PHYSICAL REVIEW E **113**, 024301 (2026)

Influence of topology on performance of pore networks in membrane filters

M. Illingworth,[1] B. Gu,[2] L. Cummings,[1] and L. Kondic[1]
[1]*Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, New Jersey 07102, USA*
[2]*Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, Massachusetts 01609, USA*

(Received 6 August 2025; accepted 12 January 2026; published 3 February 2026)

Membrane filtration, a crucial and widely used industrial process, has attracted significant attention in mathematical modeling. Previous theoretical endeavors have posited the internal structure of membrane filters as a network of cylindrical pores, with fouling characterized as the gradual accumulation of fouling particles on the inner walls of these pores. Subsequently, membrane filter performance was evaluated employing metrics such as total throughput and accumulated foulant concentration throughout the filter's lifespan; and the correlation between such metrics and certain network properties, such as porosity and tortuosity, was analyzed. In this work, we analyze the correlation between the performance of such networks and their topological characteristics, with the objective of identifying the most pertinent topological features for membrane filter design for a fixed porosity setup. Persistent homology serves as our primary instrument for quantifying topological features. Specifically, we generate persistence diagrams for the pore networks utilizing thresholding based on the pore radii. The data encoded in these persistence diagrams are then statistically correlated with the performance metrics. We identify a strong correlation between total throughput and our topological measures and demonstrate explicitly that membrane topology significantly influences filtration performance, distinct from the effects of membrane geometry. We also report a strong positive correlation between total throughput and accumulated foulant concentration and find that, for fixed porosity setups as used in this work, tortuosity does not provide a useful measure for permeability.

DOI: 10.1103/s8n8-vxzx

I. INTRODUCTION

Membrane filtration is an important and ubiquitous process in industrial applications, used in a wide variety of industrial systems such as water purification [1], air filtration [2], and wastewater treatment [3]. Membrane filters used for these purposes come in a wide variety of architectures, including single-layer thin porous films [4], multilayered porous membranes [5], large-scale continuous sheets of layered fibrous material [6–8], and deformable porous media [9].

A profusion of models have been developed in recent years to describe the structure and function of these membrane filters [10–13]. Earlier attempts to model the internal complexity of pores in membrane filters include simple homogenized theoretical models consisting of multiple layers of heterogeneous porous materials [5,14], and simple pore-branching structures that can admit porosity and/or pore-size gradients [15–18]. Recent advances in imaging techniques have made it possible to compare more accurately theoretical models to experimental findings [19,20], and to extract pore network structures from real images (see, e.g., Fig. 1, adapted from Kelly *et al.* [21]). The present work builds on earlier research that models the internal structure of membrane filters as a random network of pores [15,21–24]. In particular, we follow the approach of Gu *et al.* [23,24], who studied artificially generated random pore networks comprising cylindrical pores whose initial radii are drawn from a uniform probability distribution, with fouling modeled as an adsorption process; i.e., a gradual accretion of fouling particles on the inner walls of the pores. Simulation-based approaches were used to measure membrane filter performance, using metrics such as total filtrate throughput over the filter lifetime and accumulated foulant concentration in the filtrate, with a view to improving membrane filter design.

The main thrust of the present work involves the use of an algebraic topological tool called *persistent homology*; in particular, we use mathematical objects known as *persistence diagrams* (PDs), widely used in a burgeoning field known as *topological data analysis* (TDA). TDA is used in many applications, including image analysis [25], signal analysis [26], and the study of viral evolution [27]. Here, PDs and measures derived from them will be used to identify and probe correlations between the performance of membrane filters and their pore topology.

The paper is structured as follows: In Sec. II, we describe the generation of our pore networks and the simulation of filtration through them. This process involves modeling fluid flow through the network and the advection and adsorption of foulant particles within it. Subsequently, we provide a high-level overview of persistence diagrams and the relevant derived measures. In Sec. III, we present our main findings, focusing on exploring the correlations between topology, geometry, and the filtration performance of pore networks. Finally, in Sec. IV, we summarize the key findings and draw conclusions. Auxiliary aspects of the problem are presented in a series of Appendixes.

II. METHODS

This section describes the methods employed to carry out the research presented in the rest of the paper. Section II A

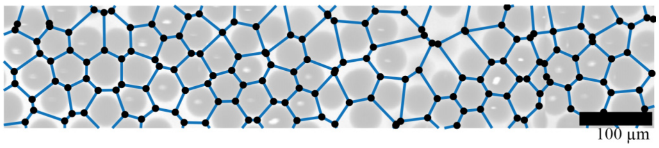

FIG. 1. A pore network representative of the network type we consider, from Kelly *et al.* [21].

discusses appropriate scalings and presents the mathematical model for particle-laden fluid flow through a single pore, where the foulant particles are advected with the flow and adsorbed at the pore wall. In Sec. II B we extend the model to arbitrary networks of pores and detail the methods used to generate the networks and characterize their properties. Finally, in Sec. II C we introduce persistence diagrams and their use as indicators of filtration performance.

A. Filtration in a single pore

We model filtration of particle-laden fluid through a membrane filter under a specified (constant) pressure drop, although the formulation can be readily modified to a specified flux scenario instead. We first describe the filtration and fouling process in a single, nearly cylindrical, pore. The fluid to be filtered, known as the feed, flows through a pore in the network, advecting foulant particles (to be removed) as it flows. The foulant particles adsorb to the pore's inner surface at a certain rate; this particle accretion is used to calculate an update to the local pore radius at each time step (described further below), until the radius goes to zero at some location and the pore is deemed closed. We leave the study of more complicated models of filtration (which may include other types of fouling, or particle desorption) for future work; in the present work, we focus on the scenario outlined above.

Figure 2 schematizes our basic network building block: a pore of length L and radius $R(Y,T)$, where Y is a local variable that measures distance from the pore inlet. The sketch depicts advection of multiple small foulant particles through the pore, which adsorb to the pore's inner surface at a rate determined by a parameter Λ, which has dimensions of length/time [23] and is a measure of the affinity between particles and membrane material. In this model of particle adsorption the foulant particles "melt" smoothly into the pore wall, maintaining a circular cross-section while conserving solid mass. We will further discuss the assumptions inherent in this adsorption model when we describe pore shrinkage below.

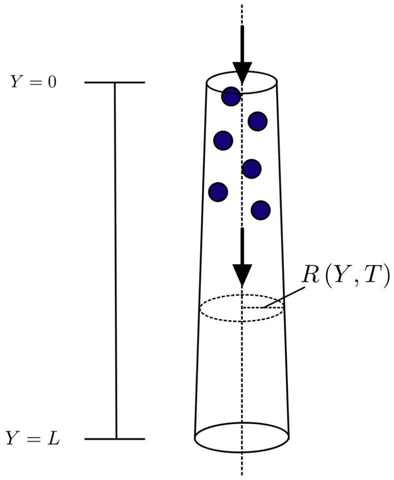

FIG. 2. A sketch of nearly cylindrical pore geometry, depicting a pore of length L and nearly constant radius $R(Y,T)$. Y is a local variable that measures distance from the pore's inlet.

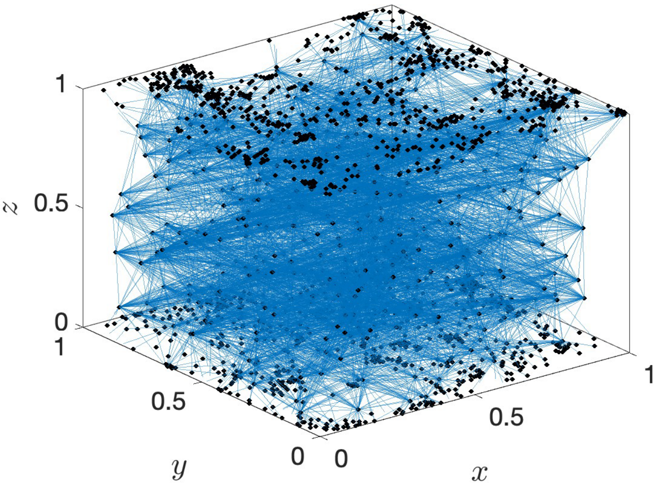

FIG. 3. A full 3D image of a typical (virtual) membrane filter network.

With these definitions in place, we introduce the scalings used to nondimensionalize our model. We assume that each pore is part of a large network, embedded in a cubical unit of membrane filter of side-length W. We scale all distances, including pore length, radius and the local length variable Y, by W, making the membrane filter element a unit cube; see Fig. 3. The pressure P at the inlet (or outlet) is scaled with the (constant) pressure P_0 at the upstream surface of the membrane filter, and foulant concentration C is scaled with the concentration C_0 of foulant in the feed solution. For the timescale we choose the characteristic time of the pore shrinkage due to particle adsorption, $W/(\alpha\Lambda C_0)$ (where α is the effective volume occupied by a particle once deposited on the pore wall, see Ref. [23]). The scaling for the flux through the membrane unit is based on the maximum flux that a pressure difference P_0 could induce through the membrane, for a fluid of viscosity μ.

The dimensional variables in the model are listed in Table I. Based on the above discussion, we define corresponding dimensionless variables and a dimensionless parameter λ that emerges, as follows:

$$(l, r, y) = \frac{1}{W}(L, R, Y), \quad p = \frac{P}{P_0}, \quad c = \frac{C}{C_0}, \tag{1}$$

$$t = \frac{\alpha\Lambda C_0}{W}T, \quad q = \frac{8\mu}{\pi W^3 P_0}Q, \quad \lambda = \frac{8\mu}{\pi W P_0}\Lambda. \tag{2}$$

Here, uppercase quantities are dimensional, and the lowercase analogs are the dimensionless equivalents; Q (q) is the flux through the pore. In the following, we present our model with respect to the dimensionless variables.

TABLE I. Dimensional quantities used in the filtration model.

L	Pore length	Q	Flux through pore
R	Pore radius	C	Concentration of foulant particles
W	Side-length of the cubic filter unit	μ	Fluid viscosity
P	Fluid pressure	Λ	Affinity between foulant particle and pore wall
P_0	Pressure at the membrane top surface	T	Time
α	Particle volume divided by 2π	C_0	Foulant concentration in the feed solution

Since the full Navier-Stokes equations for fluid flow are complicated and nonlinear, and thus computationally expensive to solve (especially if we are to simulate flow through thousands of connected pores, on many different networks), we use the Hagen-Poiseuille (H-P) approximation instead. This approximation provides a simple, linear model of fluid flow through a pore and gives the fluid flux q in terms of the pressures p_{in} and p_{out} at the pore's inlet and outlet, respectively, and the *conductance* k of the pore (a measure of how effectively the pore can transport fluid) as

$$q = k(p_{\text{in}} - p_{\text{out}}). \tag{3}$$

The H-P approximation requires that the flow have low Reynolds number (satisfied in most common liquid filtration scenarios) and that the radius of each pore be sufficiently small compared to its length. This second condition is discussed in Appendix A.

As fluid transits a pore in the axial (y) direction, it advects foulant particles, which adsorb onto the inner pore wall at a rate characterized by the dimensionless parameter λ; this process is called *adsorptive fouling*. The concentration $c(y, t)$ of foulant particles in a pore is assumed to obey the following steady-state advection equation [23]:

$$q\frac{\partial c}{\partial y} = -\lambda rc, \quad 0 \leqslant y \leqslant l, \tag{4a}$$

$$c(0, t) := c_{\text{in}}(t), \tag{4b}$$

where $c_{\text{in}}(t)$ is the concentration of foulant particles at the pore inlet and $r = r(y, t)$ is the local pore radius. The right-hand side of Eq. (4a) is a sink term that models the deposition of foulant particles at the pore wall. Inherent in this model is the assumption stated earlier, that particles deposited at a given y-location "melt" into the wall, contributing to a mass-conserving change in the local pore radius. Equations (4) admit the following analytic solution:

$$c(y, t) = c_{\text{in}}(t)\exp\left(-\frac{\lambda}{q}\int_0^y r(y', t)dy'\right), \quad 0 \leqslant y \leqslant l, \tag{5}$$

for *any* pore shape $r(y, t)$. Equation (5) tells us how much foulant is adsorbed at the pore wall, which allows us to evolve the pore's local radius, as discussed next.

The prior work on which we build [15,23,24] made the simplifying assumption that pores are initially circularly cylindrical, and remain so over time while shrinking due to fouling. While convenient (permitting significant analytical progress and efficient numerical implementation on large-scale networks), this model leads to some degree of error: in reality [as is evident from Eq. (5)] pores shrink faster at their inlets, where foulant concentration is highest. In the present work we improve on that model by assuming that pores evolve as sections of a circular cone, with local linear radius function given by

$$r(y, t) = a(t) + b(t)y, \quad 0 \leqslant y \leqslant l, \tag{6}$$

where the coefficients $a(t)$ and $b(t)$ evolve according to a system of ordinary differential equations (ODEs), derived in Appendix B. These ODEs are obtained by imposing (i) the condition that the pore shrinkage rate at the inlet ($y = 0$) is determined by the local foulant concentration and corresponding deposition rate there (which fixes $a(t)$); and (ii) a mass-conservation condition that ensures the net volume of deposited foulant particles over the entire pore equals the change in volume of the pore (which fixes $b(t)$). We also enforce the condition that $b(0) = 0$, i.e., that each pore is cylindrical at the beginning of the filtration process. This linear pore model, while still inexact, is much cheaper computationally than solving for the true nonlinear pore evolution, while substantially improving on the accuracy of the cylindrical pore approximation.

Using Eq. (6) as our pore radius function, we have the following expression for the conductance k:

$$k = \left(\int_0^l r^{-4}(y, t)dy\right)^{-1} = \frac{3a^3(t)b(t)}{1 - (a(t)/(a(t) + b(t)l))^3}, \tag{7}$$

where the first equality is the definition of conductance for an axisymmetric, variable-radius pore [16]. We can now also give an explicit version of Eq. (5):

$$c(y, t) = c_{\text{in}}(t)\exp\left(-\frac{\lambda}{q}\left(a(t)y + \frac{b(t)}{2}y^2\right)\right), \quad 0 \leqslant y \leqslant l. \tag{8}$$

Equation (8) suggests that the true pore radius profile should evolve as an exponential, since foulant particle concentration is distributed exponentially along the length of a pore. Clearly, our linear pore model is an approximation of the true pore radius evolution; however, since we constrain the pores to have small radii relative to their length, the decay of the exponential is small relative to the pore length, making its profile close to linear overall.

We note that the nondimensional affinity coefficient λ is the key parameter in our final fouling model. The choice of λ controls how deeply foulant particles penetrate the pore network: a larger value will lead to heavy fouling at the upstream side of the filter, corresponding to inefficient use of the filter and a short filter lifetime; a smaller value allows foulant particles to penetrate the filter more deeply, allowing for more efficient use of the filter and extended lifetime (possibly at the cost of more foulant particles escaping the filter). Following the work of Gu *et al.* [24], we set $\lambda = 5 \times 10^{-7}$ for all simulations in this paper. For further study of the effect of this parameter we refer the reader to earlier work by Sanaei *et al.* [16,28].

B. Pore networks

Now we extend the ideas described in Sec. II A to pore *networks*, coupled systems of many (possibly thousands of) pores, connected through junctions (see Appendix A for details on how our networks are generated). The vertices v_i in our network are enumerated, and we denote by $e_{ij} = (v_i, v_j)$ the edge joining vertices v_i and v_j; each edge corresponds to a pore in the network and each vertex to a pore junction; the set of all edges (pores) is denoted by E and the set of all vertices (junctions) by V. Flow through the network is simulated by solving the H-P equation (3) in each pore e_{ij}, with inlet pressure $p_{\text{in}} = p_i$ and outlet pressure $p_{\text{out}} = p_j$, and continuity of flux imposed at pore junctions (total incoming flux must equal total outgoing flux). This leads to a linear algebraic system for the pressure at each vertex, formulated in terms of the graph Laplacian (the discrete analog of the Laplacian). The subset of V consisting of pore inlets in the upstream membrane surface is denoted V_{in}, and the set of outlets in the downstream surface is V_{out}; the system is closed by imposing the pressures at the inlets $v_j \in V_{\text{in}}$ and outlets $v_j \in V_{\text{out}}$. With the fluid dynamics determined one can then solve for foulant concentration and pore shrinkage. For more details on how these calculations are performed, see Ref. [23] and Appendix A.

We examine two filtration performance metrics in this work. The first is *volumetric throughput* $h(t)$, which is the volume of filtered fluid (filtrate) that has passed through the membrane from the start of filtration up to any given time t. It is calculated by integrating the total flux q_{out} through all outlets $v_j \in V_{\text{out}}$, over time t:

$$h(t) = \int_0^t q_{\text{out}}(t')dt', \tag{9a}$$

$$q_{\text{out}}(t) = \sum_{v_j \in V_{\text{out}}} \sum_{v_i:(v_i,v_j)\in E} q_{ij}(t), \tag{9b}$$

where $q_{ij}(t)$ is the flux in pore e_{ij}. The second performance metric we study is *accumulated foulant concentration*, which is the number of foulant particles per unit volume of collected filtrate, accumulated up to time t:

$$c_{\text{acm}}(t) = \frac{\int_0^t c_{\text{out}}(t')q_{\text{out}}(t')dt'}{\int_0^t q_{\text{out}}(t')dt'}, \tag{10}$$

where

$$c_{\text{out}}(t) = \frac{\sum_{v_j\in V_{\text{out}}} \sum_{v_i:(v_i,v_j)\in E} c_j(t)q_{ij}(t)}{q_{\text{out}}(t)}. \tag{11}$$

Here, $c_{\text{out}}(t)$ is the instantaneous foulant concentration in the fluid exiting the filter at time t, and $c_j(t)$ is the foulant concentration at vertex v_j. In Sec. III, the time t at which we evaluate h and c_{acm} is chosen to be t_{final}, here defined as the time at which the flux through the filter falls to 1% of its value at $t = 0$ due to fouling, and the simulation is stopped. In this case we refer to h and c_{acm} as *total throughput* and *total accumulated concentration*, respectively—although in the following we often refer to them as simply *throughput* and *concentration*.

Previous research has identified certain pore network properties as being reliable indicators of filtration efficiency. Gu *et al.* [23,24] found that a very strong determinant of filter performance is *porosity* ϕ, the relative void volume of the membrane, i.e., the sum of the volumes of all the pores in a network, divided by the total volume of the filter in which the network is embedded. In terms of our nondimensional variables,

$$\phi = \pi \sum_{(v_i,v_j)\in E} r_{ij}^2 l_{ij}. \tag{12}$$

To remove confounding effects related to porosity variations, in this work we fix it to a specific value, $\phi = 0.6$, throughout, except the small-network results of Sec. III C). We note that the main features of the results do not depend on this particular choice for porosity; for example, we have carried out computations with much smaller value $\phi = 0.2$ and obtained consistent results.

To assign pore radii such that we can simultaneously mimic the pore size variations seen in real filters while keeping network porosity fixed, we first generate a random variable η_{ij} for each pore, then find a global scaling factor r_{scale} such that $r_{ij} = r_{\text{scale}}\eta_{ij}$ and $\phi = \phi_0$ (the prescribed porosity value). Given ϕ_0, we perform the following simple calculation to find the corresponding r_{scale}:

$$\phi_0 = \pi r_{\text{scale}}^2 \sum_{i,j} \eta_{ij}^2 l_{ij} \Rightarrow r_{\text{scale}} = \left[\frac{\phi_0}{\pi \sum_{i,j} \eta_{ij}^2 l_{ij}}\right]^{1/2}. \tag{13}$$

This is the procedure used to generate the radii underlying the results in Sec. III.

Note that in the earlier work, Gu *et al.* [24] drew the pore radii from a (truncated) uniform distribution, whereas the distribution of real membrane network pores is known to be skewed. The log-normal distribution is considered to be a suitable choice for such skewed distributions [29–31], so in the present work the random variables η_{ij} are drawn from a log-normal distribution. Further discussion on this point can be found in Appendix D.

Another property previously found to correlate rather strongly with filtration performance is *tortuosity* τ, the mean distance that a fluid particle in the feed travels as it transits the membrane (normalized by the membrane thickness) [15,23,24]. However, as discussed in Appendix D, tortuosity turns out not to be a good predictor of performance for the present setup which focuses (in contrast to the earlier works [15,23,24]) to a fixed porosity setup. Our understanding is that porosity value influences performance overwhelmingly, and therefore masks the effect of other filter descriptors, such as its geometry or topology.

Following Gu *et al.* [24], we consider two sources of randomness in network filters: *network variations*, the variations in particular network structures generated using the algorithm described in Appendix A; and *noise variations*, the variations in the initial random assignation of radii to a network's pores. To distinguish the effects of each of these types of variation we use two sets of data: one for the network variations, in which

1000 different networks are generated, each with different network structures and unique sets of (randomly assigned) pore radii; and one for the noise variations, in which a single "typical" network is chosen from the 1000 mentioned above, and to which 1000 different random sets of pore radii (from the same distribution) are assigned. The results in Sec. III will focus on network variations for the sake of brevity, but a comparison of network and noise variations is presented in Appendix D.

For the noise variations, Gu *et al.* [24] chose "typical" networks by finding the network whose porosity was closest to the mean; however, since we fix porosity, "typical" networks are here chosen according to their geometric and graph-theoretic properties. To this end, we define a vector u_k for each network G_k, given by

$$u_k := (|V_{\mathrm{in}}|, |V_{\mathrm{int}}|, |V_{\mathrm{out}}|, |E|, \mu_{\mathrm{deg}}, \mu_{\mathrm{len}})^{\mathrm{T}}, \tag{14}$$

where V_{in} is the set of nodes in the upper surface of G_k, V_{out} is the set of nodes in the lower surface of G_k, V_{int} is the set of nodes in the interior of G_k, μ_{deg} is the mean *degree* (the number of other vertices to which each vertex is connected) of the vertices in G_k, and μ_{len} is the mean length of G_k's pores. This vector is evaluated for all networks G_k, $k \in \{1, \dots, N\}$, the elementwise mean $\mu_X = \mathbb{E}[u_k]$ over all G_k is calculated, and finally the typical network is chosen as the one that minimizes the Euclidean distance $\|u_k - \mu_X\|_2$.

C. Computational topology measures

In this section we introduce PDs, a tool from TDA that can provide compact topological information for networks (see Appendix C for more information on PDs, and for a simple worked example showing how these diagrams are produced). In particular, we use the GUDHI library [32] to extract topological information from pore networks. To our knowledge, PDs have not yet been used to study pore networks in membrane filters; in this work we apply them to our initial, unfouled networks (in which the pores are circular cylinders), with the goal of identifying topological features that correlate with desired filtration outcomes. A rigorous mathematical description of PDs is beyond the scope of this work, therefore we give just a brief description of the relevant tools at hand. To proceed, we must first define a few key concepts.

The concept of persistence diagrams rests on a process called *thresholding*, where the features of a topological space X are selected by varying a *thresholding parameter* $\theta > 0$ (see Ref. [33] for more, including a rigorous definition of topological spaces; for our purposes, a topological space is just a pore network, and the thresholding parameter θ is taken to be the pore radius). This thresholding process produces a series of nesting subnetworks, defined here by $X_\theta = G(V_\theta, E_\theta)$, where $E_\theta \subset E$ is the set of pores $(v_i, v_j) \in E$ such that $r_{ij} \geqslant \theta$, and $V_\theta \subset V$ is the set of nodes $v_i, v_j \in V$ such that $(v_i, v_j) \in E_\theta$. In other words, X_θ is the subnetwork consisting of pores of radius at least θ, along with all the corresponding pore junctions.

In this work we use *superlevel thresholding*, where the pore radius thresholding parameter θ is decreased from infinity (in practice, some suitably large value). As it decreases, pores of radius greater than or equal to θ become topologically "visible" in the pore radius network. We can gain an intuition for this process if we imagine the appearance of a single network pore (the one with the largest radius), then the appearance of another pore (the next-largest), and so on as we decrease θ, until we recover the entire network.

We study two types of network topological features that appear under thresholding (both of which we refer to by the generic term *generator*): *connected components* and *loops*. Connected components can be thought of as subnetworks of the larger network, where each vertex can be reached by traversing a path from any other vertex in the subnetwork (i.e., the subnetwork is *path-connected* at the given threshold value). Loops (informally speaking) are paths through the network that begin and end at the same vertex. In algebraic topology, connected components are zero-dimensional objects, and loops are one-dimensional objects. Since we are working in three dimensions, one might expect that we also study the two-dimensional topological objects known as *cavities*. Cavities exist when a three-dimensional volume is *entirely* enclosed by some two-dimensional surface (imagine a hole in a typical Swiss cheese-like structure). However, within the confines of our pore network model cavities do not exist.

Connected components are *born* when a pore "appears" (i.e., when the threshold parameter θ reaches the value of that pore's radius) that is unconnected to any existing connected component. The value of $\theta = b$ at which this occurs is called a *birth number*. We also say that a connected component *dies* at *death number* $\theta = d$ when a pore appears that connects it to an older, preexisting connected component. The preexisting connected component absorbs the now-dead one. The difference $\mathcal{L} = b - d$ is called the *lifespan* of a connected component (or any other type of generator). The larger the lifespan of the generator under thresholding, the more that generator *persists*, hence the name "persistence diagram." The more persistent the generator, the more topologically significant it is. Conversely, when the lifespan of a generator is small, that generator is considered to be topological "noise."

The case of loops is somewhat different from that of connected components: a loop is born when a pore appears that completes a path from a vertex to itself. However, loops do not die, since in the present problem a loop is never fully absorbed into another loop, so loop death numbers are therefore always zero.

Having established these ideas, we can describe in more detail the concept of PDs, which plot the birth and death numbers of the generators of a network. By convention, PD_i refers to the PD depicting the ith-dimensional topological features, i.e., PD_0 is the PD for connected components, and PD_1 is the PD for loops. A worked example of these two PDs for a toy network is given in Appendix C.

We note that a loop is similar to the graph-theoretic concept of a *cycle*, which is also a path through a network that starts and ends at the same vertex. Although the two concepts may seem identical according to our informal description, the number of loops (generators) in a one-dimensional PD may not be equal to the number of cycles in a network because loops are technically linearly independent vectors; the set of all loops spans the vector space in which they exist. The set of loops is therefore a subset of the set of all cycles, which are linear combinations of loops. We will denote the number of loops as NG_1, which represents the number of generators in PD_1; more generally, NG_i is the number of generators in PD_i.

One metric for the data provided by PDs is *total persistence* (TP), a measure of the overall significance of the topological data encoded in a PD. Recall that the "persistence" of a generator in a PD is equivalent to its lifespan under thresholding. The total persistence of a PD is thus a measure of the combined lifespans of the generators, given by the following formula:

$$\mathrm{TP}_i = \left[\sum_{k=1}^{\mathrm{NG}_i} (b_k - d_k)^p\right]^{1/p}, \tag{15}$$

where $i = 0$ denotes connected components and $i = 1$ denotes loops. Thus, TP is the p-norm of the difference between the vector of birth numbers and the vector of death numbers. Typically p is chosen to be 1, and this is the value that we will use. Total persistence is admittedly a crude measure, since a large TP value may be due to a single long-lived generator, or many short-lived generators. However, TP offers a simple and readily available method of quantifying aspects of network topology and is therefore a natural measure to investigate.

Since loops never die, the above equation implies the following for TP_1:

$$\mathrm{TP}_1 = \sum_{k=1}^{\mathrm{NG}_1} (b_k - 0) = \sum_{k=1}^{\mathrm{NG}_1} b_k, \tag{16}$$

that is, TP_1 is just the sum of the birth numbers b_k.

Finally, one further measure that we will use in analyzing our results is the *mean lifespan* $\overline{\mathcal{L}_i}$ of generators in PD_i, which is defined as the mean of the lifespans of the generators in a given PD_i. Mean birth and death numbers of generators, $\overline{b}$ and $\overline{d}$, respectively, are similarly defined. Unless stated otherwise, when we refer to *mean* quantities, these quantities are always averaged over each network (or corresponding PD_i) separately.

III. RESULTS

We are now ready to discuss our findings regarding the effects of (i) the initial distribution of the pore radii, and (ii) the network topology, on filtration performance. One crucial issue that we focus on is the relationship between a filter network's geometrical features (e.g., the radii and lengths of its pores, and the number of pores per network, denoted as NP) and its topology. In particular, we consider the following question: Are topological features more predictive of filtration performance than geometrical ones? The answer to this question will depend on whether measures of network topology reduce in some way to measures of network geometry, in which case we need not concern ourselves with topology at all.

In Sec. III A we first discuss connected component-based measures and show that these are not predictive of filtration performance. Section III B focuses on the effect of loop-based measures on filtration performance; here we find dramatically different results. To gain further insight into the role of loops, in Sec. III C we investigate networks with a small number of pores NP, in which all possible networks and loops can be enumerated.

To focus on the questions and issues discussed above, we relegate some discussion to Appendixes. In particular, discussion related to the influence of the choice of pore radius distribution, which appears to be significant, is presented in Appendix D; since filters in applications are typically characterized by a radius distribution close to log-normal, we focus the discussion in the main text exclusively on log-normal radii distributions. Furthermore, we defer discussion of the role of noise variations to Appendix D, and in the main text focus on considering pore network variability. Discussion of the importance of the extent to which pore network radii vary is also relegated to Appendix E; in the main text we focus only on what we call strongly perturbed networks, where the pore radii vary considerably. The discussion in Appendix D also shows that, for filters of fixed porosity characterized by network variations, tortuosity is a not a good predictive measure; therefore in the main text we do not discuss this measure further.

A. Filtration performance: Connected component-based measures

We first discuss the influence of the topology described by the persistence of connected components on the filtration performance; for brevity we focus on (total) throughput h as a representative performance measure. In particular, we discuss how the number of generators NG_0, the mean lifespan of connected components $\overline{\mathcal{L}_0}$, the total persistence TP_0, and the mean birth and death numbers of the generators, $\overline{b}$ and $\overline{d}$, correlate with throughput (where $\overline{\mathcal{L}_0}$, $\overline{b}$, and $\overline{d}$ are averaged for each network separately). Figure 4 shows representative results, where each data point corresponds to a unique network. This figure suggests that connected components are largely ineffective for predicting filtration performance, as neither NG_0, nor $\overline{\mathcal{L}_0}$, nor TP_0 show clear correlation with h.

The results shown in Fig. 5 appear more promising at first sight: here we plot throughput h against mean birth and death numbers, $\overline{b}$ and $\overline{d}$, respectively. These results suggest a modest linear correlation for both quantities. However, further investigation reveals that $\overline{b}$ and $\overline{d}$ are each strongly correlated with mean pore radius (denoted by $\overline{r}$), a geometric quantity; see Fig. 5(c). This suggests that the modest signal that birth and death numbers provide in Figs. 5(a) and 5(b) most likely reduces to a purely geometrical phenomenon. Therefore our main question is still open: is the topology of a network relevant in determining filtration performance? As we discuss next, additional insight can be obtained by considering topological measures describing the loop structure.

B. Filtration performance: Loop-based measures

Next, we study the influence of the topology described by the persistence of loops on filtration performance. Figure 6 plots throughput h against NG_1, $\overline{\mathcal{L}_1}$, and TP_1, respectively. In contrast to the results for the connected components seen in Sec. III A, here we observe a strong correlation between h and all three topological measures. Specifically, h is a decreasing function of both NG_1 and TP_1, and an increasing function of

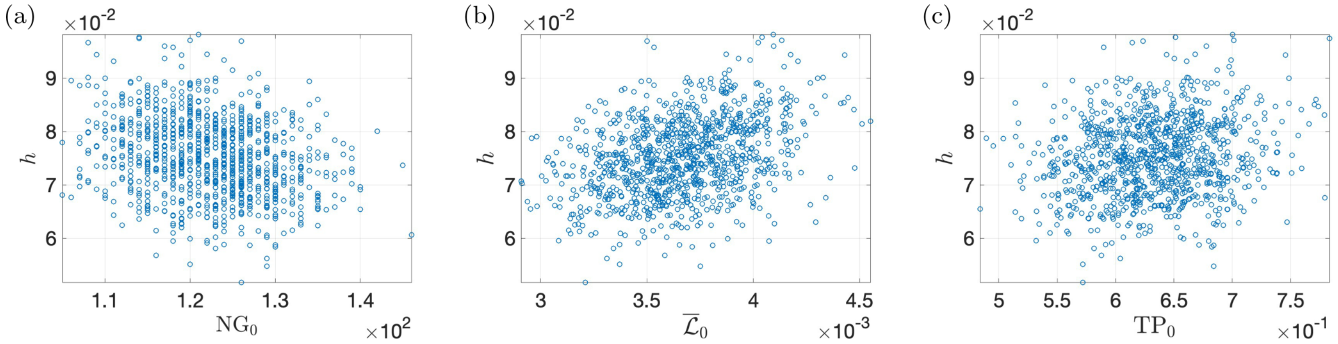

FIG. 4. Throughput h vs number of connected components NG_0 (a), mean lifespan of connected components $\overline{\mathcal{L}}_0$ (b), and total persistence of connected components TP_0 (c).

$\overline{\mathcal{L}}_1$. Figure 6(a) in particular constitutes the first key finding of this paper, since it depicts a very strong result: h is correlated with a purely topological property of filter networks (NG_1), and therefore filtration performance is indeed correlated with filter network topology. Encouraged by these findings, we focus henceforth on the investigation and discussion of the loop properties.

Cautioned by the findings for connected components above, we first address the possible mitigating effect of non-topological network properties, which might underlie the apparent correlation between filtration performance and network topology. In particular, we discuss whether the observed correlation between throughput and loop-based measures can be reduced to a purely geometric effect, or whether different network connectivity leads to different throughput, even if geometric measures are the same or similar.

Figure 7 shows that there is, in fact, a fairly strong correlation between h and each of the four geometric quantities: $\overline{r}$, mean pore resistance, mean pore area, and NP. For the most part, these relationships are as expected; for example, h is positively correlated with $\overline{r}$ [Fig. 7(a)], which can be explained by the argument that more fluid can flow through a wider pore, and wider pores take longer to close under fouling, thus increasing h. Similarly, we expect h to be negatively correlated with mean pore resistance [Fig. 7(b)], since a higher resistance decreases the flux through a pore under fixed pressure drop.

Figure 7(c) depicts the correlation between h and the mean inner surface area of the pores, per network. The relationship depicted here is the only one for which the sign is counterintuitive, since one might expect that an increase in a pore's inner surface area available for particle deposition would increase the probability of particle fouling, thus decreasing the flux through the pore more rapidly; however, with larger mean pore area, the deposition of a foulant particle has a smaller effect on fluid flow than if the mean pore area were smaller, leading to a slower increase in network resistance, and the observed positive correlation. The positive correlation between h and mean inner surface area is consistent with h being negatively correlated with mean pore resistance, since a high throughput h implies a low resistance network, which in turn implies that more particles will be advected through the network without deposition occurring. Finally, Fig. 7(d) shows a negative correlation between h and the total number of pores in the network, NP, a point we will return to below.

Overall, Fig. 7 shows that network geometry does influence filtration performance. Given the conclusions of Sec. III A, this raises the question of whether there remains some independent topological effect as well. To answer this, we first investigate the potentially confounding role of mean pore radius $\overline{r}$ in the relationship between h and NG_1, by dividing h by $\overline{r}$ (to control for its effect) and plotting the result against NG_1. The result is depicted in Fig. 8(a): there is a (negative) correlation between $h/\overline{r}$ and NG_1, thus we conclude that the effect of NG_1 cannot be reduced to a purely geometrical phenomenon. *Therefore, two networks of the same geometry (characterized here by* $\overline{r}$*), but different topology, could lead to significantly different filtration performance.*

A brief comment here on filtration performance metrics is appropriate: as noted earlier, we have so far focused on results

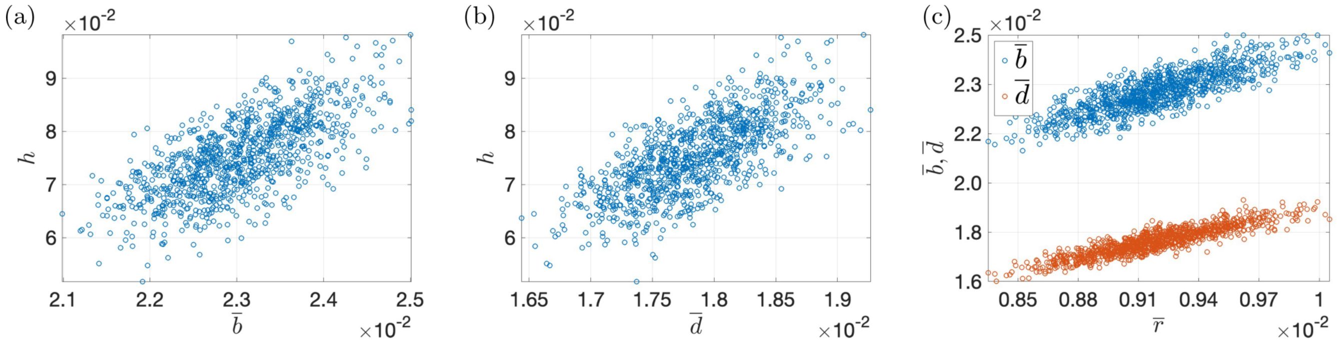

FIG. 5. Throughput h vs (a) mean birth number $\overline{b}$ and (b) mean death number $\overline{d}$ for connected components. Panel (c) depicts $\overline{b}$ (blue) and $\overline{d}$ (red) vs mean pore radius $\overline{r}$.

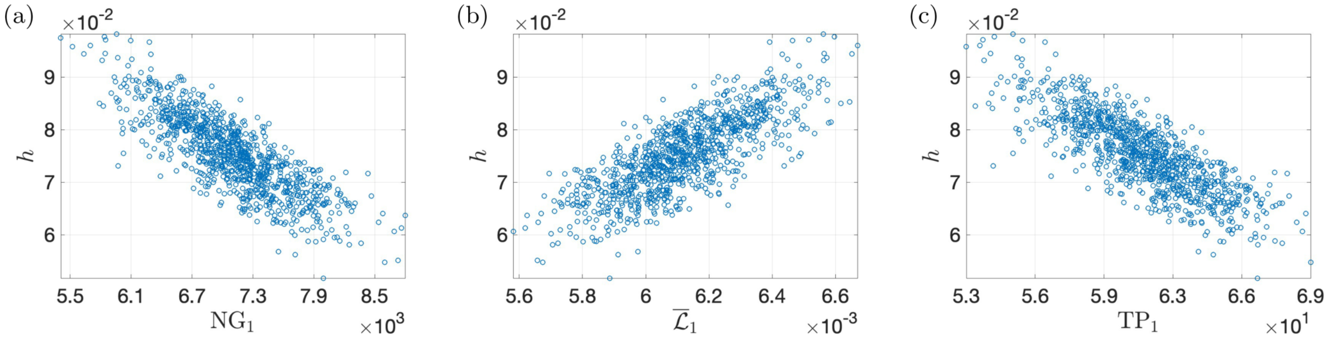

FIG. 6. Throughput h vs (a) number of loops NG_1, (b) mean lifespan of loops $\overline{\mathcal{L}}_1$, and (c) total persistence of loops TP_1.

using throughput h as a proxy for filter performance, with no regard for how clean the resulting filtrate is (measured by the quantity c_{acm}, accumulated foulant concentration). We have done this because all of the results for h presented thus far are very similar to the results for c_{acm} (not shown). Figure 8(b) exemplifies the similarity between the h results and the c_{acm} results: $c_{\mathrm{acm}}/\bar{r}$ is plotted against NG_1, and shows the same correlation as for h.

Figure 8(c) shows h versus c_{acm}, both scaled by mean radius $\bar{r}$. The main finding here is that the more filtrate one has, the more escaped foulant particles *per volume* one has as well, at least for the considered fixed-porosity setup. Bearing in mind that we want filters that find an optimal balance between h and c_{acm}, we continue the discussion of our results by focusing on h henceforth.

Having determined that $\bar{r}$ is not the sole geometric factor behind the correlation of performance with NG_1, we turn next to the following question: is NG_1 still correlated with h if we fix the number of pores NP? This question is motivated by the strong similarity between the two quantities' respective correlations with h; compare Figs. 6(a) and 7(d). Together, these suggest that the apparent relationship between h and NG_1 may reduce to a phenomenon involving NP. The rest of this section is devoted to clarifying these inter-dependencies.

First, we examine the negative correlation between h and NP depicted in Fig. 7(d). To begin to understand this relationship, in Fig. 9(a), we plot $\bar{r}$ against NP. This figure reveals strong negative correlation, which may be explained by observing that, for networks of fixed porosity, a large number of pores (high NP) means that either $\bar{r}$ or mean pore length (or both) must be small. Thus, assuming that mean pore length remains the same, a large NP implies a small $\bar{r}$; smaller pores close faster under fouling, which leads to a decrease in h, which explains the negative relationship between h and NP.

Next, we examine the correlation between NG_1 and NP. Figure 9(b) plots NG_1 against NP for all networks, and shows a strong positive correlation, with nearly unit slope, between these quantities. This indicates that, if the number of pores in a given volume of space is high enough, then on average, approximately one new loop is created (i.e., a new vector is added to the loop vector space, increasing its dimension by 1) for each new pore added to the network. This may explain why the relationship between NG_1 and h is almost identical to the relationship between NP and h, but it does not indicate whether the effect of NG_1 reduces to some effect of NP. To differentiate between the effects of NG_1 and NP, we follow the approach used in Fig. 8(a), but here we scale h by NP and and plot the result against NG_1; see Fig. 10. This figure constitutes another key finding of this paper, as it shows that the correlation is actually strengthened by controlling for NP, definitively confirming that NG_1 has a strong effect on h, which is distinct from the effect of NP. *In other words, there is indeed a purely topological effect on filtration performance, which does not depend on the effect of geometrical features.*

We also find that by assuming (different) linear fits between h and NG_1, and h and NP, we can find by straightforward calculation a nonlinear equation for h/NP vs NG_1, which is also shown in Fig. 10. At this point, we do not have any particular interpretation of the numerical values of the fitting coeffcients.

To gain further insight into this influence of NG_1 on h, distinct from the effect of NP, in the next section we examine small networks (characterized by small NP), which we use to compare the performance of filters with loops to filters without loops.

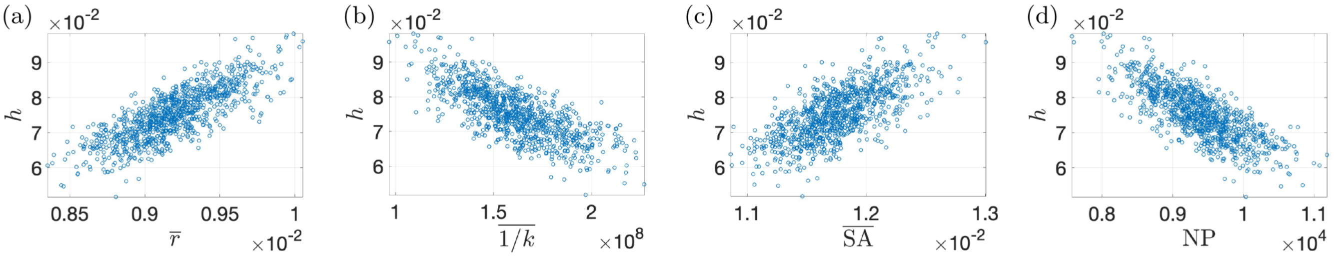

FIG. 7. Throughput h vs (a) mean pore radius $\bar{r}$, (b) mean pore resistance $\overline{1/k}$, (c) mean pore area $\overline{\mathrm{SA}}$, and (d) number of pores NP.

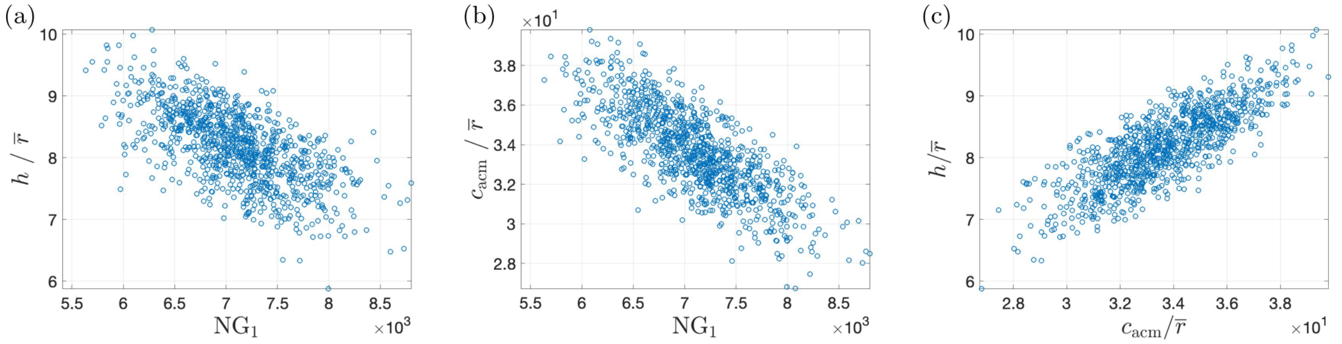

FIG. 8. (a) Throughput h and (b) concentration c_{acm}, scaled by mean pore radius $\bar{r}$, respectively, vs number of loops NG_1; (c) $h/\bar{r}$ vs $c_{\mathrm{acm}}/\bar{r}$.

C. Small networks

Here, we discuss the results of filtration simulations on small networks, whose respective NP values (recall that NP denotes "number of pores") are small integers, and which may or may not contain loops. Those networks that do contain loops we call *loopy* networks (multiply connected, topologically speaking), and those that do not we call *loopless* (simply connected). Our goal is to investigate the effect of (small) NG_1 on h by comparing with results for networks with $\mathrm{NG}_1 = 0$, to gain insight into the role of loops in full-size networks.

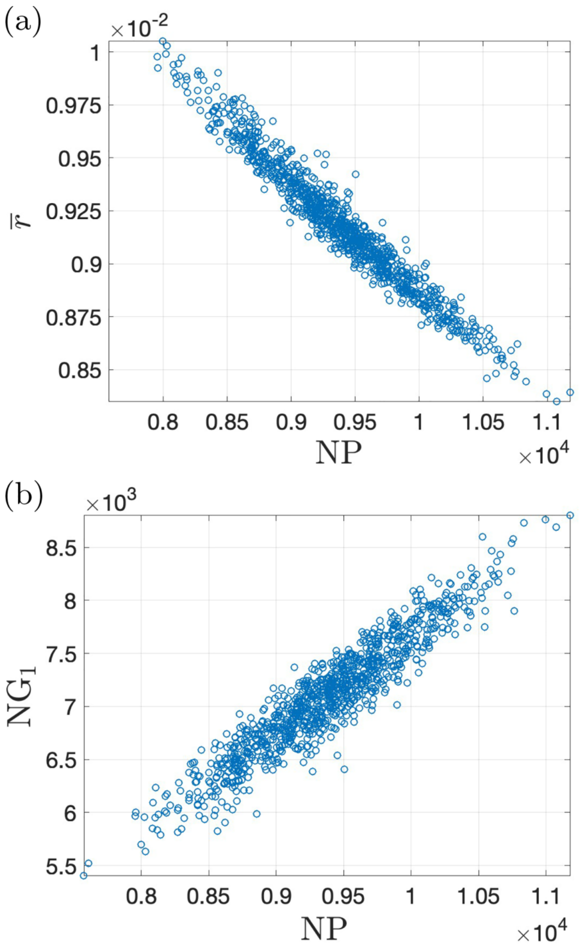

FIG. 9. (a) Mean pore radius $\bar{r}$ vs number of pores NP. (b) Number of loops NG_1 vs number of pores NP.

Here we consider all possible small networks, subject to certain constraints. First, each network must have $\mathrm{NP} \geqslant 5$, since five-pore networks are the smallest possible networks with loops that consist only of interior pores (smaller loopy networks require two pores to come from the same surface inlet or outlet, a zero-probability occurrence in our full-size networks). Further, for practical reasons, since we consider all possible networks we restrict attention to $\mathrm{NP} \leqslant 9$ (the number of possible networks grows combinatorially with NP, making larger networks too computationally expensive to investigate). We also set all pore lengths to 1 (equal to the membrane thickness), regardless of NP, to remove the effect of pore length. As with the full-size networks, porosity is fixed by scaling the radii uniformly. Keeping porosity fixed for all small networks considered, and using a constant radius of $r_5 = 0.01$ for five-pore networks, we derive the following formula for the radius of a network with NP pores:

$$r_{\mathrm{NP}} = r_5 \sqrt{\frac{5}{\mathrm{NP}}}. \tag{17}$$

Figure 11 depicts a sample of the networks with NP = 6 used in this work (schematic representation only; the pores are not drawn to scale, and real networks may be three-dimensional).

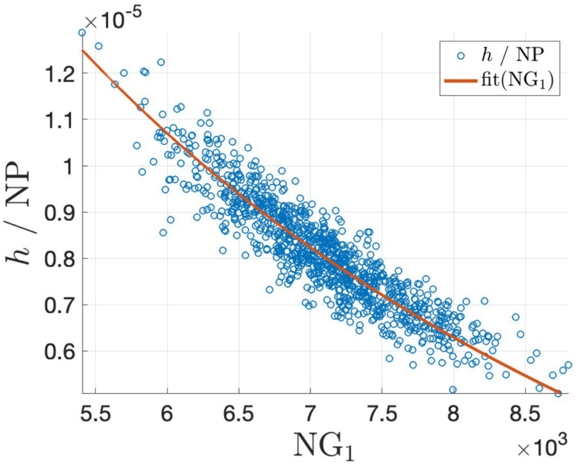

FIG. 10. Throughput h, scaled by number of pores NP, vs number of loops NG_1. The red curve depicts the nonlinear fit $h/\mathrm{NP} = \mathrm{fit(NG)}_1$, obtained as described in the text, which takes the form $\mathrm{fit}(\mathrm{NG}_1) = A + B(\mathrm{NG}_1 + C)^{-1}$, where $A = -9.85 \times 10^{-6}$, $B = 0.162$, and $C = 1789$.

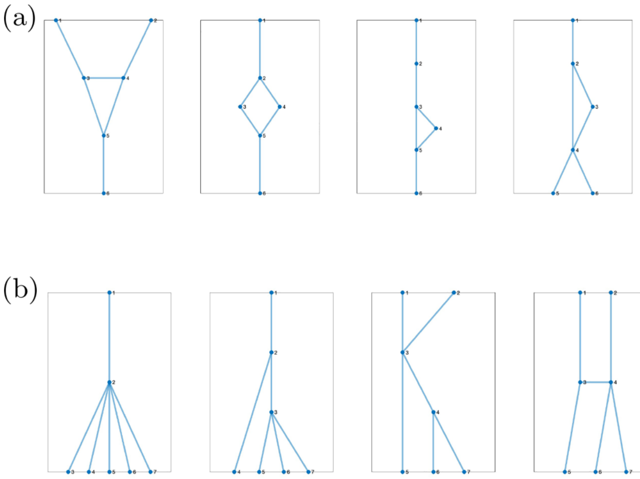

FIG. 11. A sample of the toy networks with NP = 6 used in this section, with (a) loopy networks and (b) loopless networks. (Images are schematic only; actual networks have all pore lengths equal to membrane thickness and may be three-dimensional).

Having specified the networks, we can now discuss the results. To investigate the effect of a network's loops on its filtration performance, we compare the correlation between throughput h and the number of inlets $|V_{\text{in}}|$ in the loopy and loopless cases. $|V_{\text{in}}|$ is chosen for this investigation because we found it to be strongly predictive of h in the case of small networks.

Figures 12(a) and 12(b) plot h against $|V_{\text{in}}|$ for loopy and loopless networks, respectively, where each point represents a single network. Note that the vast majority of the points in these figures overlap, making it seem as if relatively few networks are represented. We observe that in Fig. 12(a), the range of h values is consistently lower than in Fig. 12(b), for all values of NP. This suggests that h is more strongly dependent on the existence of loops than it is on NP. We also observe that there is a positive linear relationship between h and $|V_{\text{in}}|$. This conforms to our expectation that a higher number of inlets will lead to a higher value of h, since more fluid can enter and flow through the filter over its lifetime.

Note that this strong dependence of h on $|V_{\text{in}}|$ does not exist for large networks (results omitted for brevity). We assume this is because, for large networks, adding or removing an inlet pore only changes $|V_{\text{in}}|$ by a very small percentage, so the overall effect of $|V_{\text{in}}|$ on h is removed.

Figures 12(a) and 12(b) also show that h is negatively correlated with NP for both loopy and loopless networks, as the networks with fewer pores tend to yield the highest values of h. This negative correlation between h and NP invites further investigation, but we believe it to be due to the fact that, within the present setup, if a network has more pores then foulant particles will stay in the network for longer, increasing the chance of deposition, which in turn decreases throughput.

Figures 12(a) and 12(b) bring us back to the central question of whether, given NP, the number of loops NG_1 influences h. To investigate this question further, we average h separately over all loopy and loopless networks with the same NP, and plot the results against NP in Fig. 12(c) (the error bars indicate standard deviations). Given the size of the error bars, further study is needed to strengthen the results presented in Fig. 12(c). However, this figure suggests that mean throughput, $\overline{h}$, is always smaller for loopy networks than for loopless networks, regardless of NP. Although there appears to be a very small negative correlation of $\overline{h}$ with NP in the loopless case, the effect is so small as to be negligible. The lack of any significant dependence of $\overline{h}$ on NP in either case suggests not only that NG_1 has an effect on h that is distinct from the effect of NP, but that NG_1 dominates over the effect of NP.

The question of why the existence of loops has this dominating effect remains to be answered. Nevertheless, here we have found that NG_1, a purely topological property of membrane filter networks, has an effect on filtration performance that is not reducible to network geometry. We have also shown that the number of inlets is a strong predictor of h in the case of small networks, although this effect does not persist for large networks.

IV. CONCLUSIONS

The primary focus of this work is to answer a simple question: How does a filter's pore structure affect its performance? Since answering this in full generality is difficult, we focus on the specific case of filters made from a network of interconnected pores. Previous work shows that porosity plays a crucial role [23], and since its effects are reasonably well understood, in the present work we assume that the porosity is fixed. We then ask a more specific question: given

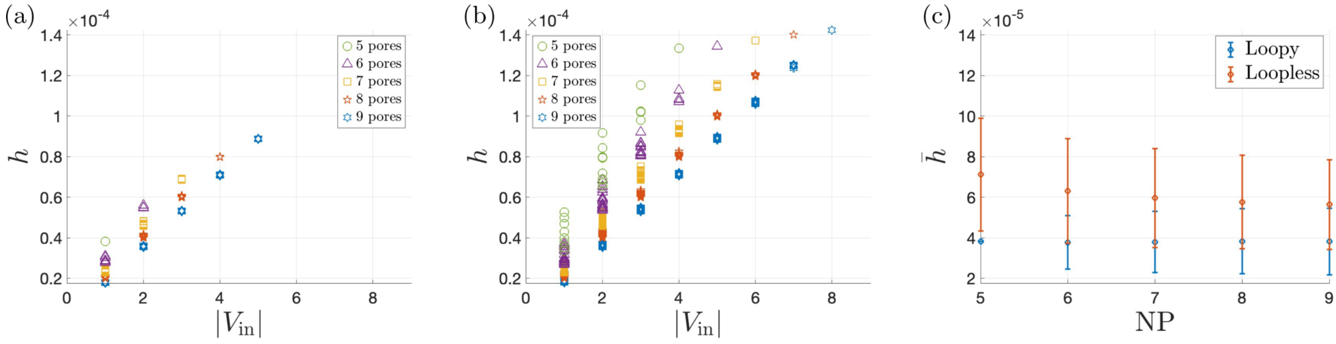

FIG. 12. Throughput h results for small networks, with pore numbers $\text{NP} \in \{5, \ldots, 9\}$. (a) Throughput h vs number of inlets $|V_{\text{in}}|$ for loopy networks; (b) throughput h vs number of inlets $|V_{\text{in}}|$ for loopless networks; (c) mean throughput $\overline{h}$ vs NP for both loopy and loopless networks.

a filter network with a known porosity, what determines its performance? To narrow the focus further, we consider filters where pore radii follow a log-normal distribution, which is known to be among the most realistic. For completeness, the effects of different distribution types and widths are discussed in Appendix D. With this focus, and assuming the pores are long and slender (permitting a simplified Hagen-Poiseuille fluid dynamics model that allows us efficiently to run many simulations with thousands of network realizations), we are prepared to answer the main question posed above.

The structure of a filter can be viewed from several different perspectives, focusing on various parameters such as the number of pores, their radii and lengths, and how they connect. Another (more complex but commonly used) measure is tortuosity, which quantifies the typical length of the path taken by the filtrate through the filter. We find that when porosity is fixed, tortuosity is not a useful performance measure. Therefore, we consider the simpler measures mentioned above, which can be roughly categorized into two groups: geometrical measures, such as pore count and size, and topological measures that describe how pores connect within the filter. While there are many ways to quantify the topology of a network, our setup—where pore radii serve as weights—can be analyzed efficiently using persistent homology (PH). The advantage of this approach over others focusing on percolation or coordination number (among others) is that weights are naturally integrated. PH measures the network's connectivity across all pore radii simultaneously and is applicable in both two and three dimensions.

Regarding the filter's throughput as a performance metric, our main findings are as follows. First, geometry influences performance: increasing the number of pores generally decreases throughput. This conclusion is statistical, given the variability across different networks. Our ability to analyze thousands of network realizations gives us confidence that this finding is widely applicable. Next, we ask: given a fixed number of pores, does the network's topology matter? Our second key result is a positive answer to this question: the topology has a significant effect on performance. Specifically, the presence of loops within the network is crucial; a higher number of loops correlates with lower throughput. Thus, both the geometry and topology of the network impact performance. Although the precise underlying reason remains unclear (in particular regarding the importance of loops), we have obtained some preliminary understanding by examining small networks with a limited number of pores.

Our study also examines additional factors, including pore radius distribution and variability, among others. These results are detailed in the extensive Appendixes, which include examples and further explanations of persistent homology measures.

While our work offers important new insights into how filter structure influences performance, many questions remain and much more could be done. The measures analyzed here—either topologically or geometrically—are based solely on the initial filter configuration. How these measures evolve over time will be one focus of future research. Finally, we also plan to investigate the positive, linear correlation between total throughput and total accumulated concentration, which indicates that, given the constraints imposed in this work, a filter that processes more fluid also produces dirtier filtrate.

ACKNOWLEDGMENTS

This work was supported by the NSF Grant No. DMS-2201627. The authors gratefully acknowledge several illuminating discussions with Dr. Jonathan Jaquette (NJIT) and with Dr. Uwe Beuscher (W.L. Gore & Associates).

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: NETWORK GENERATION AND SIMULATION METHODOLOGY

Here we present brief details of the methods used to generate the networks and simulate fluid filtration through them. For the network generation, following Gu *et al.* [24], the following algorithm is used:

(1) First generate N_{total} uniformly distributed points in a rectangular box $(x, y, z) \in [0, 0, 0] \times [1, 1, 2]$. In this work, we choose $N_{\text{total}} = 775$ (for full-size networks).

(2) To form pores, connect each point to all neighboring points within a spherical annulus of outer radius $d_{\max}$ (maximum pore length) and inner radius $d_{\min}$ (minimum pore length). The inner radius is chosen to ensure that the pore's aspect ratio (radius/length) is small enough for the Hagen-Poiseuille (H-P) approximation, described in Sec. II A.

(3) Cut the box at heights $z = 1/2$ and $z = 3/2$. The intersections of these two planes with the network edges form the inlets V_{in} and outlets V_{out} of the network. See Fig. 13 for a 2D schematic illustrating this process. This "box-cutting procedure" avoids the issue of having to specify the numbers of inlets and outlets in the membrane surfaces.

(4) Remove all inlets and outlets with length smaller than $d_{\min}$. The existence of these short pores is due to the box-cutting procedure, which does not account for the lengths of the inlets and outlets it creates.

(5) Finally, we assign an initial radius r_0 to each pore, and then we perturb randomly the respective radii of the pores according to some probability distribution. The choice of pore radius distribution is discussed in Appendix D.

To simulate filtration through the network, we apply conservation of flux at each pore junction v_i to transform the set of H-P equations for each pore to a linear algebraic system for the pressure $p(v_i)$ at each interior pore junction, formulated as the following problem:

$$(L_{\text{k}} p)(v_i) = 0, \quad v_i \in V_{\text{int}}, \tag{A1a}$$

$$p(v) = 1, \quad v \in V_{\text{in}}; \quad p(v) = 0, \quad v \in V_{\text{out}}. \tag{A1b}$$

Here, L_{k} is a $|V_{\text{int}}| \times |V_{\text{int}}|$ matrix known as the (interior) *graph Laplacian*, i.e., the discrete analog of the continuous

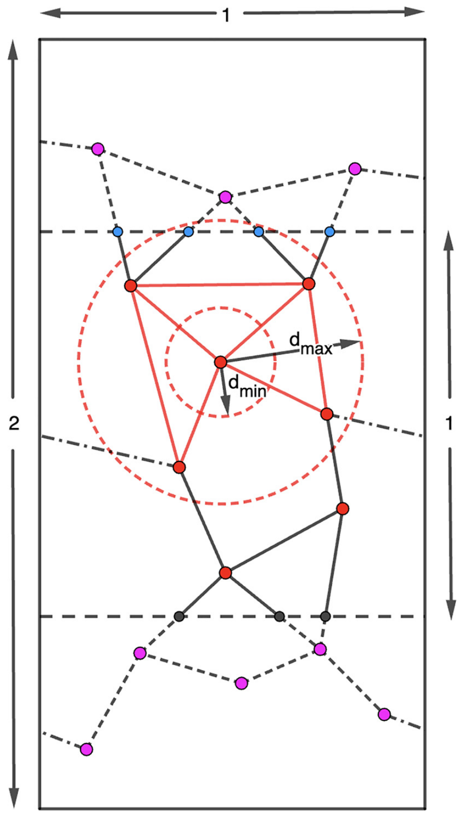

FIG. 13. A 2D schematic of the network generation protocol. The blue dots are the top surface pore inlets, the red dots are interior pore junctions, the black dots are the bottom surface pore outlets, the pink dots are pore junctions that get removed by the box-cutting procedure, and $d_{\min}$ ($d_{\max}$) is the minimum (maximum) pore length. The dashed edges in the network represent pores that were deleted by the box-cutting procedure, and dotted-dashed lines represent periodic connections.

Laplacian, weighted by the conductances k_{ij} of the networks' pores. In general, a weighted graph Laplacian L_W is given by

$$L_W = D_W - W, \tag{A2}$$

where D is the W-weighted *degree matrix* of a network, a diagonal matrix given by

$$D_{ij} = \begin{cases} \sum_{k=1}^{|V|} W_{ik}, & i = j, \\ 0, & \text{otherwise}. \end{cases} \tag{A3}$$

W is the W-weighted *adjacency matrix*, a matrix with non-negative entries W_{ij} if $(v_i, v_j) \in E$. We illustrate the concept of the weighted graph Laplacian using the simple network depicted in Fig. 14 as an example, where $W_{ij} = k_{ij}$ and k_{ij} is the conductance of pore (v_i, v_j):

$$L_k = \begin{pmatrix} k_{12} + k_{23} + k_{24} & -k_{23} & -k_{24} \\ -k_{23} & k_{23} + k_{34} & -k_{34} \\ -k_{24} & -k_{34} & k_{24} + k_{34} + k_{45} \end{pmatrix}. \tag{A4}$$

Once the pressures $p(v_i)$ at the interior pore junctions are calculated, they are then used to construct a matrix Q of the fluxes through each pore, using Eq. (3).

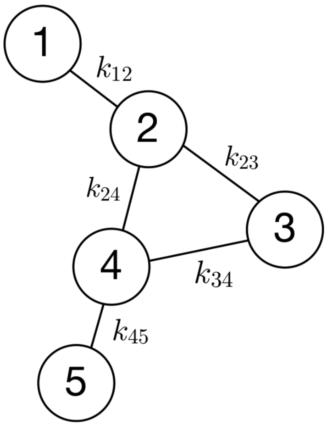

FIG. 14. A simple network to illustrate its graph Laplacian per Eq. (A4). Nodes 1 and 5 are boundary nodes at the upper and lower surfaces of the filter, respectively, at which pressure is prescribed. Nodes 2, 3, and 4 are interior nodes whose pressure we solve Eq. (A1) for. k_{ij} is the conductance of pore (v_i, v_j).

Next, the foulant concentration $c(v_i)$ at each interior pore junction is computed by solving an inhomogeneous version of Eq. (A1a), modified to model the advection and adsorption of foulant particles. The problem is given by

$$L_q^{\text{in}} c = s^T c_0, \tag{A5a}$$

$$s_{ij} = q_{ij} \exp\left(\frac{-\lambda \int_0^{l_{ij}} r_{ij}(y,t)dy}{q_{ij}}\right), \tag{A5b}$$

$$c_0 = (1, \ldots, 1, 0, \ldots, 0)^T, \tag{A5c}$$

where s_{ij} captures the advection and adsorption of foulant particles at the exit of the pore (v_i, v_j), and c_0 is a boundary condition specifying that $c(v) = 1$ for $v \in V_{\text{in}}$. The matrix $L_q^{\text{in}} = D_q^{\text{T}} - s^{\text{T}}$ is the q-weighted *in-degree advection Laplacian*, where D_q^{T} is a q^{T}-weighted degree matrix that encodes the *incoming* flux q_{ij} through pore inlets (thus including the direction of the flow through the pores), which is why we call L_q^{in} the "in-degree" Laplacian. We solve this linear system to obtain the concentration c_i at each vertex v_i. Finally, concentration is used to update the pore radii, using the model for advection and adsorption of foulant particles given in Sec. II A. More detail on the derivation of Eqs. (A1) and (A5) and the process of solving them on large networks can be found in Ref. [23].

APPENDIX B: MASS CONSERVATION FOR PORE EVOLUTION

We assume that pores have a straight axis with local length coordinate y, about which the pore profile is circular, with variable radius $r(y, t)$. In this work, we introduce a mass conservation scheme, which can be used for any pore radius function $r(y, t)$, though we implement it only for linear pore profiles. In the following, we consider a single pore, inside which foulant particles are deposited. The deposition of the particles causes a decrease in the volume of the pore, and this decrease is exactly equal to the sum of the effective volumes of the foulant particles. Since we assume the pores maintain their circular cross-section as they shrink, we implicitly assume that deposited particles distribute their volume uniformly onto the pore wall. This relation between particle and pore volumes is represented by a conservation equation, in which the left-hand side represents the volume of the deposited foulant particles, and the right-hand side the change in volume of the pore:

$$\int_0^l r(y,t)c(y,t)dy = -\int_0^l r(y,t)\frac{\partial r}{\partial t}(y,t)dy. \tag{B1}$$

Using the expression for $c(y,t)$ given in Eq. (5), we can evaluate the left-hand side of Eq. (B1), which will be denoted by $c_{\text{dep}}(t)$:

$$c_{\text{dep}}(t) := \frac{c_{\text{in}}(t)q(t)}{\lambda}\left[1 - \exp\left(-\frac{\lambda}{q(t)}\int_0^l r(z,t)dz\right)\right], \tag{B2}$$

where $c_{\text{in}}(t) = c(0,t)$, as defined in Eq. (4b). We can also use Eq. (B1) to derive an ODE relating pore radius evolution and foulant concentration:

$$\int_0^l r(y,t)\left[\frac{\partial r}{\partial t}(y,t) + c(y,t)\right]dy = 0 \tag{B3a}$$

$$\Rightarrow \frac{\partial r}{\partial t}(y,t) = -c(y,t). \tag{B3b}$$

Equation (B3b) gives a general formula for radius evolution, valid for any suitable radius function $r(y,t)$.

We now focus on the specific case in which a linearly varying pore radius is assumed, given by Eq. (6), $r(y,t) = a(t) + b(t)y$. Equation (B1) then becomes

$$c_{\text{dep}}(t) = \left[a(t)\dot{a}(t)l + \dot{a}(t)b(t)\frac{l^2}{2} + \dot{b}(t)\left(a(t)\frac{l^2}{2} + b(t)\frac{l^3}{3}\right)\right]. \tag{B4}$$

Equation (B3b) implies that $\dot{a}(t) = \dot{r}(0,t) = -c_{\text{in}}(t)$. Using this in Eq. (B4) and rearranging to solve for $\dot{b}(t)$, we arrive at the following system:

$$\dot{a}(t) = -c_{\text{in}}(t), \tag{B5a}$$

$$\dot{b}(t) = \frac{c_{\text{in}}(t)l(6a(t) + 3b(t)l) - 6c_{\text{dep}}(t)}{3a(t)l^2 + 2b(t)l^3}. \tag{B5b}$$

Thus, the radius function for a given pore can be updated at each time step. In our implementation, the update is calculated using the forward Euler method.

Physically, $r(y,t)$ is expected to be decreasing pointwise with respect to time. However, in a naive implementation, $r(y_0,t)$ can begin to increase in t at some point y_0. The increasing of r is due to the fact that our linear pores are an approximation of what would in real network pores be a nonlinear pore profile, caused by the nonlinear distribution of foulant particles (as discussed in Sec. II B). This discrepancy between the true shape and the modeled linear shape can conflict with the mass-conservation condition, which sometimes compensates by making a pore increase in size locally. To avoid such unphysical behavior in our simulations, we implement a "freezing algorithm" for pores where it has been detected that $\dot{r}(l,t) > 0$ (here we choose $y_0 = l$, but the choice is arbitrary since any point y_0 will do). We implement the freezing algorithm by setting $\dot{r}(l,t) = 0$:

$$\frac{\partial r}{\partial t}(l,t) = 0 \Rightarrow \dot{a}(t) + \dot{b}(t)l = 0. \tag{B6}$$

Solving for $\dot{b}(t)$ and plugging it into Eq. (B4) gives the following system of ODEs:

$$\dot{b}(t) = -\frac{\dot{a}(t)}{l}, \quad \dot{a}(t) = -\frac{6c_{\text{dep}}(t)}{3a(t)l + b(t)l^2}. \tag{B7}$$

Using Eq. (B6) and $\dot{a}(t) = -c_{\text{in}}(t)$, the condition for activating the freezing algorithm can be expressed as

$$\dot{b}(t)l > c_{\text{in}}(t). \tag{B8}$$

Thus, when $\dot{r} < 0$ everywhere along the length of a given pore, we use the system (B5), and when $\dot{r} > 0$ somewhere along the pore's length, we use the system (B7). In either case, mass is conserved.

In simulations, most pores end up requiring this freezing algorithm at some point during their evolution. As a result, pores rarely close fully within a feasible time frame, and waiting until flux falls to zero through the membrane as the stopping criterion is impractical. Instead, we end our simulations when the flux reaches 1bl of its initial value.

APPENDIX C: PERSISTENCE DIAGRAMS: AN EXAMPLE

Here we give an example of how persistence diagrams are used to extract information about a network, using the toy network in Fig. 15(a). This is a very sparse version of the type of network we consider, with just two inlets, three outlets, four interior pore junctions, and numbers (here taken to be between 0 and 1, but in practice they are the dimensionless pore radii in our specific problem) assigned to each edge. Figure 15(b) shows the corresponding persistence diagram PD_0 for connected components, and Fig. 15(c) shows PD_1 for the loops.

We consider the connected components first as the thresholding parameter θ is decreased continuously from 1 to 0. Three distinct connected components are born at $\theta = 0.9$, $\theta = 0.8$, and $\theta = 0.7$, respectively. Note that in Fig. 15(b), the point at (0.9,0), corresponding to edge (3,4) in Fig. 15(a), lies directly on the birth axis. This is because, by convention, the first connected component to be born never dies. The first death occurs at $\theta = 0.5$, when the connected component that had previously consisted of the edge (6,9) is absorbed into the connected component originally consisting of (3,4); therefore we say it "dies" at $\theta = 0.5$ (in line with the convention, the older component lives on, becoming larger, while the younger one that joined it is said to die). At $\theta = 0.4$, the connected component born at $\theta = 0.7$ dies as well. At this point, all connected components in the toy network have been absorbed into the one that was born first, at $\theta = 0.9$.

Turning now to loops using the same thresholding process, the first one (which can be written as $\{(1,3), (3,8), (8,5), (5,1)\}$, although we will denote it by l_1) is born at $\theta = 0.3$. The next loop, $l_2 = \{(1,3), (3,5), (5,1)\}$, is born at $\theta = 0.1$, when the edge (3,5) appears. Note that both

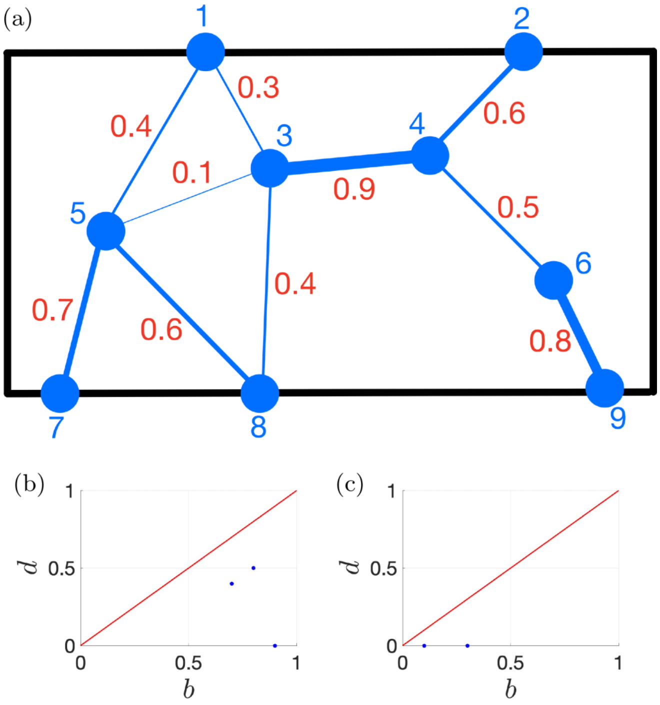

FIG. 15. (a) A simple network, where each edge is equipped with a number. In our case, these numbers are pore radii, but their meaning can be more abstract. Panels (b) and (c) are the persistence diagrams associated with connected components and loops, PD_0 and PD_1, respectively, for the network depicted in panel (a).

points in Fig. 15(c) lie on the birth axis, since, as mentioned before, by convention loops do not die.

Note also that only one loop is born at $\theta = 0.1$, not two. One might naively think that $l_3 = \{(3,8),(8,5),(5,3)\}$ would be another loop born at $\theta = 0.1$, but this is not the case. In fact, we could have chosen l_3 to be our second loop, rather than l_2, but not both. This is because a *cycle* is a path through a network that starts and ends at the same node, whereas a *loop*—while being almost the same concept, at least superficially—is technically a vector in a linearly independent basis that spans all the cycles in a network. So, while all l_i are cycles, only two can be loops, because otherwise the basis would not be linearly independent. The choice of basis is arbitrary, and thus we could have one of three sets of loops, where either l_2 or l_3 would be loops, and the remaining one would be a cycle in the span of the two loops. Note that l_1 must be a loop, since no other loop was born at $\theta = 0.3$.

APPENDIX D: EFFECT OF PORE RADIUS DISTRIBUTION AND THE ROLE OF TORTUOSITY

Here we consider the effect of an important geometric feature of pore networks: the distribution of pore radii. In previous theoretical work (e.g., Refs. [15,24]), pore radii were assumed to be distributed uniformly, but the distribution of pore radii of real network filters is known to be skewed, and the log-normal distribution is generally considered to be a suitable choice to represent such skewed distributions [29–31]. Therefore, the results presented in Sec. III are based on pore networks with radii drawn from a log-normal distribution. Here, we present a brief statistical investigation demonstrating the differences between the uniform and log-normal pore radius distributions—for both noise and network variations, as described in Sec. II B—particularly in terms of their effects on total throughput and tortuosity.

For the uniformly distributed pore radius perturbations, we follow Gu *et al.* [24], who used cylindrical pores with radii

$$r = r_0(1+\eta), \quad \eta \sim \mathrm{Unif}(-\beta, \beta), \tag{D1}$$

where β is the *perturbation strength* (values $0.06 \leqslant \beta \leqslant 0.75$ were used in Ref. [24]) and r_0 is the mean pore radius. To compare this pore radius distribution to the log-normal one that we will use, we introduce the *coefficient of variation* α, defined as the standard deviation of a given distribution divided by its mean:

$$\alpha = \sigma/\mu. \tag{D2}$$

We use the parameter β in Eq. (D1) to find the appropriate value of α for the uniform distribution case:

$$\alpha = \frac{1}{r_0}\sqrt{\mathrm{Var}[r_0(1+\eta)]} = \frac{1}{r_0}r_0\sqrt{\mathrm{Var}[1+\eta]} \tag{D3a}$$

$$= \sqrt{\mathrm{Var}[\eta]} = \frac{\beta-(-\beta)}{2\sqrt{3}} = \frac{\beta}{\sqrt{3}}, \tag{D3b}$$

where the second inequality in Eq. (D3b) follows from the definition of variance for the uniform distribution. We use this coefficient of variation α to define a log-normal pore radius distribution analogous to the uniform one, as we now describe.

A random variable X is log-normal if $\ln X \sim \mathcal{N}(\mu, \sigma^2)$, i.e., $\ln X$ is normal with some mean μ and variance σ^2. We define our log-normal pore radii by

$$r = r_0\gamma, \quad \gamma \sim \mathrm{Lognormal}(\mu, \sigma^2). \tag{D4}$$

Mean pore radius for the uniform distribution case is r_0, hence we prescribe this also for the log-normal pore radii. This requires that the random variables γ_i (where i ranges over the indices of the radii) have a mean of 1, so that

$$\mathbb{E}[r_i] = \mathbb{E}[r_0\gamma_i] = r_0\mathbb{E}[\gamma_i] = r_0. \tag{D5}$$

Following from Eq. (D2), this choice of mean fixes the standard deviation for both the uniform and log-normal distributions as well, to $\sigma = \alpha$. Setting the standard deviation of the log-normal distribution to $\sigma = \alpha = \beta/\sqrt{3}$ gives us the relation we need for comparing uniform and log-normal results.

Having arrived at the relation between the two distributions, we can now examine how the choice of pore size distribution affects throughput, h and tortuosity, τ (see Sec. II B for the definition of τ). For this study we will use $\beta = 0.75$, for two reasons: first, this is one of the values of β used by Gu *et al.* [24]; and second, the corresponding α value is approximately 0.433, which is close to the coefficient of variation ($\alpha = 0.447$) of one of the parameter choices studied in Ref. [31].

We first consider the relationship between h and τ in the case of noise variations. Figure 16(a) depicts this relationship for 1000 independent pore radius perturbations on the typical network, generated as described in Sec. II B, where for each perturbation a simulation was run through to t_{final}, with

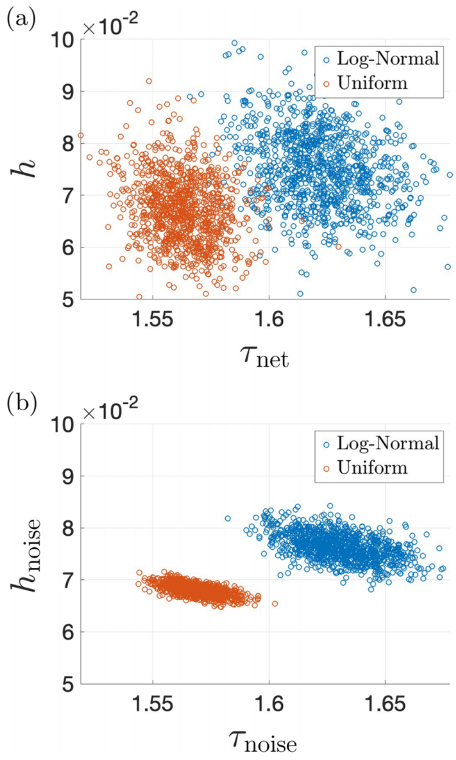

FIG. 16. Throughput h vs tortuosity τ for uniform and log-normal pore radius distributions under (a) noise variations and (b) network variations.

$\alpha = 0.443$. Here we see that h and τ have approximately the same negative, relatively strong correlation for both uniform and log-normal pore radius perturbations, although the correlation is weaker in the log-normal case. We also observe that log-normal pore radius perturbations result in a higher mean and variance for h than in the uniform case. It is clear from the results that log-normal and uniform pore-radius perturbations lead to significantly different results in the case of noise variations.

Figure 16(b) shows h versus τ for uniform and log-normal pore radius distributions under network variations, i.e., the pore radii of each of 1000 unique networks are perturbed twice, once according to the uniform distribution and once according to the log-normal distribution, also with $\alpha = 0.443$. Here the results are similar to those presented in Fig. 16(a) in only one respect: both h and τ have higher mean and variance in the log-normal case than in the uniform case, although the increase in mean and variance is not as high as in the case of noise variations. However, the correlation we see in the case of noise variations is completely absent in the case of network variations. Tortuosity τ is therefore not a good predictor of h under network variations (which are the more physically relevant type of perturbations).

Figures 16(a) and 16(b) show that τ is not a good predictor of total throughput h under either noise or network variations. In the former case h is essentially constant, and it is therefore insensitive to changes in τ; and in the latter case, network variations influence h so strongly that no correlation between h and τ exists. Overall, τ is therefore not a useful predictor of filtration performance within the constraints of our study.

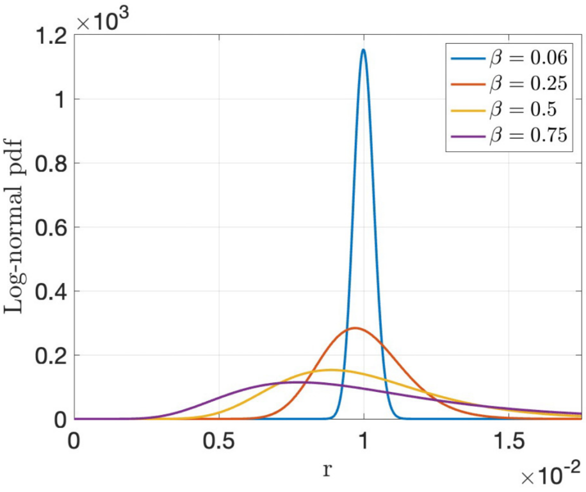

FIG. 17. A comparison of log-normal distributions with respect to radius r, with mean $\mu = 0.01$ and variance $\sigma^2 = (\alpha\mu)^2 = \frac{(\beta\mu)^2}{3}$, for $\beta = \{0.06, 0.25, 0.5, 0.75\}$.

This finding contrasts with the conclusions drawn by Gu *et al.* [24], who found that τ *was* a strong predictor of filtration performance. However, network porosity was not controlled for in that work, and since fixing porosity to a prescribed value is the only substantial change we have made to the network generation procedure used in Ref. [24], we conclude that τ fails to predict filtration performance for networks of the same porosity.

Returning to the higher mean and variance of h seen in the log-normal case (compared to the uniform case), we conjecture that the higher mean is due to the fact that log-normal distributions generally have a large tail for large r (see Fig. 17 for an example log normal distribution with $\beta = 0.75$). This wide tail results in more large pore radii than there are in the uniform case, which in turn results in higher flux, leading to an increase in h and τ, on average. Similarly, we suspect that the wide tail is also responsible for the higher variance seen in h and τ, because the support of the log-normal distribution is much wider than that of the uniform distribution.

There are two main takeaways here: first, τ is not a useful predictor of filtration performance. However, what we lose in the predictive power of τ, we gain in the predictive power of TP_1 and NG_1, (see Sec. III B for a discussion of the predictive power of loop-based topological measures). Second, it is important to use log-normal pore radius distributions (or a similar skewed distribution) in filtration simulations, for two reasons: first, real filter pore radii are believed to be distributed approximately log-normally; and second, there is a measurable difference in both filter performance and network geometry when log-normal pore radii are used instead of uniform pore radii. Therefore, we will only use log-normal pore radii in this and future work.

APPENDIX E: EFFECT OF PORE SIZE PERTURBATION STRENGTH

Gu *et al.* [24] found that the parameter β, the perturbation strength in the uniform pore radius distribution, Eq. (D1), is

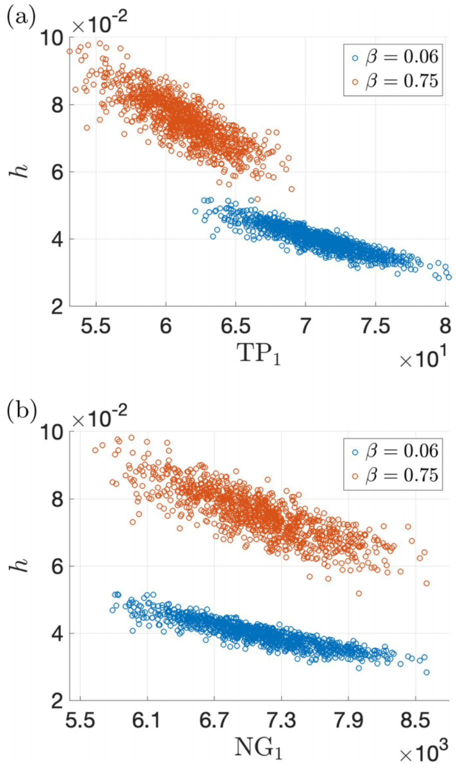

FIG. 18. Throughput h vs (a) total persistence of loops TP_1, and (b) number of loops NG_1, for $\beta = 0.06, 0.75$.

positively correlated with throughput h, and thus it is a good predictor of filtration performance in the case of uniform pore radius perturbation. Here we investigate the effect of β on topological measures in the case of log-normal pore radius distributions.

Perturbation strength has a pronounced effect on log-normal pore radius distributions, since the shape of the distribution changes significantly with changes in the variance, and the variance is controlled by β (see Appendix D for more on the relationship between β and radius variance). This is illustrated in Fig. 17, which depicts a family of log-normal distributions with mean $\mu = 0.01$ and variance $\sigma^2 = (\alpha\mu)^2 = \frac{(\beta\mu)^2}{3}$, for $\beta = \{0.06, 0.25, 0.5, 0.75\}$ (the β values used by Gu *et al.* [24]). When β is small, the log-normal distribution looks like a narrow Gaussian, whereas when β is large, it spreads out, developing a wide tail to the right.

Figures 18(a) and 18(b) depict the same relationships as in Figs. 6(c) and 6(a), respectively (the red circles), but with the addition of results using small β for comparison. Figure 18(a) confirms that there is a negative relationship between h and TP_1, and that the correlation between them is strong for both β values—although the correlation is weaker in the case of large β, since large β means a higher pore radius variance. The increased variance in pore radius causes a corresponding increase in the variance of h, which weakens the correlation between h and TP_1. However, the fundamental relationship between h and TP_1 is essentially the same for either β value, up to a shift along both axes of Fig. 18(a). The vertical shift is due to a wider log-normal distribution for larger β (Fig. 17), allowing for larger pores, which in turn lead to higher throughput h. The horizontal shift is also due to the wider pore radius distribution at higher β, since this effect also leads to more small pores, thus making TP_1 smaller.

Figure 18(b) depicts the correlation between h and NG_1 for small and large β, and it tells a similar story to that of Fig. 18(a): h is a decreasing function of NG_1, and the two are strongly correlated, with larger h in the case of larger β. As noted above, larger β leads to larger variance in h, thus decreasing the strength of the correlation. However, as in Fig. 18(a), the fundamental relationship between h and NG_1 is the same for both β values, up to a vertical shift caused by the effect of larger pore radii. The horizontal shift seen in Fig. 18(a) is not present here, since NG_1 is invariant to pore radius perturbations.

We suspect that the more loops there are, the more paths there are for foulant particles to flow through the network, and thus there is a higher chance that a foulant particle will adhere to the inner wall of a pore. The radius of the pore then shrinks, causing an increase in its resistance, which in turn causes h to decrease.

We can surmise from Fig. 18 that, while β does have an effect on the magnitude and variance of h, it can be disregarded in the study of the effect of topology on filtration performance. Since $\beta = 0.75$ is suspected to give a realistic pore radius variance, this is the value we use in Sec. III.

---

[1] S. Akram, V. Naddeo, Z. A. Rehan, M. Zahid, A. Rashid, and W. Razzaq, A comprehensive review on polymeric nano-composite membranes for water treatment, J. Membr. Sci. **8**, 1 (2018).

[2] G. Liu, M. Xiao, X. Zhang, C. Gal, X. Chen, L. Liu, S. Pan, J. Wu, L. Tang, and D. Clements-Croome, A review of air filtration technologies for sustainable and healthy building ventilation, Sustain. Cities. Soc. **32**, 375 (2017).

[3] C. J. Huang, B. M. Yang, K. S. Chen, C. C. Chang, and C. M. Kao, Application of membrane technology on semiconductor wastewater reclamation: A pilot-scale study, Desalination **278**, 203 (2011).

[4] C. L. Chang and M. S. Chang, Preparation of multi-layer silicone/PVDF composite membranes for pervaporation of ethanol aqueous solutions, J. Membr. Sci. **238**, 117 (2004).

[5] A. Onur, A. Ng, W. Batchelor, and G. Garnier, Multi-layer filters: Adsorption and filtration mechanisms for improved separation, Front. Chem. **6**, 417 (2018).

[6] A. Patanaik, R. D. Anandjiwala, and L. Boguslavsky, Development of high efficiency particulate absorbing filter materials, J. Appl. Polym. Sci. **114**, 275 (2009).

[7] Y. C. Ahn, S. K. Park, G. T. Kim, Y. J. Hwang, C. G. Lee, H. S. Shin, and J. K. Lee, Development of high efficiency nanofilters made of nanofibers, Curr. Appl. Phys. **6**, 1030 (2006).

[8] R. S. Barhate and S. Ramakrishna, Nanofibrous filtering media: Filtration problems and solutions from tiny materials, J. Membr. Sci. **296**, 1 (2007).

[9] Q. Ma, L. Zhao, H. Su, Z. Chen, and Q. Xu, Numerical study of mass transfer and desorption behaviors in deformable porous

media using a coupling lattice Boltzmann model, Phys. Rev. E **102**, 023309 (2020).

[10] X. Du, Y. Shi, V. Jegatheesan, and I. U. Haq, A review on the mechanism, impacts and control methods of membrane fouling in MBR system, Membranes **2020**, 10 (2020).

[11] E. Iritani and N. Katagiri, Developments of blocking filtration model in membrane filtration, KONA Powder Part. J. **33**, 179 (2016).

[12] O. Iliev, R. Kirsch, Z. Lakdawala, S. Rief, and K. Steiner, Modeling and simulation of filtration processes, in *Currents in Industrial Mathematics: From Concepts to Research to Education*, 1st ed. (Springer, Berlin, Heidelberg, 2015), pp. 163–228.

[13] T. Sparks and G. Chase, *Filters and Filtration Handbook*, 6th ed. (Butterworth-Heinemann, Oxford, UK, 2016).

[14] D. Fong, L. J. Cummings, S. J. Chapman, and P. Sanaei, On the performance of multilayered membrane filters, J. Eng. Math. **127**, 23 (2021).

[15] I. M. Griffiths, I. Mitevski, I. Vujkovac, M. R. Illingworth, and P. S. Stewart, The role of tortuosity in filtration efficiency: A general network model for filtration, J. Membr. Sci. **598**, 117664 (2020).

[16] P. Sanaei and L. J. Cummings, Membrane filtration with complex branching pore morphology, Phys. Rev. Fluids **3**, 094305 (2018).

[17] Y. S. Polyakov, Depth filtration approach to the theory of standard blocking: Prediction of membrane permeation rate and selectivity, J. Membr. Sci. **322**, 81 (2008).

[18] M. Sahimi and A. O. Imdakm, Hydrodynamics of particulate motion in porous media, Phys. Rev. Lett. **66**, 1169 (1991).

[19] W. Sun, T. Chen, C. Chen, and J. Li, A study on membrane morphology by digital image processing, J. Membr. Sci. **305**, 93 (2007).

[20] M. J. Blunt, B. Bijeljic, H. Dong, O. Gharbi, S. Iglauer, P. Mostaghimi, A. Paluszny, and C. Pentland, Pore-scale imaging and modelling, Adv. Water Resour. **51**, 197 (2013).

[21] G. Kelly, N. Bizmark, B. Chakraborty, S. S. Datta, and T. G. Fai, Modeling the transition between localized and extended deposition in flow networks through packings of glass beads, Phys. Rev. Lett. **130**, 128204 (2023).

[22] A. U. Krupp, I. M. Griffiths, and C. P. Please, Stochastic of membrane filtration, Proc. R. Soc. A **473**, 20160948 (2017).

[23] B. Gu, L. Kondic, and L. J. Cummings, A graphical representation of membrane filtration, SIAM J. Appl. Math. **82**, 950 (2022).

[24] B. Gu, L. Kondic, and L. J. Cummings, Network-based membrane filters: Influence of network and pore size variability on filtration performance, J. Membr. Sci. **657**, 120668 (2022).

[25] P. Bendich, H. Edelsbrunner, and M. Kerber, Computing robustness and persistence for images, IEEE Trans. Vis. Comput. Graph. **16**, 1251 (2010).

[26] J. A. Perea and J. Harer, Sliding windows and persistence: An application of topological methods to signal analysis, Found. Comput. Math. **15**, 799 (2015).

[27] J. Minhow Chan, G. Carlsson, and R. Rabadan, Topology of viral evolution, Proc. Natl. Acad. Sci. USA **110**, 18566 (2013).

[28] P. Sanaei and L. J. Cummings, Flow and fouling in membrane filters: Effects of membrane morphology, J. Fluid Mech. **818**, 744 (2017).

[29] U. Beuscher (private communication).

[30] U. Beuscher, Modeling sieving filtration using multiple layers of parallel pores, Chem. Eng. Technol. **33**, 1377 (2010).

[31] P. R Johnston, Revisiting the most probable pore-size distribution in filter media: The gamma distribution, Filtr. Separat. **35**, 287 (1998).

[32] GUDHI library—Topological data analysis and geometric inference in higher dimensions, https://gudhi.inria.fr/.

[33] J. R. Munkres, *Elements of Algebraic Topology*, 1st ed. (CRC Press, Menlo Park, CA, 1984).