JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

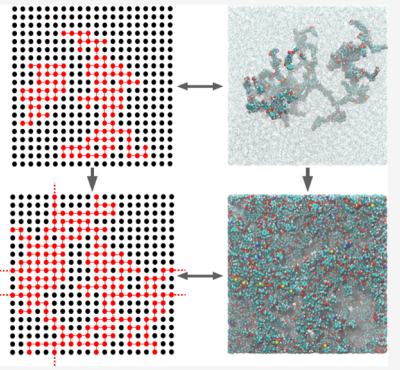
# An Exact Algorithm to Detect the Percolation Transition in Molecular Dynamics Simulations of Cross-Linking Polymer Networks

Mattia Livraghi,[§] Kevin Höllring,[§] Christian R. Wick, David M. Smith, and Ana-Sunčana Smith*

Cite This: https://doi.org/10.1021/acs.jctc.1c00423 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Periodic molecular dynamics simulations are developing to a routine tool for the investigation of complex, polymeric materials. A typical application is the simulation of the curing reaction of covalently cross-linked polymers, which provides detailed understanding of network formation at the molecular scale, with examples including gelation and glass transitions. In this article, we delineate the connection between percolation theory and gel-point detection in periodic polymeric networks. Specifically, we present an algorithm that can detect the onset of percolation during cross-linking of polymers in periodic molecular dynamic simulations. A sample implementation is provided at https://github.com/puls-group/percolation-analyzer. As an example, we apply the algorithm to simulations of an epoxy resin undergoing curing with an amine hardener. We also compare results with indirect gel point measurements obtained from monitoring the growth of the largest mass and the onset of secondary cycles.





## INTRODUCTION

In recent years, the technological importance of cross-linked thermoset materials has been growing, with examples including low-cure-temperature composites for the aerospace industry,[1] molecularly imprinted polymers,[2] rubbery ion electronic components,[3] and drug delivery hydrogels.[4] The synthesis of a covalently cross-linked thermoset polymer affords the so-called curing reaction, which includes a very specific reaction extent known as the gel point. This is typically characterized by sharp changes in both soluble fraction and viscosity,[5−7] caused by the transition from a viscous liquid of single molecular entities to an insoluble covalently cross-linked solid. Recently, the available experimental methods to analyze this transition have been extensively reviewed.[8] The importance of gelation for several industrial processes is also well-known.[9] For example, many molding techniques for composite materials rely on the injection of liquid precursors; since the shape of the composite is irreversibly set after the occurrence of gelation, knowledge of the time window preceding the gel point is essential. Other manufacturing processes for which gelation is highly relevant include foaming, the design of self-repairing materials, and 3D printing. From a microscopic point of view, the Flory−Stockmayer[10,11] theory of polymer growth explains the gel transition by the appearance of a gel-like molecule of macroscopic extent, which spans the whole volume of the mixture and whose mass can be measured in macroscopic units, thus looking "infinite" from the monomer-weight scale.[5]

Lately, Molecular Dynamics (MD) simulations have proven to be an effective tool to investigate the curing reaction for thermoset polymers, paving the way for a more detailed understanding of the gelation process at the molecular level. In MD simulations, gelation is typically determined indirectly by seeking a sharp change in some physical or structural properties of the polymerizing system, which is supposed to signal the appearance of a volume-spanning, mass-dominant macromolecule. Thus, a common indirect mass-based measurement consists of determining the point at which the heaviest molecular group starts to significantly outweigh its runner-up. To this end, both a visual estimate[12] and the point of inflection of the largest-mass buildup[13] have been employed. A more analytical mass-based criterion is represented by the reduced molecular weight (RMW), which is defined as the molecular weight average of all the molecules in the system, except the largest one. The RMW rises until the largest group begins to predominate, after which it declines. On these grounds, its maximum point has been used to measure gelation.[14,15]

From a more structural point of view, the gel-like molecular group naturally englobes most of the still unreacted functionalities present in the system, thus marking the rise of intramolecular reactions (also known as secondary cycles). As a consequence, the inception of intramolecular polymerization is the other common indirect property employed for gel point detection.[16]

Received: April 30, 2021

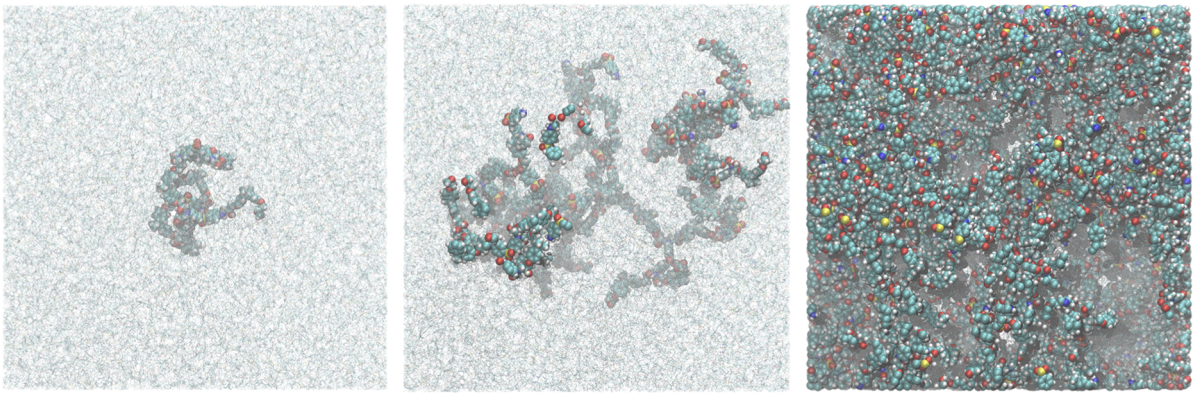



**Figure 1.** Visualization of the largest molecular group in one of our example epoxy systems (frontal view). The system consists of a periodic box containing 2000 DGEBA and 1000 DDS molecules (127 000 atoms). The shape of the box is approximately cubic, as a result of isotropic pressure control, with a side length of about 12 nm. Only the largest, reacted molecular group is portrayed as balls and sticks, while everything else is shown as wireframes. (a) corresponds to 20% curing, (b) to 50%, and (c) to 63%. (c) coincides with the point of percolation: the appearance of a single, spanning, highly massive macromolecule is evident.

Indirect methods, however, face challenges when applied to sufficiently complex polymer networks. In particular, they have limited precision, because both the emergence of a single, highly massive group and the inception of secondary cycles can only be estimated up to a range of curing extents. Also, they have limited accuracy, since the appearance of the gel-like macromolecule is related to, but not necessarily sharply marked by, a significant increase in either the mass or the number of intramolecular reactions of a group. Consequently, indirect methods are more likely to demonstrate larger uncertainties with increasing the system size, albeit with larger simulations finite size effects should diminish. Hence, one may not be able to harness the benefit of increased computational effort.

To circumvent these limitations, we here introduce a new strategy for gel point detection in molecular dynamics simulations of cross-linking materials. Instead of looking for changes in some sentinel properties, our strategy is to directly recognize the appearance of a molecular group spanning the entire simulation box. To define the problem mathematically, we identify such a spanning macromolecule with a so-called percolating cluster, exploiting the natural relation existing between gelation in a polymerizing material and percolation in a lattice of infinite size. Because atomistic simulations are necessarily of finite and microscopic size, our definition of percolation must rely on periodic boundary conditions to simulate an infinite system by means of periodicity. Then, we present an algorithm capable of detecting a percolating molecular cluster in one, two, and three dimensions, in a manner that is invariant under linear transformations of atomic coordinates. We argue that our solution is accurate within the limitations imposed by system size and periodicity, while its precision is naturally independent of the size of the network. We further argue that it can be applied to both atomistic and coarse grained models, employing both orthorhombic and triclinic simulation cells.

Although our procedure is not system specific, we illustrate it through an application to five independently constructed atomistic epoxy systems, comprised of diglycidyl ether of bisphenol A (DGEBA) and 4,4′-diaminodiphenyl sulfone (DDS) monomers undergoing cross-linking (Figure 1). This system offers challenges of both topological and dynamical nature. The former are reflected by the presence of numerous aromatic structures and cross-linking-derived intramolecular loops, which requires the algorithm to correctly and efficiently handle looplike connectivity. The latter are rooted in those force-field and cross-linking interactions whose complex interplay determines the growth of the network. In this sense, the DGEBA/DDS mixture is ideally suited to show that our method can be successfully applied to a complex, atomistic system of relevant interest to the computational chemistry community.

## RELATION BETWEEN GELATION AND PERCOLATION

Consider a square lattice of vertices (Figure 2) in which every two neighboring vertices are joined to each other by a bond

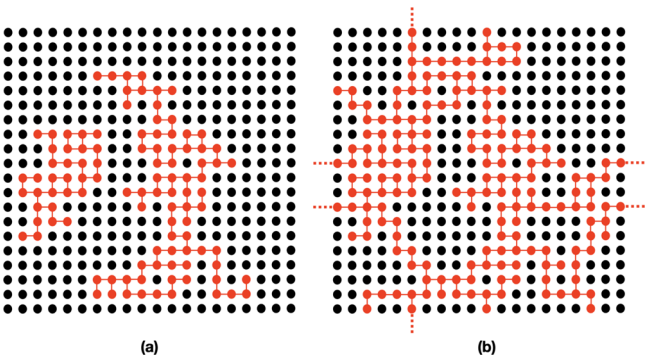



**Figure 2.** A 20 × 20 lattice of vertices, showcasing examples of two separate clusters of finite size (a) and of a percolating cluster (b).

with probability $p$. A group of vertices connected altogether forms a so-called *cluster*. When the lattice is infinite in size, a cluster that extends infinitely in all of the dimensions of the lattice is called a *percolating cluster*.[17] Such a cluster appears for the first time when $p$ reaches a critical value, known as the *percolation threshold* $p_c$. Indeed, at this stage a hypothetical fluid poured on top of the system would no longer find itself an open path through the network, thus being unable to percolate. Both the existence of this critical value and the almost-sure (i.e., with a probability of one) uniqueness of the


https://doi.org/10.1021/acs.jctc.1c00423
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

corresponding percolating cluster can be mathematically proven.[18]

Now, suppose that the size of the lattice ideally matches the number of molecules in a polymerizing system (the order of magnitude of the Avogadro number). Then, the relation between a lattice of bonded vertices and a cross-linking material (Figure 3) naturally emerges by identifying each

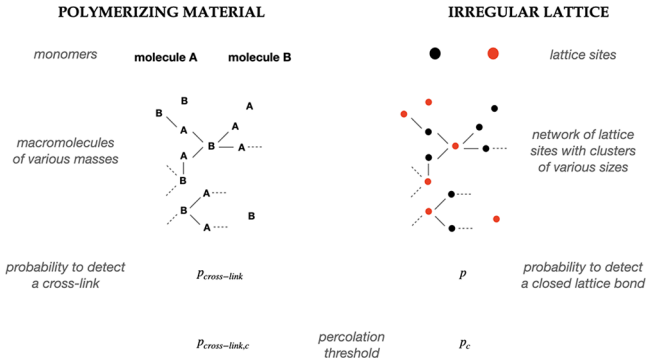



**Figure 3.** Abstracting a polymerizing chemical system into an irregular lattice of bonded vertices.

original small monomer with one vertex of an irregularly spaced lattice and, thus, macromolecules (groups of monomers bonded together) with clusters. The probability to detect a closed bond in the lattice finds its chemical counterpart in the probability $p_{cross\text{-}link}$ to detect a cross-link bond between two monomers in the system.

As stated above, the point of gelation is defined as the extent of reaction at which a macroscopic network or molecule emerges. We can relate gelation to percolation theory by requiring the formation of an equally spanning network of chemical bonds,[17] vast enough to look infinite from the monomer scale. A spanning network of connected monomers is naturally identified with a percolating cluster, which was defined above as a spanning set of bonded lattice vertices.[19] Therefore, the value of $p_{cross\text{-}link}$ at which a percolating molecule appears for the first time is exactly the gel point criterion we seek. In the SI, we show that this quantity is intimately connected to the extent of curing which is more convenient to work with (see Section S-1).

Such a strong comparison between gelation and percolation only holds if the simulated molecular system/lattice is of macroscopic size from the point of view of single monomers. Strictly speaking, because MD simulations are currently confined to the nanometer scale, this condition cannot be achieved (for example, the linear size of one of our DGEBA molecules is about 0.1 nm, so only 2 orders of magnitude below the box side, of approximately 11 nm).

However, we can work around this limitation by exploiting periodic boundary conditions to postulate a notion of percolation in a finitely sized but infinitely repeating system. Specifically, we say that a cluster is percolating if, by starting from one of its vertices (monomers) and walking along its bonds, we encounter two different periodic copies of the same initial vertex, and we do so in a manner that warrants a three-dimensional periodic structure. Hence, in a periodic MD system that is not truly infinite in size, the periodic percolation threshold directly resembles the physical significance of gelation, i.e., the emergence of a unique network of connected monomers that periodically and three-dimensionally spans the simulation box. This periodic threshold constitutes a direct structural criterion for gel point detection. The point it identifies would coincide exactly with the actual point of percolation—the gel point—in the limit of a macroscopic system, bringing the advantage that the quality of our periodic criterion is only limited by system size. It is thus predicted to improve as the simulated system gets larger. Furthermore, the periodic percolation threshold exclusively hinges on the connectivity of the simulated particles across different periodic images, regardless of the nature of the particles themselves and of the geometrical shape of the repeating simulation cell. This aspect advantageously implies that this criterion can be equally applied to both atomistic and coarse-grained systems, utilizing periodic cells of any symmetry, including the orthorhombic and triclinic systems.

In the next section, we elaborate on the definition of percolation in a periodic system and introduce a periodicity-based algorithm that can efficiently detect such a structure. Importantly, the output of such an algorithm will be shown to be univocal (i.e., a single percolation point is identified in each case), making this approach naturally precise, independently of the size of the simulated system.

## RIGOROUS ALGORITHM TO DETECT PERCOLATION IN A PERIODIC SYSTEM

To identify and characterize the nature of a percolation within a system with periodic boundary conditions, we employ techniques from basic linear algebra as well as variations of simple graph algorithms known in computer sciences. The basic idea underlying the algorithm is that, in a system with periodic boundary conditions, one only considers a finite amount of vertices (atoms) and a growing number of edges (bonds) between them. We can therefore observe an infinite, percolating cluster (molecule) if and only if, by following the edges within a connected cluster, we encounter two different periodic copies of the same vertex. The connecting vector between two copies of the same vertex within a cluster is then referred to as a *period* of that cluster, with the set of all its periods being representative of its overall periodic structure.

As all linear combinations with integer coefficients of periods are again periods, we can attribute a periodic dimension to a cluster by determining the algebraic vector-space dimension of the generally infinite set of periods, as is well-known from linear algebra. Due to the discrete nature of the molecular graph, there will be a representative set of nonzero period vectors with the same dimension as the set of all periods of the overall cluster, of which we can determine the dimension using basic linear algebra in finite computational time. Finding this representative set of vectors for each cluster in an efficient manner is the main idea behind our algorithm.

The algebraic properties of the vector space dimension allow for application of the described algorithm to any triclinic periodic bounding box and not exclusively to one adopting orthorhombic pbc.

**Computational Methods.** Starting from an arbitrary initial vertex, we can explore the entire cluster containing the starting vertex, by iteratively following all outgoing edges of an already explored subcluster, which equates to a breadth-first search (BFS) in computational science terms. If we do not encounter two different periodic copies of the same vertex by following edges, there is no percolation. If we do encounter two different copies, there might be a percolation occurring in any number of dimensions from 1 to the total dimension of the


https://doi.org/10.1021/acs.jctc.1c00423
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

system. However, attempting to break down the test for percolation independently for each spatial dimension does not provide any insight into the overall periodic dimension of a full percolating cluster (see Figure S1 for an illustration of the lack of rotational invariance of this separated approach). We therefore need to determine the dimension of the full set of periods of a cluster and cannot look at the periods in each dimension independently.

To gain more insight into the structure of the percolating cluster, whenever we encounter another copy of a vertex that we have visited previously, we keep track of the connecting vector between it and the first encountered vertex of its kind. Here the vector between the positions of two observed copies is a period of the overall lattice structure of the percolating cluster and needs to be added to the overall set of periods for the cluster.

This also means that we do not need to consider outgoing bonds from any such copy of an already encountered vertex in our exploration, limiting the set of vertices to visit in a possibly infinitely expanding cluster.

Let us use the configuration as seen in Figure 4 for an example on a simplified system. The system has periodic boundary conditions only in one dimension and consists of a total of 7 vertices labeled 0 through 6. We start our exploration of the cluster at the vertex 0, which we immediately encounter for the first time. In the first step of the BFS we visit all vertices at a distance of 1 edge from vertex 0 which makes us encounter the vertices 1, 2, 3, and 4 each for the first time. In the second step of the BFS we move on to all vertices at a distance of 2 edges from vertex 0, which leads to the first visit of vertices 5 and 6 as well as encountering copies of vertices 1, 2, 3, and 4 in neighboring periodic cells. We then take the distance vectors between each of these copies and the first-encountered copies of these vertices and keep them as members of the sets of periods. As we are in one dimension, these periods are equivalent to differences in cell indices, where negative and positive cell indices identify, respectively, the cells to the left and right of the starting one, which leaves us with members $\pm 1$ of the set of periods. In a third step we only follow the outgoing edges of vertices 5 and 6, due to which we only encounter the already visited original copies of 5 and 6, which contribute no further nonzero periods to our set of periods. We now have followed all edges within the cluster, encountering all vertices of the cluster at least once, completing the gathering step of the algorithm.

In the end, we can determine the algebraic dimension $d$ of the overall set of collected periods via an appropriate algorithm. Any optimization where we only keep track of a set of linearly independent vectors may also be applied to reduce run time and memory requirements.

In three-dimensional space, this dimension $d$ will help us distinguish between percolating clusters that have the structure of a long string of molecules ($d = 1$), those that have the structure of a sheet ($d = 2$), and those that span the entire periodic space with a full lattice structure ($d = 3$). The point in time when we first observe a cluster of full lattice dimension $d = d_{max}$ can be uniquely defined as the percolation point of the system, where $d_{max}$ denotes the number of dimensions with periodic boundary conditions (pbc). In a typical three-dimensional simulation box in the shape of a parallelepiped with pbc in all directions, this amounts to $d_{max} = 3$.

Going back to the above example in Figure 4, we have a set of periods with at least one nonzero entry. As $d_{max} = 1$ in this simplified case and any nonzero period amounts to at least one-dimensional periodicity of the cluster, we obtain a percolation of maximum dimension $d = 1$ for the studied cluster. Therefore, we could have stopped our BFS after step 2 of the analysis, since the dimension of the set of periods could not possibly have grown any further.

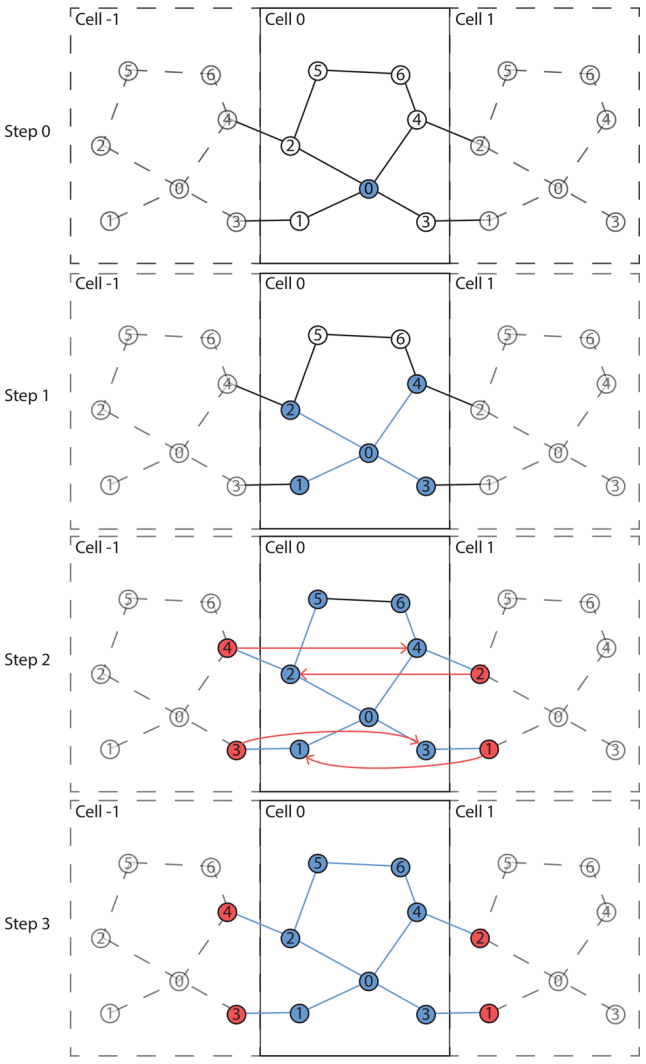



**Figure 4.** Graphical illustration of the percolation detection algorithm in a two-dimensional system with periodic boundary conditions in only one dimension ($d_{max} = 1$). Percolation in one dimension (the maximum allowed) is eventually observed. Cell 0 indicates the arbitrary origin cell with the arbitrary origin at atom 0. Each circle/vertex represents an atom and each edge a bond. Periodic copies are drawn with reduced opacity, and copies of edges are drawn with dashed lines. Cells −1 and 1 are the first copies in their respective direction. Red vertices indicate the detection of a duplicate vertex, leading to truncation of the BFS; the respective periodicity vectors are marked in red. The edges already considered in the BFS are colored blue, as are the first-encountered copies of each vertex.

For a more rigorous and formal definition of the full algorithm and its individual steps, please refer to the SI.

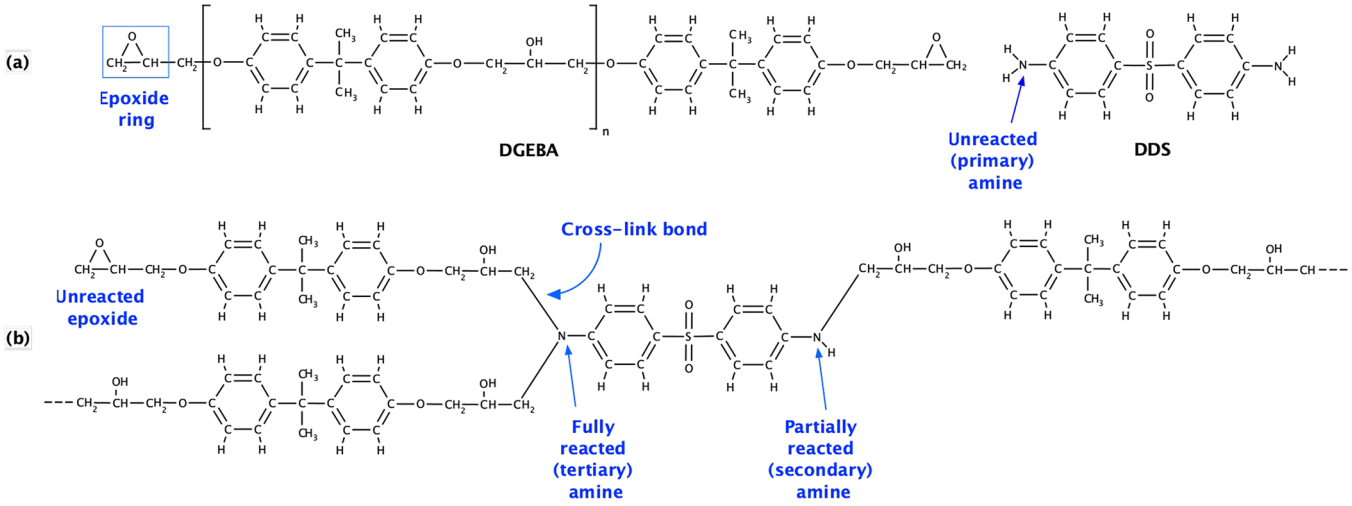
**Properties of the Algorithm.** *Invariance of the Results.* As proven in the SI (see Section S-3.3), the following property holds:

**Theorem** (Invariance of detected percolation dimension). The lattice dimension of a connected cluster of vertices as



https://doi.org/10.1021/acs.jctc.1c00423
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

**Scheme 1. (a) Unreacted Epoxy (DGEBA) of Length $n$ and the Amine Hardener (DDS) and (b) Possible Configuration within the Cured DGEBA-DDS Covalent Network, Showing Both Reacted and Unreacted End Groups**




obtained by the percolation algorithm described previously is invariant under affine linear transformations.

Another property immediately resulting from the proof of the above theorem is the following:

**Corollary** (Independence of starting vertex). The lattice dimension of a connected cluster of vertices as obtained by the percolation algorithm is independent of the chosen starting vertex for the algorithm.

We are therefore safe to start exploring any possible cluster from any of its constituting vertices without affecting the detected periodic dimension. Furthermore, neither the exact positioning of the periodic boundary conditions nor translations of the entire system within the periodic boundaries can influence the outcome of the algorithm, making it satisfyingly stable in its applicability.

*Algorithmic Complexity.* The terminal condition of not following outgoing edges originating from copies of encountered vertices causes the overall number of edges and vertices in the exploration process to be bounded by the number of edges $e_c$ and vertices $v_c$ in the cluster. This guarantees that the proposed algorithm will always terminate. It also bounds the worst-case run time complexity of the algorithm at $O(e_c + v_c)$, since the periodicity dimension check can be carried out in constant time for a fixed overall system dimension $d_{max}$. For a total number of edges $e$ and vertices $v$ in the entire system, this leaves us with a complexity $O(e + v)$ for the execution on the state of the whole system at a fixed point in time.

This also means that the algorithm shares the same linear complexity as those employed for simple cluster detection as well as for cluster mass detection, both of which also amount to a BFS with varying constant-time requirements for the processing of each encountered vertex. This property makes it a well-suited alternative to the algorithms presently employed in the analysis of periodic thermoset-hardening simulations, both from a theoretical and a computational point of view.

With only minor variations, the described algorithm can also be applied to the detection of gelation in finite-size and nonperiodic systems. For this purpose one would introduce a virtual particle for each dimension of the system and introduce connections between vertices within a certain distance of the edge of the system in that direction to the virtual particle, introducing a virtual periodicity condition. Running the described percolation detection then provides the number of dimensions in which a particle spans the entirety of the spatial extent of the finite system in that respective dimension, with similar implications for the gelation of the system when percolation in all dimensions is observed as in the periodic system.

## DETECTING PERCOLATION IN A DGEBA/DDS MODEL SYSTEM

In this section, we illustrate an application of our algorithm by computing the point of periodic percolation for a relatively large, highly cross-linked epoxy system. We also compute the gel point by means of the traditional methods based on the growth of the largest molecule's mass and on the onset of secondary cycles, comparing these results to our percolation measurement. However, we do not intend to offer a comparison of such measurements with any experimental data; rather, we endeavor to highlight the deficiencies suffered by indirect methods when they are applied to large molecular networks.

Chemically, epoxies are reactive prepolymers defined by the presence of an epoxide ring at either end of the molecule (Scheme 1). When mixed at a sufficiently high temperature with a hardener, such epoxide gets covalently cross-linked, or "cured", with the end group of the hardener, which is often an amine. The physical qualities of the material depend on the resulting highly cross-linked covalent network, thus making epoxies an ideal model system for the investigation of network properties.

Each example epoxy system contains 1000 hardener DDS (4,4′-diaminodiphenyl sulfone) and 2000 DGEBA (diglycidyl ether of bisphenol A) resin molecules, of length $n$ = 0, for a total of 127 000 atoms. The chosen 2:1 ratio is stoichiometric and meant to resemble laboratory conditions. The LAMMPS[20] (Dec. 2018 stable) MD package and the GAFF[21] force field were employed, utilizing partial atomic charges derived via a RESP[22] fragment-based approach. A 1 fs time step and periodic boundary conditions were applied. The cutoff for the calculation of Coulomb interactions in direct space and for VDW interactions was 10 Å. All systems were initially minimized and then equilibrated for 2 ns in the NPT ensemble at 503 K[23] and 1 atm, using Nosé−Hoover thermostat and


https://doi.org/10.1021/acs.jctc.1c00423
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

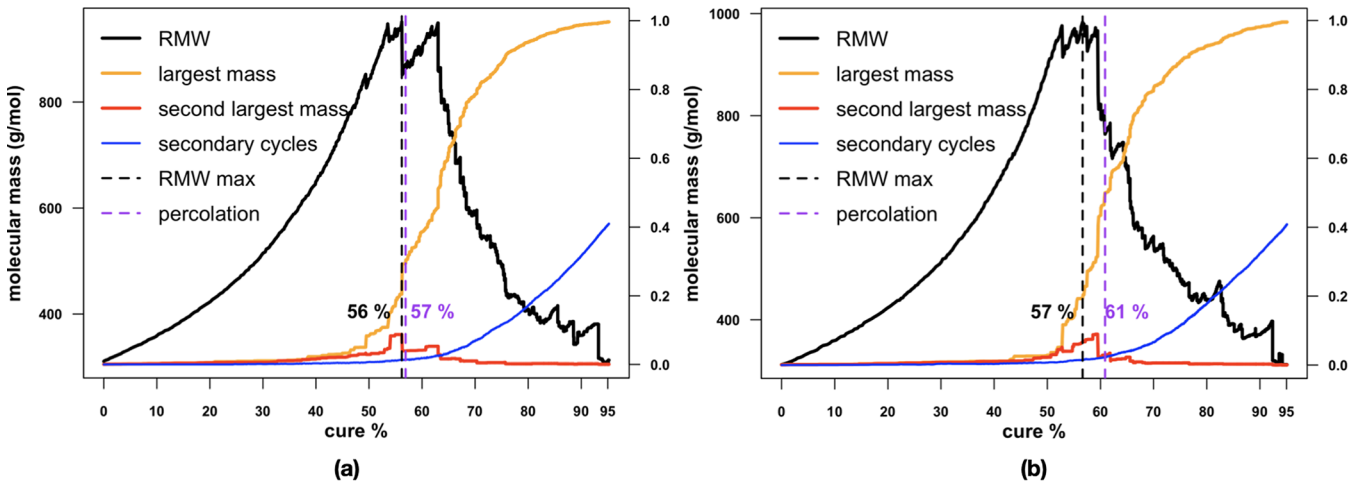



**Figure 5.** Gel point measurements for two epoxy systems based on various strategies. Black curve: reduced molecular weight; orange curve: normalized largest mass; red curve: normalized second largest mass; blue curve: normalized number of secondary cycles.

barostat. Following this, the curing reaction was simulated in the NPT ensemble by means of the *fix bond/react*[24] algorithm shipping with LAMMPS, reaching at least a 95% extent of reaction. Further simulation details can be found in the SI. A visualization of one of our systems at three different curing extents is provided in Figure 1.

The point of percolation in our five model systems was detected in the (61 ± 3)% range of reaction extents. The existence of a nonzero standard deviation for this result may seem surprising. In fact, we have shown that the point of periodic percolation is well-defined mathematically and that it can be determined by a rigorous algorithmic procedure, implying that there can be no uncertainty related to each percolation measurement. However, because the five systems were each constructed with different initial conditions, they are expected to have slightly different points of percolation as well, leading to the observed uncertainty. It is important to realize that such ensemble disparities, associated with statistical fluctuations in system construction, are supposed to disappear in the limit of a macroscopic simulated system. As a matter of fact, in this limit, the point of periodic percolation tends to the actual point of percolation, which is uniquely determined by the topological properties of the network, as we argued earlier.

The situation for mass-based methods is quite different. The results for two of our epoxy models are shown in Figure 5, while results for all other systems are reported in Figure S1 of the SI. Mass-based methods try to indirectly detect the gel point by seeking the emergence of a single macromolecule that considerably outweighs all others. This event is usually recognized either by the point of inflection of the largest-mass buildup (orange curve in Figure 5) or by the point at which this curve considerably rises over the growth of the second largest mass (red curve).

Both plots in Figure 5 show that the largest group is able to take over the second-largest one while the latter is also still growing (at a lesser rate), leaving doubts about which point should mark a "significant" difference between the two. As a matter of fact, any value in the curing range 50−60% could reasonably be chosen for such a purpose, as depicted by the error bar in Figure 6. The figure compares gel point measurements as obtained by percolation with those from indirect methods. Choosing the inflection point of the heaviest molecule growth is certainly analytically convenient, yet

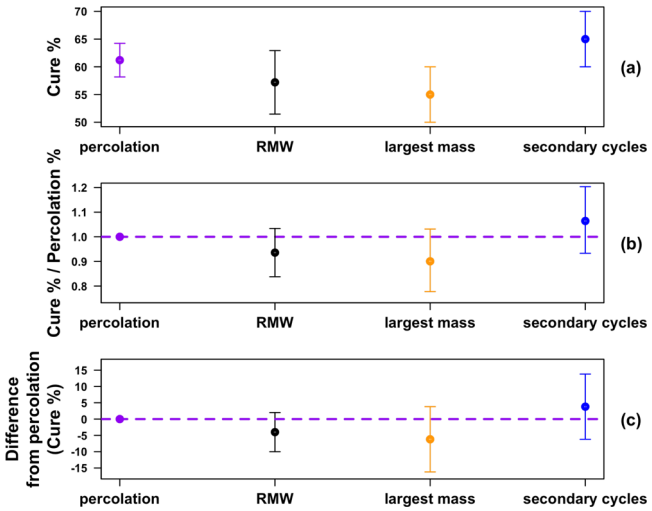



**Figure 6.** Comparing the gel point percolation criterion to indirect measurements for all five systems. (a) Gel point values are expressed as cure percentages. The percolation error bar is the standard deviation of the results for the five systems. The RMW error bar is obtained by error propagation, estimating an 8% and 10% uncertainty for the peaks of the black curves in Figure 5a,b, respectively, and no error bars for the other three systems. Error bars for largest mass and secondary cycles both have a 10% width, respectively centered around 55% and 65% curing, as discussed in the text. These midpoints are purely a guide to the eye. (b) Gel point measurements from all methods have been rescaled by their corresponding percolation values, to highlight how much indirect results precede or follow the percolation threshold. Error bars are then obtained as above. (c) The average percolation value has been subtracted from each of the other results, to gauge to what extent indirect measurements agree with the point of percolation.

arbitrary. Moreover, its identification would require a good quality functional fit, based on undesirable smoothing of a significant portion of the data as well. Alternatively, one could rely on a simpler but also more subjective visual estimate,[13] which would again locate the gel point in the same 50−60% range.

Seeking a more quantitatively precise criterion, the reduced molecular weight (RMW) has been introduced in order to measure the average molecular mass of all the groups except


https://doi.org/10.1021/acs.jctc.1c00423
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

the largest one. The maximum point of this quantity is supposed to mark the appearance of the single highly massive group that mass-based methods seek. In Figure 6, the peak of the RMW identifies the interval (57 ± 6%) for the gel point.

The RMW is definitely a more analytic and less subjective criterion, but its quality crucially depends on it exhibiting a single, sharp peak, followed by a steady fall. This, however, need not be the case: both peaks in the black curve of Figure 5a are of comparable heights and correspond to negative jumps in the second-largest group of equally comparable magnitudes, thus making the choice between them ambiguous. Moreover, looking at Figure 5b, one can identify multiple local maxima contributing to a single global peak, smeared over about a 10% curing range. Since all extrema in this interval are of comparable heights, there is no clear reason for preferring the global maximum in this case. These two observations lead to uncertainties of 8% and 10%, respectively, on which the peak of the black curve should be used to mark the gel point in Figure 5a,b. The consequence of error propagation is then a large 6% error bar in Figure 6a, obtained assuming no uncertainty for the other three epoxy systems, in which the RMW shows a much clearer single peak (Figure S2 of the SI).

Although indirect methods based on mass have been successfully applied to systems with a relatively small number of monomers,[12−16] our results have clearly uncovered the shortcomings suffered by these strategies when applied to sufficiently large molecular networks.

On the one hand, we have shown that multiple massive macromolecules are able to coexist for a perceivable amount of time, before a single group eventually takes over. This causes the RMW to exhibit ambiguity due to the presence of multiple comparable peaks and prevents the existence of a sharp point after which a single macromolecule predominates in mass. As a consequence, mass-based methods are affected by limited precision, as visible from the large error bars of Figure 6.

On the other hand, from Figure 5 (and Figure S2) we are able to notice how the percolation point discernably follows the peak of the RMW in all cases, with curing delays ranging from 1 to 6%. This is evidence that the largest macromolecule is able to considerably outweigh all other groups while its bond network still does not truly span the system. Such underestimation is systematic, indicating that mass concentration in a single group may at least be partially blind to those structural changes involved in percolation. As a result, even if mass-based methods were precise enough to provide a single point rather than a range, they could still fail in accuracy, in light of the fact that mass accumulation and percolation can indeed be separate events. Nevertheless, Figure 6b,c shows that the point of percolation does fall within the black and orange error bars, indicating that mass-based measurements could at least be in agreement with percolation results in this case.

Assuming a more structural perspective, the onset of secondary cycles is the other indirect criterion usually applied to detect the gel point. This is motivated by the fact that the gel-like macromolecule, once emerged, must necessarily enclose most of the unreacted end groups remaining in the system, leading to a sudden rise in the number of intramolecular cross-linking events. A secondary cycle can be identified by detecting a closed loop connecting a cross-linked monomer back to itself.

From the blue curves in Figure 5, we can observe how the buildup in the number of secondary cycles consistently shows two distinct linear regimes at low and high curing extents. By intersecting the corresponding tangent lines, one may determine at which point the slope change occurs and use it to mark the inception of intramolecular reactions. The greatest disadvantage of this method lies in the very sensitive dependence of the intersection abscissa[25] upon the slopes of the two lines. Because such slopes depend on the widths of the two linear regimes, which must be subjectively determined by the experimenter, once again only a curing range can be identified. A quick visual estimate locates the intersection abscissa in the 60−70% interval, about 10% to the right of our previous mass-based approximations.

Differently from mass-based methods, the onset of intramolecular reactions is expected[16] to be in better agreement with percolation, since this method is indirect but at least sensitive to changes in network topology. Our data seem to confirm this expectation: Figure 6b,c shows that the average point of percolation falls within the blue error bar, corresponding to secondary cycles. Although we noticed earlier that the same is true for the black and orange error bars, Figure S2 of the SI adds the observation that, with one exception, the point of percolation always falls within the same 60−70% uncertainty interval that the onset of intramolecular reactions does. The same cannot be said for the 50−60% interval identified by the largest mass method, since only one percolation data point lies there.

## CONCLUSIONS

We have detailed how the emergence of a percolating molecular cluster, in a lattice of infinite size, is the natural structural criterion to detect the appearance of a gel-like macromolecule in a polymerizing material. To make this criterion tractable to MD investigation, we have introduced a notion of percolation for the case of a finitely sized but infinitely periodic molecular lattice, based on the appearance of a three-dimensional periodic structure in the simulated system. Furthermore, we have implemented an algorithm that can both efficiently and univocally identify such a structure. This constitutes a rigorous and accurate procedure to detect gelation, whose precision is only affected by statistical fluctuations in system construction. Because such fluctuations are supposed to average out in the limit of a macroscopic system, the quality of our criterion is expected to improve at greater system sizes.

By stark contrast, we have revealed how indirect methods based on mass suffer from severe pitfalls when applied to large polymerizing systems. These are mainly due to the impossibility of reliably identifying the rise of a single mass-predominating macromolecule and to the time gap existing between the appearance of such group and the appearance of a truly percolating structure. At the system size we investigated, this gap is discernible but still small enough for mass-based measurements to agree with percolation results. However, we would expect it to widen as larger molecular networks with even vaster conformational spaces are investigated. Finally, we showed that the same size limitation does not affect the inception of intramolecular reactions, which depends on network topology rather than mass, thus providing better agreement with percolation results. Nevertheless, in analogy with mass-based strategies, this method is too an indirect criterion, and its precision is greatly limited by the fact that the onset of secondary cycles cannot be univocally identified.


https://doi.org/10.1021/acs.jctc.1c00423
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c00423.

Further details on the parameters for the simulated systems and the design of the algorithm for the detection of percolation as well as the mathematical proof of the stated properties of invariance of said algorithm (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

**Ana-Sunčana Smith** − *PULS Group, Institute for Theoretical Physics and Interdisciplinary Center for Nanostructured Films (IZNF), Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), 91058 Erlangen, Germany; Group of Computational Life Sciences, Division of Physical Chemistry, Ruđer Bošković Institute, 10000 Zagreb, Croatia;* orcid.org/0000-0002-0835-0086; Phone: +49 91318570565; Email: ana-suncana.smith@fau.de; Fax: +49 91318520860

### Authors

**Mattia Livraghi** − *PULS Group, Institute for Theoretical Physics and Interdisciplinary Center for Nanostructured Films (IZNF), Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), 91058 Erlangen, Germany;* orcid.org/0000-0003-2344-3907

**Kevin Höllring** − *PULS Group, Institute for Theoretical Physics and Interdisciplinary Center for Nanostructured Films (IZNF), Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), 91058 Erlangen, Germany;* orcid.org/0000-0002-9497-3254

**Christian R. Wick** − *PULS Group, Institute for Theoretical Physics and Interdisciplinary Center for Nanostructured Films (IZNF), Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), 91058 Erlangen, Germany; Competence Unit for Scientific Computing (CSC), Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), 91058 Erlangen, Germany;* orcid.org/0000-0003-3365-0942

**David M. Smith** − *Group of Computational Life Sciences, Division of Physical Chemistry, Ruđer Bošković Institute, 10000 Zagreb, Croatia;* orcid.org/0000-0002-5578-2551

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.1c00423


### Author Contributions

§M.L. and K.H. contributed equally to this work. A.-S.S. and D.M.S. conceived and supervised the study. K.H. devised the algorithm to identify the percolation point in a periodic graph. C.R.W. and M.L. designed the molecular model system. M.L. built and ran all the simulations, as well as performed the analysis based on existing protocols. C.R.W. supervised the simulations and proper data curation. M.L. and K.H. implemented the new algorithm, developed the numerical tool, performed the analysis and wrote the first draft. All authors contributed to the interpretation of data and writing the final paper.

### Funding

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 377472739/GRK 2423/1-2019 and 363055819/GRK 2415. The authors are very grateful for this support.

### Notes

The authors declare no competing financial interest.

A sample implementation of the algorithm detailed in this work including source code is provided at https://github.com/puls-group/percolation-analyzer.

## ■ REFERENCES


(1) Genidy, M. S.; Madhukar, M. S.; Russell, J. D. An investigation of cure induced stresses in low cure temperature thermoset polymer composites. *J. Reinf. Plast. Compos.* **1999**, *18*, 1304−1321.

(2) Beyazit, S.; Tse Sum Bui, B.; Haupt, K.; Gonzato, C. Molecularly imprinted polymer nanomaterials and nanocomposites by controlled/living radical polymerization. *Prog. Polym. Sci.* **2016**, *62*, 1−21.

(3) Lee, K. H.; Kang, M. S.; Zhang, S.; Gu, Y.; Lodge, T. P.; Frisbie, C. D. Cut and stick" rubbery ion gels as high capacitance gate dielectrics. *Adv. Mater.* **2012**, *24*, 4457−4462.

(4) Oliva, N.; Conde, J.; Wang, K.; Artzi, N. Designing hydrogels for on-demand therapy. *Acc. Chem. Res.* **2017**, *50*, 669−679.

(5) Flory, P. J. *Principles of polymer chemistry*; Cornell University Press: Ithaca, NY, 1953.

(6) Grenier-Loustalot, M.-F.; Grenier, P.; Horny, P.; Chenard, J.-Y. Reaction mechanism, kinetics and network structure of the DGEBA-DDS system. *Br. Polym. J.* **1988**, *20*, 463−476.

(7) Serrano, D.; Harran, D. On the increase of viscoelastic modulus with advancement of reaction of an epoxy resin. *Polym. Eng. Sci.* **1989**, *29*, 531−537.

(8) Danielson, S. P.; Beech, H. K.; Wang, S.; El-Zaatari, B. M.; Wang, X.; Sapir, L.; Ouchi, T.; Wang, Z.; Johnson, P. N.; Hu, Y.; et al. Molecular Characterization of Polymer Networks. *Chem. Rev.* **2021**, *121*, 5042−5092.

(9) Vidil, T.; Tournilhac, F.; Musso, S.; Robisson, A.; Leibler, L. Control of reactions and network structures of epoxy thermosets. *Prog. Polym. Sci.* **2016**, *62*, 126−179.

(10) Flory, P. J. Molecular size distribution in three dimensional polymers. II. Trifunctional branching units. *J. Am. Chem. Soc.* **1941**, *63*, 3091−3096.

(11) Stockmayer, W. H. Theory of molecular size distribution and gel formation in branched-chain polymers. *J. Chem. Phys.* **1943**, *11*, 45−55.

(12) Estridge, C. E. The effects of competitive primary and secondary amine reactivity on the structural evolution and properties of an epoxy thermoset resin during cure: A molecular dynamics study. *Polymer* **2018**, *141*, 12−20.

(13) Varshney, V.; Patnaik, S. S.; Roy, A. K.; Farmer, B. L. A molecular dynamics study of epoxy-based networks: cross-linking procedure and prediction of molecular and material properties. *Macromolecules* **2008**, *41*, 6837−6842.

(14) Cheng, K.-C.; Chiu, W.-Y. Monte Carlo simulation of polymer network formation with complex chemical reaction mechanism: kinetic approach on curing of epoxides with Amines. *Macromolecules* **1994**, *27*, 3406−3414.

(15) Chen, Y.-C.; Chiu, W.-Y. Polymer chain buildup and network formation of imidazole-cured epoxy/phenol resins. *Macromolecules* **2000**, *33*, 6672−6684.

(16) Hädicke, E.; Stutz, H. Comparison of the structure of step-growth networks obtained by Monte Carlo simulation and branching theory. *J. Appl. Polym. Sci.* **2002**, *85*, 929−935.

(17) Stauffer, D.; Aharony, A. *Introduction to percolation theory*; Taylor & Francis: London, 1992.

(18) Grimmett, G. R. *Percolation*, 2nd ed.; Springer: Berlin, Heidelberg, 1999.

(19) Stauffer, D. Can percolation theory be applied to critical phenomena at gel points? *Pure Appl. Chem.* **1981**, *53*, 1479−1487.

(20) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.

(21) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(22) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(23) White, S. R.; Mather, P.; Smith, M. Characterization of the cure-state of DGEBA-DDS epoxy using ultrasonic, dynamic mechanical, and thermal probes. *Polym. Eng. Sci.* **2002**, *42*, 51−67.
(24) Gissinger, J. R.; Jensen, B. D.; Wise, K. E. Modeling chemical reactions in classical molecular dynamics simulations. *Polymer* **2017**, *128*, 211−217.
(25) Patrone, P. N.; Dienstfrey, A.; Browning, A. R.; Tucker, S.; Christensen, S. Uncertainty quantification in molecular dynamics studies of the glass transition temperature. *Polymer* **2016**, *87*, 246−259.


https://doi.org/10.1021/acs.jctc.1c00423
*J. Chem. Theory Comput.* XXXX, XXX, XXX−XXX