 425 (1951).
- [72] I. Benjamin, Molecular structure and dynamics at liquid–liquid interfaces, *Annual review of physical chemistry* **48**, 407 (1997).
- [73] Z. Meng and K. Xia, Persistent spectral-based machine learning (perspect ml) for protein-ligand binding affinity prediction, *Science Advances* **7**, eabc5329 (2021).
- [74] A. Takiyama, T. Teramoto, H. Suzuki, K. Yamashiro, and S. Tanaka, Persistent homology index as a robust quantitative measure of immunohistochemical scoring, *Scientific reports* **7**, 1 (2017).
- [75] K. Xia and G.-W. Wei, Persistent homology analysis of protein structure, flexibility, and folding, *International journal for numerical methods in biomedical engineering* **30**, 814 (2014).
- [76] J. Mirth, Y. Zhai, J. Bush, E. G. Alvarado, H. Jordan, M. Heim, B. Krishnamoorthy, M. Pflaum, A. Clark, and H. Adams, Representations of energy landscapes by sublevelset persistent homology: An example with n-alkanes, *The Journal of Chemical Physics* **154**, 114114 (2021).
- [77] X. Lin, J. H. Lorent, A. D. Skinkle, K. R. Levental, M. N. Waxham, A. A. Gorfe, and I. Levental, Domain stability in biomimetic membranes driven by lipid polyunsaturation, *The Journal of Physical Chemistry B* **120**, 11930 (2016).
- [78] G. Spreemann, B. Dunn, M. B. Botnan, and N. A. Baas, Using persistent homology to reveal hidden covariates in systems governed by the kinetic ising model, *Physical Review E* **97**, 032313 (2018).
- [79] Y. Hu, P. Ounkham, O. Marsalek, T. E. Markland, B. Krishnamoorthy, and A. E. Clark, Persistent homology metrics reveal quantum fluctuations and reactive atoms in path integral dynamics, *Frontiers in chemistry* **9**, 57 (2021).
- [80] T. Qaiser, Y.-W. Tsang, D. Taniyama, N. Sakamoto, K. Nakane, D. Epstein, and N. Rajpoot, Fast and accurate tumor segmentation of histology images using persistent homology and deep convolutional features, *Medical image analysis* **55**, 1 (2019).

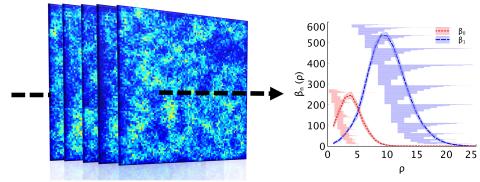


FIG. 10. This is a TOC image