Quantifying force networks in particulate systems

Miroslav Kramár [a,*], Arnaud Goullet [a], Lou Kondic [b], Konstantin Mischaikow [a]

[a] *Department of Mathematics, Hill Center-Busch Campus, Rutgers University, 110 Frelinghusen Rd, Piscataway, NJ 08854-8019, USA*
[b] *Department of Mathematical Sciences, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA*

HIGHLIGHTS

- We present mathematical models for analyzing force distributions in particulate systems.
- Persistent homology is used to compare the force networks in different granular systems.
- We consider the stability of the persistence diagrams with respect to experimental error.

ARTICLE INFO

ABSTRACT

We present mathematical models based on persistent homology for analyzing force distributions in particulate systems. We define three distinct chain complexes of these distributions: *digital*, *position*, and *interaction*, motivated by different types of data that may be available from experiments and simulations, e.g. digital images, location of the particles, and the forces between the particles, respectively. We describe how algebraic topology, in particular, homology allows one to obtain algebraic representations of the geometry captured by these complexes. For each complex we define an associated force network from which persistent homology is computed. Using numerical data obtained from discrete element simulations of a system of particles undergoing slow compression, we demonstrate how persistent homology can be used to compare the force distributions in different systems, and discuss the differences between the properties of digital, position, and interaction force networks. To conclude, we formulate well-defined measures quantifying differences between force networks corresponding to the different states of a system, and therefore allow to analyze in precise terms dynamical properties of force networks.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Particulate systems consisting of a large number of particles have attracted significant attention in the last decades. Despite significant research on these systems, their properties are still not well understood and some of them appear to be rather elusive. The fact that the forces do not propagate uniformly in systems made of interacting particles has been established in a number of different systems, including granular matter, colloids, gels, emulsions and foams, see, e.g., [1–4]. It is well accepted that the interparticle forces play a key role in determining the mechanical properties of static and dynamic systems; see e.g. [5] for an extensive review of the role of interaction networks in the context of amorphous solids. However there are no universal methods for describing and quantifying relevant aspects of the interparticle forces. For example, even the commonly used notion of 'force chain' – which we take to mean a connected set of particles interacting by a larger than average force – is not generally defined. One important goal of this paper is to present a method that can be used to describe precisely the global properties of force networks in both static and dynamic settings.

Forces between interacting particles have been considered extensively from statistical point of view, in particular in the context of dense granular matter (DGM). For example, the works by Radjai and collaborators, see, e.g. [6,7], discussed the differences in the probability density functions of strong and weak forces (distinguished by the forces being larger or smaller than the average one) arising in simulations; Behringer and collaborators explored these forces in the experimental systems built from photoelastic particles, see e.g. [4]. Possible universality of the force distributions has been considered [8], as well as the connections between force and contact networks [9]. These works have provided a significant insight into the statistical properties of the force distributions but

* Corresponding author. Tel.: +1 9082792251.
E-mail addresses: miroslav@math.rutgers.edu (M. Kramár),
arnaud.goullet@gmail.com (A. Goullet), kondic@njit.edu (L. Kondic),
mischaik@math.rutgers.edu (K. Mischaikow).

by design do not focus on the structural properties of force networks.

Only recently, attempts have been made to move beyond purely statistical description and consider in more detail the properties of these networks. Examples of recent studies include works by Tordesillas and collaborators, see [10–12] and the references therein. These studies include extensive discussion of local properties of networks of forces based on the forces that particles experience and on their connectivity, including appropriately defined force chains and force cycles with a particular emphasis on cycles of length 3 and 4. Furthermore, these studies introduce mesoscopic network properties such as degree, clustering coefficient and centrality which describe particle arrangements. Averaging these properties over the entire network allows to discuss the connection between the changes observed in the macroscopic network properties to the underlying structural rearrangements of the material.

Alternative approaches use network-type of analysis to discuss the properties of force networks [13–15]. These works provide a significant new insight and confirm that the properties of force networks are relevant in the context of propagation of acoustic signals [13], fracture [14], and compression and shear [15]. Topology based approach has been considered as well, with focus on the contact network topology in isotropically compressed [16] and tapped granular media [17]. A similar approach is considered in our recent work [18], where we discuss connectivity of force networks, including the dependence of the number of connected components and holes/loops (quantified by the Betti numbers), on the (normal) force between the particles. While that work uncovered some intricate properties of force networks and allowed to connect the results of topology based analysis to the ones obtained using standard percolation-based approach, it was still based essentially on counting components and loops at fixed magnitudes of force. As such, it thus does not provide an understanding of how these geometric structures persist through different magnitudes of the force.

In [19] we introduced the use of persistent homology [20,21] to DGM. More recently, these ideas have been employed in the context of tapped systems [22]. Conceptually persistent homology is preferable to the above mentioned Betti number analysis. By design persistent homology measures the same geometric structures as the Betti number analysis, but simultaneously records how these structures appear, disappear or persist through different magnitudes of the force. Thus, two networks of forces could produce identical information on the level of Betti numbers, i.e., the number of connected components and loops, but still have distinct global structures in the sense that as one varies the magnitudes of the forces the relationships between the connected components and loops are different. Therefore, the results presented in [19] provide better quantification of the properties of considered force networks and shed new light on the differences between the systems that differ by their frictional properties and particle size distributions.

It should be noted however, that persistent homology is an abstract tool. Hence, there is considerable freedom as to how it can be employed. In this paper we provide a firm mathematical background for using persistent homology in the context of DGM. In addition, we discuss different concepts for constructing and comparing the persistence diagrams. This allows us to compare the features of different force networks both locally and globally and hence is complementary to the approaches, discussed above, that consider local properties of force networks. Furthermore, the ability to compare different force networks is crucial for quantifying the dynamical properties of DGM.

In the next section we give an overview of persistence homology and the structure of the paper.

2. Overview

In this paper we introduce the concept of a force network, which is designed to model force interactions between the particles. The definition varies depending on available form of the data, but every force network is described by a scalar function $f : D \rightarrow \mathbb{R}$. The domain D models the particles and the function f models the forces. Persistent homology is used to reduce the function f to a collection of points in the plane. This collection of points is called a *persistence diagram* and denoted by $PD(f)$. Each point in the persistence diagram encodes a well defined geometric feature of f.

It is useful to view persistent homology as a mapping from scalar functions to persistence diagrams, e.g. $f \mapsto PD(f)$. Stated more formally, persistent homology can be viewed as a function from a space of scalar functions to a space of persistence diagrams. A fundamental result is that with appropriate metrics on the space of functions and on the space of persistence diagrams, persistent homology is a continuous function [21]. At least theoretically this implies that bounded noise or small errors in the measurement of the DGM will lead to a small change in the associated persistence diagram.

This theoretical potential combined with the successful applications presented in [19] suggests the need for a careful analysis of the practical details of applying persistent homology to DGM. There are at least three specific issues that need to be addressed:

1. Given a particular form of the experimental or numerical data, how can one perform the persistent homology computations?
2. Having chosen a method by which the persistent homology computations are being performed, how robust is the resulting persistence diagram as a function of experimental or numerical noise or errors?
3. How can the information provided by the persistence diagrams be used to analyze DGM?

Addressing these issues in the context of DGM is the main focus of this paper.

The first step in the construction of the force network is to establish the domain D on which the function f representing the force interactions is defined. A contact network seems to be a natural candidate for the domain D. Indeed, if positions and shapes of the particles are known, then one can construct a contact network. If the data is in the form of a digital image, then building a contact network is more complicated. In Section 3 we start by introducing *digital* and *position* networks that are closely related to contact networks. We investigate their stability with respect to measurement errors and show that their topology can considerably differ from the topology of the physical system they represent. Therefore we propose an alternative domain, the *interaction* network. This is an abstract mathematical concept and its topology is not related to the topology of the physical system it represents. However, it provides a fixed domain for describing the force networks in DGM.

Section 4 introduces homology, which can be crudely interpreted as a tool for counting connected components, loops and cavities. The advantages of homology are that it supports efficient algorithms, can be used in higher dimensions, and allows one to compare components, loops, and cavities over different spaces. Section 5 introduces force networks, clarifying the connection between the type of available data and formulation of appropriate network. Section 6 focuses on persistence homology, our principal tool for analyzing the force networks. The interaction network can be used in the setting of numerical simulations or experiments (see, e.g., [4]), where complete information about the forces between adjacent particles is known. However, for many experiments only the total force experienced by a particle may be available [23]. This necessitates the use of a digital or position network.

In Section 7 we discuss the space of persistence diagrams, in particular the appropriate metrics on the space, and we provide a theorem that justifies the claim that the interaction force networks are optimal.

In Section 8 we conclude with a review of the developed concepts in the context of DGM data obtained by discrete element based simulations (DES). We choose to work with numerical simulations since all the data is available with high precision and therefore we can process them through all three force network constructions (digital, position, and interaction). This allows for greater clarity in interpreting the geometric meaning of the persistence diagrams, greater ease in comparing the results of the different networks, and simplicity in testing for stability with respect to perturbations. The reader who is familiar with the language of persistent homology may wish to skip directly to this section, before examining the details of the constructions. We note that the focus of this section is not on reaching general conclusions about the force networks in DGM, but on describing how the tools of persistence homology *can* be used to extract detailed information about these networks. For this purpose, we provide a few selected examples of the simulation data and discuss the application of persistence homology to these examples. More interpretation-oriented discussion that focuses on the connection between persistence homology and physical properties of DGM is given in [19].

There are several points that we encourage the reader to keep in mind while reviewing Section 8. First, we provide examples of two dimensional force networks in order to facilitate the reader's intuition about the described features, however our analysis is based exclusively on the information contained in the persistence diagrams. This is particularly important in the context of three dimensional systems, where visual inspection may be impossible. Second, in our examples we mostly concentrate on the magnitude of the normal force between the particles, but in principle any function that assigns a scalar value to every edge can be used. We illustrate this point by briefly considering tangential forces. Finally, although this paper is focused on DGM, the constructions are independent of the details of particle–particle interaction and could as well be applied to any other system consisting of interacting particles. The software used to build the various force networks [24] and to compute persistence diagrams [25] is available in the public domain.

3. Particle networks

In the Introduction the force network is described by a scalar function $f : D \rightarrow \mathbb{R}$. The first step towards using algebraic topology to characterize the geometric structures associated with DGM is to represent the domain D as a finite complex, defined below. We introduce three complexes motivated by the type of data commonly obtained from experiments or simulations. Consider Fig. 1(a) that shows a small portion of an image of a collection of photoelastic disks. We interpret this information in three ways:

Digital. This figure arises from a digital image and thus the data can be viewed as a collection of a large number of pixels.

Positions. Since this is a controlled experiment involving circular disks, the configuration of the particles (locations of their center points and their radii) can be determined.

Interactions. The particles are made of photoelastic material and thus the light intensities within the particles can be used to determine the normal forces between the particles.

We note that the information required increases considerably as digital, position, interaction complexes are considered, respectively. Our approach is to encode each of these data types into different force networks and one of the goals of this paper is to make clear the difference of the geometric information that can be extracted. One not particularly surprising conclusion is that the in-
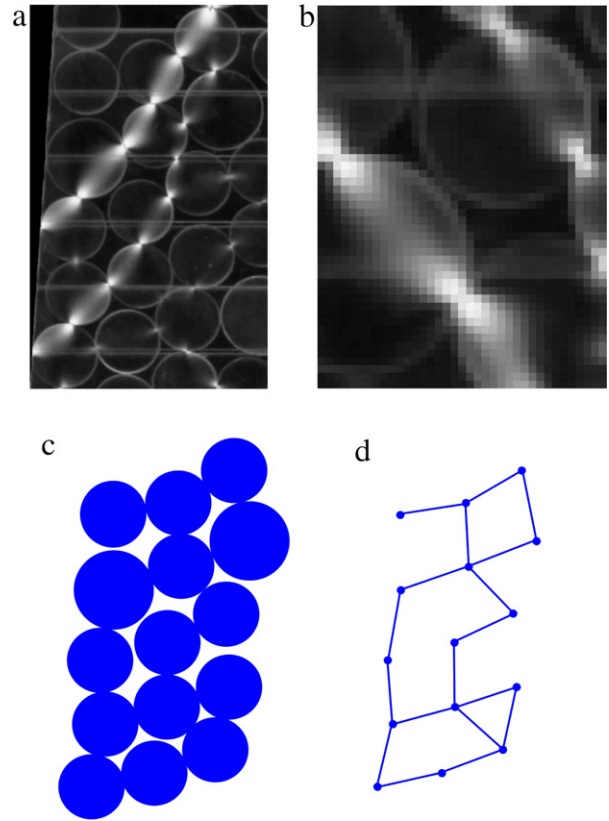

Fig. 1. Different representations of the particle networks derived from the experimental data. For simplicity we neglect the particles that intersect the edges of the picture. (a) Small portion of a digital image of an experimental system (courtesy of R.P. Behringer, unpublished results). (b) Detail of a digital image. (c) Digital complex. The blue pixels represent two dimensional cubes present in the complex. (d) Position complex. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

teraction data provides the best information and the digital data the worst, but it is worth quantifying these differences. With this in mind we begin with several formal definitions. Our focus is on physical systems, thus for the most part we restrict our discussion to two and three dimensional complexes (see [21,26] for a more general discussion).

To interpret the pixel data we make use of cubical complexes. Observe that once the pixel data is obtained the actual size used to represent each pixel is no longer an issue. Thus, for the sake of simplicity of discussion and without loss of generality we assume that the pixel data is embedded in $\mathbb{R}^2$ with each pixel being represented by a square defined by the integer lattice. More precisely, a 2-*dimensional cube* (*pixel*) is a square of the form $[n, n+1] \times [k, k+1]$, a 1-*dimensional cube* (*edge*) is a unit interval of the form $[n, n] \times [k, k+1]$ or $[n, n+1] \times [k, k]$, where $n, k \in \mathbb{Z}$, and a 0-*dimensional cube* (*vertex*) is a point with integer coordinates. A two dimensional *cubical complex* CN is a collection of 0,1 and 2-dimensional cubes that satisfy the following property: if $\sigma \in$ CN and $\sigma' \subset \sigma$, then $\sigma' \in$ CN. This property guarantees that for every 2-dimensional cube (edge) in CN its edges (vertices) belong to CN as well.

Definition 3.1. Given a digital image of particles $\{p_i \mid i = 0, \ldots, I\}$ the *digital complex* CN_D is the cubical complex consisting of squares $\{\sigma_j\}$ where each square σ_j represents a single pixel associated with some particle.

How pixels are associated with particles is intentionally left vague in Definition 3.1. The actual association depends on the particular characteristics of the imaging device, filtering, thresholding,

etc., used to obtain and process the data. Conceptually, the most straightforward approach is to discretize the domain of the image into squares, identify the squares with pixels, and declare the pixel to represent a particle if the associated square intersects the particle. A cubical particle network CN_D is shown in Fig. 1(c). We provide more detail about the construction later.

Remark 3.2. In this paper we consider only two-dimensional examples. However, the same ideas can be applied to particles in $\mathbb{R}^3$ where the three dimensional images are represented as voxels. In this case one builds a cubical complex by representing each voxel as a unit cube of the integer lattice in $\mathbb{R}^3$ (see [26] for the general theory).

In the case of particles with simple shape the contact network of [15] or the unweighted (binary) network of [13] can be used to represent the particles. We represent these networks in terms of simplicial complexes which are defined as follows. We begin with a finite set of vertices $CN^{(0)} := \{v_i \mid i = 0, \ldots, I\}$. An *n-dimensional simplex* in CN is a subset of $CN^{(0)}$ consisting of $n + 1$ vertices. The set of *n*-dimensional simplices in CN is denoted by $CN^{(n)}$. Given the set of vertices $CN^{(0)} := \{v_i \mid i = 0, \ldots, I\}$ it is customary to denote the 0-dimensional simplices by $\langle v_i \rangle$, the 1-dimensional simplices by $\langle v_i, v_j \rangle$, and the 2-dimensional simplices by $\langle v_i, v_j, v_k \rangle$. One and two dimensional simplices are referred to as *edges* and *triangles*. A *simplicial complex* CN is a collection of simplices that satisfies the following property: if $\sigma \in CN$ and $\sigma' \subset \sigma$, then $\sigma' \in CN$.

Definition 3.3. Given a collection of circular disks $\{p_i \mid i = 0, \ldots, I\}$, location of their centers $\{x_i \in \mathbb{R}^2 \mid i = 0, \ldots, I\}$, and their radii $\{r_i \mid i = 0, \ldots, I\}$ the associated *position complex* CN_P is the simplicial complex consisting of vertices $\{v_i \mid i = 0, \ldots, I\}$, where each vertex v_i is identified with particle p_i, and edges $\langle v_i, v_j \rangle$ if and only if $\|x_i - x_j\| \leq r_i + r_j$.

For the sake of clarity Definition 3.3 is presented in the context of the examples considered in this paper. More generally, one can consider spherical particles positioned in $\mathbb{R}^d$, $d = 3$ being the most relevant for physical applications. As presented, the position complex is an abstract simplicial complex; that is, there is no specific geometric object associated with it. In the context of this work we can always geometrize the complex by declaring the vertices to be the points $\{x_i \in \mathbb{R}^2 \mid i = 0, \ldots, I\}$, and the edges to be the line segments connecting the points. From now on we rarely distinguish between the abstract simplicial complex and its geometric realization.

Having defined these complexes, a reasonable first question is whether they correctly capture the topology of the particle configuration. We begin with the following positive result under the assumption that the particles cannot deform under the pressure induced by contacts with other particles.

Proposition 3.4. *Given a collection of circular hard disks $\{p_i \mid i = 0, \ldots, I\}$, location of their centers $\{x_i \in \mathbb{R}^2 \mid i = 0, \ldots, I\}$, and their radii $\{r_i \mid i = 0, \ldots, I\}$ the associated position complex CN_P is homotopic to the union of the regions occupied by the particles, $\bigcup_{i=0}^{I} p_i$.*

The proof follows from retracting the set of particles onto the geometric realization of CN_P (see for example Fig. 2). We do not provide details of the proof because this result is of limited importance. In any experiment or numerical simulation the locations of the particles can only be given up to some specified precision. If we assume the particles to be hard, then two particles p_i and p_j are in contact if and only if $\|x_i - x_j\| = r_i + r_j$. Clearly, arbitrarily small errors in x_i and x_j can lead to an inequality which indicates that the particles are not in contact. The same argument applies to arbitrarily small errors in the measurements of the radii of the disks. Assuming that the particles are soft makes this result slightly more
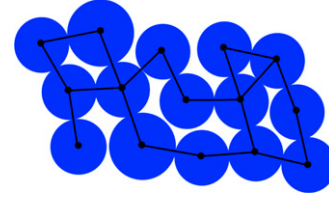

Fig. 2. With complete information the position complex CN_P has the same homotopy type as the configuration space of the particles $\bigcup_{i=0}^{I} p_i$. The proof involves collapsing the particles onto graph.

robust, but this is tempered by the fact that this stability depends on the existence of sufficiently large normal forces. We attempt to quantify these comments in Section 8.2.

To measure the topological fidelity of the digital complex requires the choice of a rule for determining if a pixel is included in the complex or not. For the sake of clarity we continue with the conceptually simple rule that the pixel belongs to the complex if the associated square intersects some particle.

Fig. 3 shows the digital complexes associated with different particle configurations. One can see that the failure of the digital complex CN_D to correctly capture the topology of the particle configuration can be quite dramatic. Even the simplest setting of two particles with a high pixel resolutions does not guarantee a correct topological description. Fig. 3(a) and (c) demonstrate that both the number of connected components and loops can be counted incorrectly. Fig. 3(a) and (b) show that there is no particular direction to the error in the loop count. The fact that the number of components of the digital complex is never larger than that of the configuration of particles arises from our assumption on how to identify pixels with particles. In particular our approach leads to an artificial expansion of the area covered by each particle. Thus, two separate particles can appear to be in contact, but two particles that are in contact can never appear to be separated.

To keep things in perspective we remind the reader that even though it is clear that the digital complex can fail to record the correct topology in a variety of ways it is the easiest means of collecting data and is applicable in situations where the forces between the grains cannot be directly measured and without a priori assumptions about the geometry and rigidity of the grains.

Consider the position network CN_P for the particle conformation shown in Fig. 4(a), which we refer to as a *crystalline structure* since the particles are packed as densely as possible. If we restrict our definition of the position complexes to graphs (one dimensional simplicial complexes), then CN_P has nine loops all of which involve three particles. This should be contrasted with Fig. 4(b) in which there are only 4 loops, but three of them are associated with particles that are not packed as densely as possible. Since in a perfect densely packed crystalline structure made up of disks of the same size all loops would be made up of exactly 3 particles we refer to a loop involving four or more particles as a *defect*. We use this definition of a defect even for systems built from variable size particles.

In this paper we have chosen to focus on defects, i.e. we want to avoid counting loops that can be expressed in terms of three particles. This can be done naturally using the flag complex, $CN_P^{\blacktriangle}$, defined as follows. Set

$$CN_P^{\blacktriangle(0)} := CN_P^{(0)} \quad \text{and} \quad CN_P^{\blacktriangle(1)} := CN_P^{(1)}$$

and

$$\langle v_i, v_j, v_k \rangle \in CN_P^{\blacktriangle(2)}$$

if and only if

$$\left\{ \langle v_i, v_j \rangle, \langle v_i, v_k \rangle, \langle v_j, v_k \rangle \right\} \subset CN_P^{(1)}.$$

The newly added triangles $\langle v_i, v_j, v_k \rangle$ exactly fill in the loops formed by three particles, see Fig. 4(c).
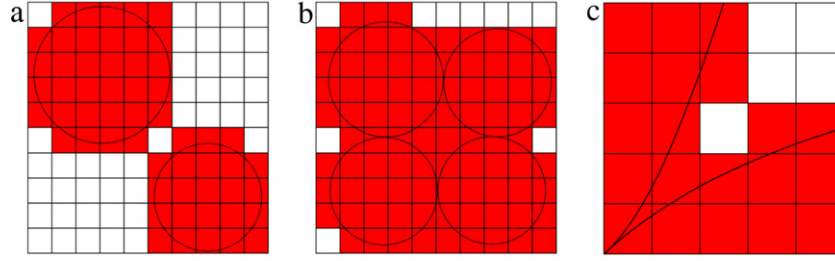
Fig. 3. Failure of the digital complex to correctly capture the topology of the particle configuration. (a) The particle configuration consists of two components and contains no loops. The associate digital complex has one component and one loop. (b) The particle configuration contains one loop while the digital complex contains none. (c) The particle configuration contains one component and no loops. The associate digital complex has one component and one loop. Furthermore, doubling the resolution does not remove the unwanted loop.
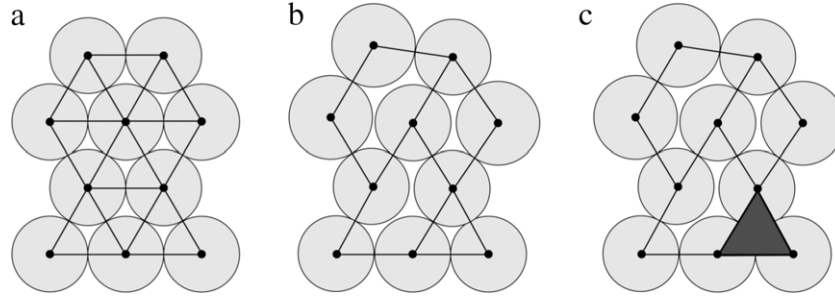

Fig. 4. Two configuration spaces with position complexes CN_P. (a) Crystalline structure with many loops formed by three particles. (b) Noncrystalline structure with fewer loops than (a). (c) Flag complex $CN_P^▲$ derived from (b). Observe that the only loops which remain are associated with defects.

It is worth noting that attempting an analogous construction of a flag complex in the setting of cubical complexes that arise from digital complexes will not work. A square can be missing from CN_D because of the phenomena indicated in Fig. 3(a) or (c), but it can also be missing because it represents the loop formed by four distinct particles.

As is discussed at the beginning of this section, the complexes CN_D and CN_P are introduced in order to be able to apply algebraic topological tools to the characterize the geometric structures of DGM. Unfortunately, these complexes are not necessarily robust with respect to perturbations since arbitrarily small changes in the locations of the particles can lead to the loss of cubes or edges in CN_D or CN_P, respectively. With this in mind we introduce the following complex.

Definition 3.5. Given a collection of particles $\{p_i \mid i = 0, \dots, l\}$ the *interaction complex* CN_I is the simplicial complex consisting of vertices $\{v_i \mid i = 0, \dots, l\}$ where each vertex v_i is identified with particle p_i, all edges $\langle v_i, v_j \rangle$, and all triangles $\langle v_i, v_j, v_k \rangle$.

The interaction complex itself does not have any meaningful geometric interpretation. However, if the forces between the grains are available, then we can make use of CN_I as the domain for the force complex (see Section 6). This in turn allows us to prove continuity for the persistence diagrams (Corollary 7.3). The implications of continuity, or lack thereof in the case of CN_D and CN_P, is made clear in Section 8.2.

4. Homology

We pause in our development of the networks to review a few fundamental definitions from the classical theory of homology with a focus on the simple setting of the digital, position and interaction complexes introduced in Section 3. For a more general discussion the reader is referred to a standard text in algebraic topology or to [21,26] for descriptions more closely associated with the data analysis.

Recall that position complexes CN_P and interaction complexes CN_I are simplicial complexes. This leads to our use of simplicial homology. Recall that $CN^{(n)}$ denotes the set of n-dimensional simplices in the simplicial complex CN. The *n-chains* of CN are defined to be the vector space

$$C_n(CN) := \left\{ \sum_{\sigma \in CN^{(n)}} m_\sigma \sigma \mid m_\sigma \in \mathbb{Z}_2 \right\}. \tag{1}$$

Since we are working with planar arrangements of particles it is sufficient to use $\mathbb{Z}_2$ coefficients, i.e. the set $\{0, 1\}$ with the standard binary addition and multiplication operations. Observe that $C_n(CN)$ is the vector space over $\mathbb{Z}_2$ with basis elements consisting of the n-dimensional simplices.

The associated *boundary maps* are linear maps (these are often represented as matrices using the simplices as bases) $\partial_n \colon C_n(CN) \to C_{n-1}(CN)$ ($C_{-1}(CN) := 0$) defined on the simplices as follows

$$\partial_0 \langle v_i \rangle := 0$$
$$\partial_1 \langle v_i, v_j \rangle := \langle v_i \rangle + \langle v_j \rangle$$
$$\partial_2 \langle v_i, v_j, v_k \rangle := \langle v_i, v_j \rangle + \langle v_i, v_k \rangle + \langle v_j, v_k \rangle.$$

A direct calculation making use of the linearity and the use of $\mathbb{Z}_2$ coefficients show that $\partial_{n-1} \circ \partial_n = 0$, e.g.

$$\begin{aligned}
\partial_1 \circ \partial_2 \langle v_i, v_j, v_k \rangle &= \partial_1(\langle v_i, v_j \rangle + \langle v_i, v_k \rangle + \langle v_j, v_k \rangle) \\
&= \partial_1 \langle v_i, v_j \rangle + \partial_1 \langle v_i, v_k \rangle + \partial_1 \langle v_j, v_k \rangle \\
&= \langle v_i \rangle + \langle v_j \rangle + \langle v_i \rangle + \langle v_k \rangle + \langle v_j \rangle + \langle v_k \rangle \\
&= 0.
\end{aligned}$$

The boundary maps can be used to identify components and loops. To do this we focus on *cycles*, these are chains which are sent to the 0 vector under ∂_n. More formally,

$$Z_n(CN) := \ker \partial_n.$$

Observe that $\langle v_i \rangle \in Z_0(CN)$ and $\langle v_i, v_j \rangle + \langle v_i, v_k \rangle + \langle v_j, v_k \rangle \in Z_1(CN)$.

The power of homology is that we are able to move from geometric data to an algebraic format from which we can then extract geometric information. For example, the algebraic statement that
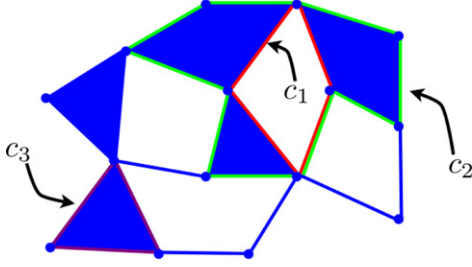
Fig. 5. Three different 1-dimensional chains. Each of these chains corresponds to a loop and hence is a cycle. The red chain c_1 and the green chain c_2 correspond to the same loop in the particle network. The brown loop c_3 does not correspond to any loop and can be contracted to a point. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$\langle v_i \rangle \in Z_0(\text{CN})$ can be interpreted as a statement that $\langle v_i \rangle$ identifies a component of CN. Similarly, $\langle v_i, v_j \rangle + \langle v_i, v_k \rangle + \langle v_j, v_k \rangle \in Z_1(\text{CN})$ can be identified with the path of edges $\langle v_i, v_j \rangle, \langle v_i, v_k \rangle, \langle v_j, v_k \rangle$ that makes up a loop. To emphasize the relationship between the algebra and geometry consider the simplicial complex indicated in Fig. 5. There are three chains that form loops and hence cycles indicated in red, green and brown.

For obvious reasons it is important not to over count components or loops. In particular, if an edge $\langle v_i, v_j \rangle$ belongs to CN, then $\langle v_i \rangle$ and $\langle v_j \rangle$ belong to the same component and therefore we wish to identify them. This can be done algebraically by the relation $\partial_1 \langle v_i, v_j \rangle = \langle v_i \rangle + \langle v_j \rangle$. Similarly, if a 2-dimensional simplex $\langle v_i, v_j, v_k \rangle \in \text{CN}$, then the loop $\langle v_i, v_j \rangle, \langle v_i, v_k \rangle, \langle v_j, v_k \rangle$ does not enclose a loop and thus should not be counted. Again, this can be detected algebraically by the relation $\partial_2 \langle v_i, v_j, v_k \rangle = \langle v_i, v_j \rangle + \langle v_i, v_k \rangle + \langle v_j, v_k \rangle$. Observe that the relations in these examples are obtained via images of the boundary operator. This leads to the definition of the *boundaries* of CN,

$$B_n(\text{CN}) := \partial_{n+1}\left(C_{n+1}(\text{CN})\right).$$

Referring to the complex depicted in Fig. 5 observe that there exists $\bar{c}$ such that $\partial_2 \bar{c} = c_1 + c_2$ which implies that the cycles c_1 and c_2 represent the same loop in the complex. This motivates the following definition. The n-th *homology group* of the simplicial complex CN is defined by

$$H_n(\text{CN}) := \frac{Z_n(\text{CN})}{B_n(\text{CN})},$$

the vector space of equivalence classes of cycles identified by boundaries. To be more specific given a cycle $z \in Z_n(\text{CN})$ the associated homology class $[z] = [z]_{\text{CN}}$ is the equivalence class of all cycles of the form $z + b$ where $b \in B_n(\text{CN})$.

The dimension of $H_n(\text{CN})$ is called the n-th *Betti number* $\beta_n(\text{CN})$. $\beta_0(\text{CN})$ counts the number of components and $\beta_1(\text{CN})$ counts the number of loops which encircle a void. If we were working with DGM in three dimensions then $\beta_2(\text{CN})$ would indicate the number of cavities.

A fundamental property of homology is that if two topological spaces are homotopic, then they have the same homology groups. A corollary of this is that under the hypothesis of Proposition 3.4 the Betti numbers of CN_P agree with the Betti numbers of the space defined by $\cup_{i=0}^{l} p_i$. Since the hypotheses of this proposition are rather strong, e.g. exact knowledge of locations and radii, in general, given numerical or experimental data we do not expect that these Betti numbers agree.

Recall that the interaction complex does not have a meaningful geometric interpretation. In fact, independent of the number and arrangement of the particles, the homology of CN_I is very simple,

$$\beta_n(\text{CN}_I) \cong \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n > 0. \end{cases} \quad (2)$$

Thus the Betti numbers tell us that CN_I has a single connected component and does not have any loops that encircle a void.

We do not present the details of computing homology with cubical complexes. Conceptually the ideas are the same, though the boundary operators are slightly different. The interested reader is referred to [26] for a complete presentation. Even more generally, simplicial and cubical complexes are examples of chain complexes and the individual simplices or cubes are examples of *cells*.

The reader may be somewhat underwhelmed by the fact that we have constructed a significant amount of algebra to essentially count components and loops, especially since there are extremely efficient graph theoretic algorithms for performing these operations. However, the algebra allows us to relate components and loops in different complexes. Recall that components and loops are measured by elements of $H_0(\text{CN})$ and $H_1(\text{CN})$. Thus, given two distinct chain complexes CN and CN′ we need to be able to relate homology classes of $H_k(\text{CN})$ with homology classes of $H_k(\text{CN}')$. This is done via the following algebraic construction. Linear maps $\phi_n : C_n(\text{CN}) \to C_n(\text{CN}')$ are *chain maps* if

$$\partial'_n \phi_n = \phi_{n-1} \partial_n$$

for all n where ∂_n and ∂'_n are the boundary maps for $C_n(\text{CN})$ and $C_n(\text{CN}')$. A fundamental result is that if ϕ_n is a chain map, then ϕ_n induces a linear *map on homology* $\phi_n : H_n(\text{CN}) \to H_n(\text{CN}')$ defined by

$$\phi_n([z]_{\text{CN}}) := [\phi_n(z)]_{\text{CN}'}.$$

As is made clear in Section 6 for the purposes of this paper it is sufficient to note that if $\text{CN} \subset \text{CN}'$, then the inclusion map induces, for each dimension, a chain map and hence a map on homology.

5. Force networks

As is indicated in the Introduction it is well accepted that the geometry of force chains plays an important role in determining the macroscopic properties of dense granular material. In this section we expand on the complexes constructed in Section 3 to include the forces between the particles into this mathematical framework. In the present work we mainly focus on the normal force, that is, the component of the force projected on the line connecting the centers of interacting particles. We will view the magnitude of the normal force as a scalar field defined over the complex, i.e., a function $f : \text{CN} \to \mathbb{R}$. There is one constraint on the definition of f that arises from the use of persistent homology to capture the geometry of the force chains.

To understand this constraint assume for the moment that we are given a complex CN and a scalar field $f : \text{CN} \to \mathbb{R}$. Since for particulate systems one often considers particles interacting by strong/weak forces, we are interested in the geometry of a part of the complex on which the forces exceed a specified level. Thus we define a *force network* to be the super level set

$$\text{FN}(f, \theta) := \{\sigma \in \text{CN} \mid f(\sigma) \geq \theta\} \quad (3)$$

which corresponds to the part of the particle network experiencing force larger than θ. We use homology to quantify the geometry of $\text{FN}(f, \theta)$. Hence $\text{FN}(f, \theta)$ has to be a complex for every value of θ. This means that if $\sigma \in \text{FN}(f, \theta)$ and $\sigma' \subset \sigma$, then $\sigma' \in \text{FN}(f, \theta)$. Thus, in our construction of f we need to insure that this condition is satisfied. This leads to the following definition.

Definition 5.1. Given a complex CN, a function $f : \text{CN} \to \mathbb{R}$ is *algebraically monotone* if $f(\sigma') \geq f(\sigma)$ for every $\sigma', \sigma \in \text{CN}$ such that $\sigma' \subset \sigma$.

It is left to the reader to check that if f is algebraically monotone, then $\text{FN}(f, \theta)$ is a complex for every value of θ.

Definition 5.2. Given a complex CN and an algebraically monotone function $f : \text{CN} \to \mathbb{R}$, the associated *force network filtration*
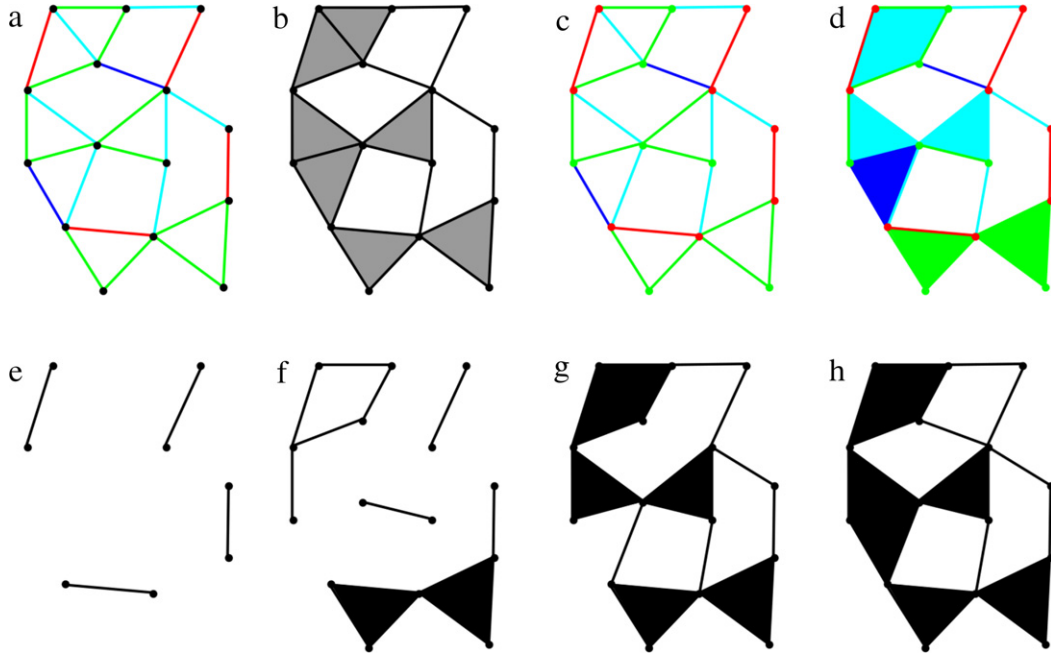
Fig. 6. A representation of a simple interaction network FN_I. (a) Vertices represent the particles and the edges correspond to the non zero force between the particles. An increasing value of the force is denoted by blue, cyan, green, and red. (b) Collection of simplices on which the function $\mathsf{CN}_I : f \to \mathbb{R}$ is positive. Extension of the function f to the vertices (c) and to the 2-dimensional simplices (d). (e)–(h) The complexes $\mathsf{FN}_{\bullet}(f, \theta_i)$ for positive θ equal to θ_4 (red), θ_3 (green), θ_2 (cyan) and θ_1 (blue), respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

is the collection of all force network complexes

$$\{\mathsf{FN}(f, \theta) \mid \theta \in \mathbb{R}\}.$$

The construction of $f : \mathsf{CN} \to \mathbb{R}$ depends on the available information. The weaker assumption, which we associate with digital or position complexes, is that for each particle p_i we can estimate the magnitude of the force ψ_i on p_i. The stronger assumption, which leads to the use of an interaction complex, is that we can estimate the magnitude of the force $\psi_{i,j}$ between any two particles p_i and p_j. The function f is defined in two steps. First we define f for cells of a certain dimension depending on the type of complex. Then we uniquely extend the definition to all the cells. The construction of the extension guarantees that f is an algebraically monotone function.

Digital force networks. Recall that a 2-dimensional cube $\sigma \in \mathsf{CN}_D^{(2)}$ if it intersects at least one particle p_i. We define

$$f(\sigma) = \max \{\psi_i \mid \sigma \cap p_i \neq \emptyset\}$$

for $\sigma \in \mathsf{CN}_D^{(2)}$.

Position force networks. For each $\langle v_i \rangle \in \mathsf{CN}_P^{(0)}$ corresponding to the particle p_i we define

$$f(\langle v_i \rangle) := \psi_i.$$

Interaction force networks. For the interaction network CN_I the natural starting point for the definition of f is on the edges $\langle v_i, v_j \rangle \in \mathsf{CN}_I^{(1)}$,

$$f(\langle v_i, v_j \rangle) := \psi_{i,j}.$$

For a complex $\mathsf{CN}_{\bullet}$, $\bullet \in \{P, I\}$ we extend the definition of the function f from $\mathsf{CN}_{\bullet}^{(i)}$ to the cells $\sigma \in \mathsf{CN}_{\bullet}^{(j)}$ for $j < i$ by

$$f(\sigma) = \max \{f(\sigma') \mid \sigma \subset \sigma', \ \sigma' \in \mathsf{CN}_{\bullet}^{(i)}\}.$$

Extension to cells $\sigma \in \mathsf{CN}_{\bullet}^{(j)}$ for $j > i$ and $\bullet \in \{P, I\}$ is defined by

$$f(\sigma) = \min \{f(\sigma') \mid \sigma' \subset \sigma, \ \sigma' \in \mathsf{CN}_{\bullet}^{(i)}\}.$$

We use the following proposition to summarize the above discussion and constructions.

Proposition 5.3. *Given a complex* $\mathsf{CN}_{\bullet}$, $\bullet \in \{D, P, I\}$ *and* f *defined as above, the associated super level set* $\mathsf{FN}_{\bullet}(f, \theta)$ *is a complex for all values of* $\theta \in \mathbb{R}$.

Since we are assuming that there is only a finite number of particles in our system, any force network filtration $\{\mathsf{FN}_{\bullet}(f, \theta) \mid \theta \in \mathbb{R}\}$ contains only finitely many distinct complexes. We can use homology, in particular the Betti numbers, to characterize the geometry of each of the distinct force networks in the force network filtration.

To gain intuition into the force networks consider the interaction force network indicated in Fig. 6. Fig. 6(a) represents a collection of particles. The particles are represented by the vertices and the shown edges correspond to the non-zero forces between the particles. Fig. 6(b) shows only the simplices of CN_I for which the value of the function $f : \mathsf{CN}_I \to \mathbb{R}$ is positive. The value of f on the edges is determined by the forces between the particles. In Fig. 6(a) the non-zero forces are color coded. In order of increasing value, the force is denoted by θ_1 (blue), θ_2 (cyan), θ_3 (green) and θ_4 (red). The value of the function f is extended to the vertices in Fig. 6(c) and to the 2-dimensional simplices in Fig. 6(d). Fig. 6(e)–(h) indicate the associated force network for non negative values of θ. For $\theta \leq 0$ the complex $\mathsf{FN}_I(f, \theta) = \mathsf{CN}_I$, and as explained in Section 4, it consists of a single connected component that does not encircle any loops.

For example, referring to the force network filtration of Fig. 6 we can extract the following data:

$$(\beta_0(\mathsf{FN}_I(f, \theta_4)), \beta_1(\mathsf{FN}_I(f, \theta_4))) = (4, 0)$$
$$(\beta_0(\mathsf{FN}_I(f, \theta_3)), \beta_1(\mathsf{FN}_I(f, \theta_3))) = (4, 1)$$
$$(\beta_0(\mathsf{FN}_I(f, \theta_2)), \beta_1(\mathsf{FN}_I(f, \theta_2))) = (1, 3)$$
$$(\beta_0(\mathsf{FN}_I(f, \theta_1)), \beta_1(\mathsf{FN}_I(f, \theta_1))) = (1, 4).$$

It is worth noting that the H_0 homology information for $\mathsf{FN}_I(f, \theta_4)$ and $\mathsf{FN}_I(f, \theta_3)$ agree and yet the structure of the components has changed dramatically. Two distinct connected components become one and a new connected component is formed. To capture this information we make use of the fact that these complexes
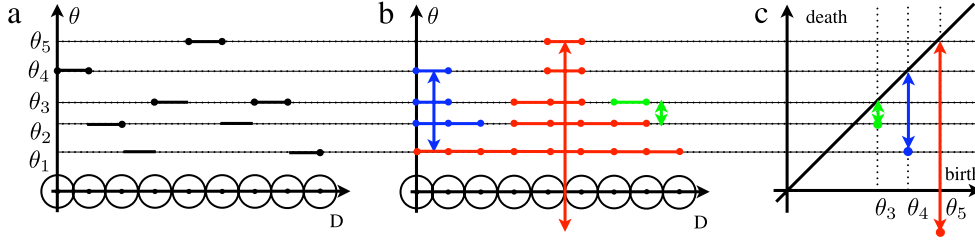
Fig. 7. (a) A linear chain of particles is shown on the horizontal axis. Magnitude of the normal force between the adjacent particles is given by the value of the step function above the edge connecting the particles' centers. (b) Structure of $FN_I(f, \theta)$ for different values of θ. For example, $FN_I(f, \theta_5)$ consist of a single connected component formed by one edge and two vertices. The set $FN_I(f, \theta_3)$ contains three distinct connected components corresponding to the different geometric features indicated by the double arrows. (c) β_0 persistence diagram for the interaction force network of the system shown in (a). The color of the points matches the color of the features they represent. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

are nested by inclusion. This leads to the concept of persistent homology.

6. Persistent homology

Given a force network filtration

$$FN(f, \theta) := \{\sigma \in CN \mid f(\sigma) \geq \theta\}$$

generated by a finite number of particles, there is a finite number of values

$$0 = \theta_0 < \theta_1 < \cdots < \theta_K = \max_{\sigma \in CN} f(\sigma)$$

such that $\theta_k = f(\sigma)$ for some $\sigma \in CN$. Though the Betti numbers characterize the topology of a given force network $FN(f, \theta_k)$, the vector space structure of homology plays an essential role in that it allows us to compare the topology of $FN(f, \theta_k)$ with any other force network $FN(f, \theta_j)$. Given $\theta_i < \theta_j$, $FN(f, \theta_j) \subset FN(f, \theta_i)$ and hence there is an inclusion map

$$\iota_{\theta_i, \theta_j} : FN(f, \theta_j) \to FN(f, \theta_i).$$

As is indicated at the end of Section 4, this defines maps

$$\iota_{\theta_i, \theta_j *} : H_*(FN(f, \theta_j)) \to H_*(FN(f, \theta_i))$$

on each homology group H_*. It is important to note that $\iota_{\theta_1, \theta_2 *}$ need not be an inclusion map on the level of the homology groups.

Persistent homology makes use of the above mentioned maps to compare topological features within different force networks. The first observation, while trivial, is essential for our discussion and follows directly from the fact that $FN(f, \theta) = \emptyset$ for all $\theta > \theta_K$.

Lemma 6.1. *If $\theta > \theta_K$, then $H_*(FN(f, \theta)) = 0$.*

Now consider a value θ_k such that $v \in H_n(FN(f, \theta_k))$ and let $v \neq 0$. If $n = 0$ or 1, then v provides information about the existence of components or loops, respectively, in $FN(f, \theta_k)$. In light of Lemma 6.1, there exists a unique largest threshold $\theta_b(v) \geq \theta_k$ with the property that there exists $v_b \in H_n(FN(f, \theta_b))$ such that $\iota_{\theta_k, \theta_b *}(v_b) = v$. The geometric feature associated with v is said to have been *born* at level $\theta_b(v)$.

It is also possible that for some $\theta < \theta_k$, $\iota_{\theta, \theta_k *}(v) = 0$. In this case we define

$$\theta_d(v) := \max \left\{ \theta_j \mid \iota_{\theta_j, \theta_k *}(v) = 0 \right\}$$

and we say that the geometric feature associated with v *dies* at level $\theta_d(v)$. Given our construction, not every geometric feature needs to die. In particular, for $n = 0, 1$,

$$H_n(FN(f, 0)) \cong H_n(CN)$$

which, as the examples in this paper indicate, need not be trivial for digital and position complexes. We make use of the following convention

if $\iota_{0, \theta_k *}(v) \neq 0$, then $\theta_d(v) = -1$.

A remarkable fact [20,21] is that given a finite filtration it is possible to choose a consistent set of bases for $H_n(FN(f, \theta_k))$, $k = -1, \ldots, K$ such that each basis element has a well defined birth and death levels (θ_b, θ_d). By Eq. (2), if $v \in H_1(FN_I(f, \theta_k))$, then $\theta_d(v) \geq 0$ and there exists a unique element $\bar{v} \in H_0(FN_I(f, \theta_k))$ such that $(\theta_b(v), \theta_d(\bar{v})) = (\theta_K, -1)$.

The collection of all pairs (θ_b, θ_d) associated with the n-th homology group for the force networks are used to construct the β_n persistence diagram for the scalar field $f : CN \to [0, \infty)$. To provide some intuition concerning the process of going from an interaction network CN_I to a persistence diagram, consider a single chain of particles shown along the horizontal axis in Fig. 7(a). If two particles are not in contact, then the force acting between them is zero. Stated more formally, for any edge $\langle v_i, v_j \rangle \in CN_I$ if the particles corresponding to the vertices v_i and v_j are not in contact, then $f(\langle v_i, v_j \rangle) = 0$. Otherwise $f(\langle v_i, v_j \rangle)$ is defined by a step function such as that shown in Fig. 7(a). The extension of f to the vertex $\langle v_i \rangle$ is obtained via the definition in Section 5 and given by the black dot above the center of the particle corresponding to $\langle v_i \rangle$. Note that $f(\langle v_i, v_j, v_k \rangle)$ is always zero. Fig. 7(b) shows the sets $FN_I(f, \theta)$ for different values of θ. The set $FN_I(f, \theta) = \emptyset$ for $\theta > \theta_5$ and $FN_I(f, \theta_5)$ consist of a single connected component formed by one edge and two vertices. Finally for $\theta \leq 0$ the set $FN_I(f, \theta) = CN_I$ has a single connected component. The β_0 persistence diagram for the interaction force network is shown in Fig. 7(c).

We now explain what can be inferred about the interaction force network from the β_0 persistence diagram shown in Fig. 7(c). The fact that there are no points with birth coordinate larger than θ_5 indicates the absence of components experiencing force larger than θ_5. The set $FN_I(f, \theta_5)$ consists of a single connected component. Another connected component appears at θ_4. These two components merge together at θ_1. In the language of persistent homology the connected component is born at θ_4 and dies at θ_1 as indicated by the point (θ_4, θ_1) in the persistence diagram. The connected component that appears at θ_5 persists for all values $\theta \leq \theta_5$. In the persistence diagram it is represented by the point $(\theta_5, -1)$. There is one more geometric feature of the function f described by the point (θ_3, θ_2). This geometric feature corresponds to a pair consisting of a local maxima with value θ_3 and a local minima with value θ_2. This pair is visualized by the shortest double arrow in Fig. 7. Also the point (θ_4, θ_1) corresponds to a pair consisting of a local maxima and minima. The special point $(\theta_5, -1)$ encodes the value of the global maxima.

We now return to the example of the interaction force network associated with Fig. 6 of Section 5. This network is more complex than the one analyzed above. Not surprisingly the persistence diagrams for this network, shown in Fig. 8, contain more points. For the same reason as before, the β_0 persistence diagram, see Fig. 8(a), does not contain any points with the birth coordinate larger than θ_4. Four points with the birth coordinate θ_4 correspond to the four connected components that appear in $FN_I(f, \theta_4)$ (Fig. 6(e)). The
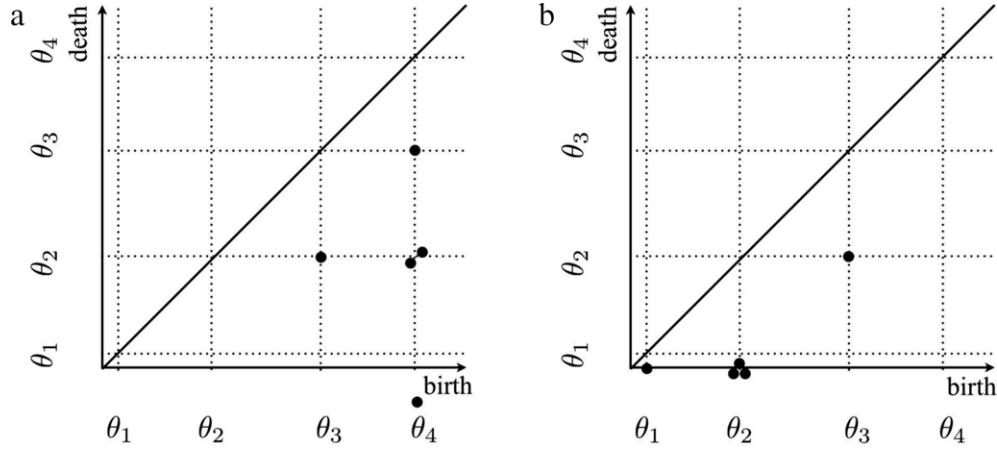
Fig. 8. (a) β_0 and (b) β_1 persistence diagrams for the force network shown in Fig. 6(c).

death coordinates of the points differ. This indicates that the components merge for different values of θ. The first merging appears for θ_3 (Fig. 6(f)) and is represented by the dot (θ_4, θ_3). Moreover a new component appeared at the level θ_3 and consequently merged with a preexisting component at θ_2 (Fig. 6(g)) as indicated by the dot (θ_3, θ_2). Also another two components that appear at θ_4 disappear at θ_3 hence there are two copies of the point (θ_4, θ_2). Finally there is only one connected component for all $\theta < \theta_2$. This component appears for $\theta = \theta_4$ and does not disappear. In the persistence diagram it is represented by $(\theta_4, -1)$.

The fact that the interaction force network contains loops can be inferred from the β_1 persistence diagram. The first loop appears at θ_3 (Fig. 6(f)) and is filled by triangular cells at θ_2 (Fig. 6(g)) as shown by the point (θ_3, θ_2) in Fig. 8(b). Another three loops appear at θ_2 and persist for all positive thresholds. Due to the definition of $\mathsf{FN}_I(f, \theta)$ all the loops are filled in for $\theta = 0$. So these three loops are represented by three copies of the point $(\theta_2, 0)$. The last loop appears at θ_1 (Fig. 6(h)) and also persists for all positive thresholds hence the point $(\theta_1, 0)$ belongs to the β_1 persistence diagram.

In general for the β_0 diagram a birth level $\theta_b(v)$ corresponds to the value of a local maximum that is associated with the birth of a connected component measured by an element $v \in H_0(\mathsf{FN}_I(f, \theta_b))$. As θ decreases this component grows until it meets, at a point associated with a local minimum or saddle, another component. Assume this other component is measured by the homology class v' and that the value of the local minimum (or saddle) is $\underline{\theta}$. If $\theta_b(v) < \theta_b(v')$, then $\theta_d(v) = \underline{\theta}$. In this case, $\theta_b(v) - \theta_d(v)$ measures the difference in height between the local maximum and local minimum and hence this difference can be used as a measure of how robust a feature is.

For the β_1 diagram a birth level $\theta_b(v)$ of the loop corresponding to an element $v \in_0(\mathsf{FN}_I(f, \theta_b))$ is the smallest value of f along this loop. If the loop is filled in by particles forming a crystalline zone, then $\theta_d(v)$ is the smallest value of f inside the region encompassed by the loop. If the interior of the loop is not completely filled in by a crystalline structure, then there must be a defect encircled by this loop. So the loop cannot be filled in with the triangles. Therefore, if we use a digital or position force network, then the loop never dies and $\theta_d(v) = -1$. For the interaction network all the loops are filled in at $\theta = 0$ and $\theta_d(v) = 0$.

We close this section with a formal definition of the persistence diagrams.

Definition 6.2. Let $\Theta = \{\theta_k \mid k = -1, \ldots, K\}$ and

$$\{\mathsf{FN}(f, \theta_k) \mid \theta_k \in \Theta\}$$

be a force network filtration over a complex CN. The associated *n-th persistence diagram* $\mathsf{PD}_n(f, \mathsf{CN}, \Theta)$ is the multiset consisting of the following points:

1. one point for each n-th persistence point (θ_b, θ_d);
2. infinitely many copies of points (θ, θ) on the diagonal.

Condition (1) of Definition 6.2 arises because distinct geometric features can appear and disappear at the same thresholds and thus there may be multiple copies of the same persistence pair. The necessity of condition (2) is made clear in Section 7.

We conclude this section with an observation. Let $\mathscr{PD}_n$ denote the set of all n-th persistence diagrams and $\mathscr{PD}$ the set of all persistence diagrams. Given a chain complex CN, let $M(\mathsf{CN}, [0, \infty))$ denote the set of monotone maps on CN. We can view persistence diagrams as a function

$$\mathsf{PD} \colon M(\mathsf{CN}, \mathbb{R}) \to \mathscr{PD} \tag{4}$$

or equivalently a collection of functions $\mathsf{PD}_n \colon M(\mathsf{CN}, \mathbb{R}) \to \mathscr{PD}_n$ defined by

$$\mathsf{PD}_n(f) = \mathsf{PD}_n(f, \mathsf{CN}, \Theta)$$

where $\Theta = \{\theta_k \mid k = -1, \ldots, K\}$ consists of the finite set of values obtained by f along with the convention that $\theta_{-1} = -1$ and $\theta_0 = 0$.

7. The space of persistence diagrams

The results concerning topological fidelity of the complexes $\mathsf{CN}_\bullet$ have, up to this point, been mostly negative. The introduction of persistence allows us to present positive results. In and of itself this suggests that the force network filtrations and their associated persistent homology provide more appropriate metrics for understanding force networks than the measurements performed at single thresholds. To obtain continuity results that guarantee that small changes in the measurement of the forces in the DGM lead to small changes in the persistence diagrams requires us to be able to measure the distance between persistence diagrams.

Definition 7.1. Let $\mathsf{PD} = \{\mathsf{PD}_i\}_{i=0}^n$ and $\mathsf{PD}' = \{\mathsf{PD}_i'\}_{i=0}^n$ be two collections of persistence diagrams. The *bottleneck distance* between PD and PD' is defined to be

$$d_B(\mathsf{PD}, \mathsf{PD}') = \max_{0 \le i \le n} \inf_{\gamma : \mathsf{PD}_i \to \mathsf{PD}_i'} \sup_{p \in \mathsf{PD}_i} \|p - \gamma(p)\|_\infty,$$

where $\|(a_0, b_0) - (a_1, b_1)\|_\infty := \max \{|a_0 - a_1|, |b_0 - b_1|\}$ and γ ranges over all bijections. Similarly, the *degree-q Wasserstein distance* is defined as

$$d_{W^q}(\mathsf{PD}, \mathsf{PD}') = \left[\sum_{i=0}^n \inf_{\gamma : \mathsf{PD}_n \to \mathsf{PD}_n'} \sum_{p \in \mathsf{PD}_n} \|p - \gamma(p)\|_\infty^q \right]^{1/q}.$$

As is indicated in [21], equipped with either the bottleneck or degree-q Wasserstein distance $\mathscr{PD}$ is a metric space. From now on we always assume $\mathscr{PD}$ is one of these metric spaces.
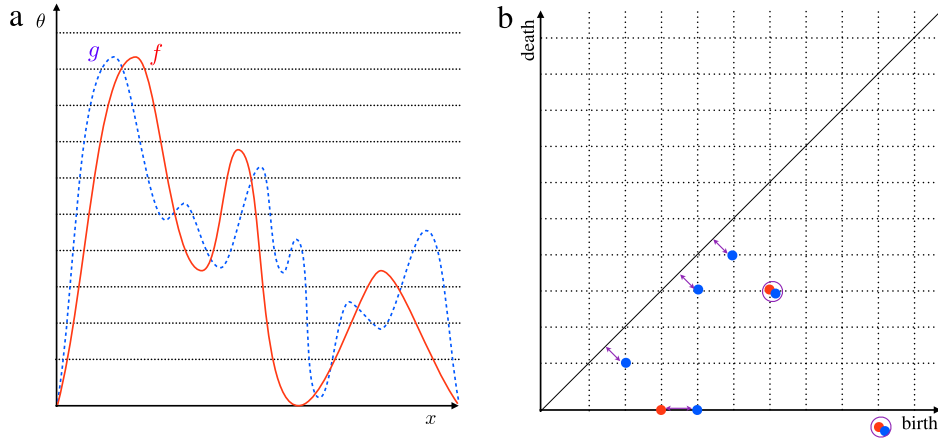
Fig. 9. (a) Two functions. Dashed (blue) represents a noisy perturbation of solid (red). (b) Associated persistence diagrams along with matching of persistence points satisfying the definition of bottleneck distance. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
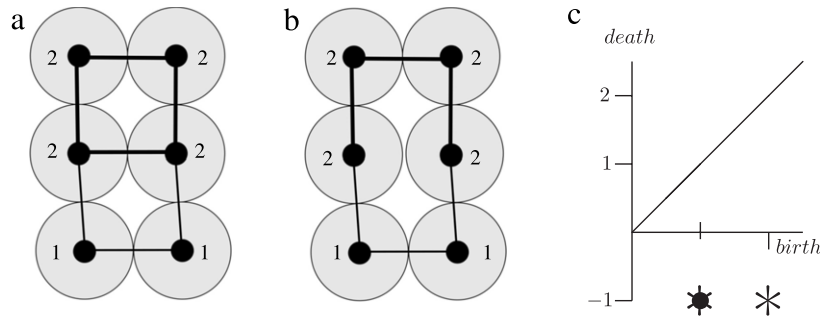

Fig. 10. An arbitrarily small change in positions can lead to an order one change in the bottleneck distance. Six particles with the magnitude of the force field indicated. Thick edges have force value 2, thin edges have force value 1. (a) Because of configuration of particles we see two loops. One loop appears at $\theta = 2$, the second loop appears at $\theta = 1$. (b) A perturbation of the configuration in (a) but the forces on the particles do not change. Only one loop appears at $\theta = 1$. (c) β_1 persistence diagrams. The stars at $(1, -1)$ and $(2, -1)$ correspond to the persistence points for (a) and the dot at $(1, -1)$ is the single persistence point for (b).

Fig. 9 shows two functions and their persistence diagrams. The function g is a small perturbation of f. The ability to match points in persistence diagrams with points on the diagonal, as shown in Fig. 9(b), suggests that small perturbations lead to small distances between persistence diagrams. In fact, it is proven in [21] that given a complex CN and two monotone functions $f, g : CN \rightarrow \mathbb{R}$ the bottleneck distance satisfies

$$d_B(\mathrm{PD}(f), \mathrm{PD}(g)) \leq \sup_{x \in X} |f(x) - g(x)|. \tag{5}$$

A similar result holds for the degree-q Wasserstein distance [21, Section VIII.3]. A more formal statement is as follows.

Theorem 7.2. *Given a complex* CN *let* $M(CN, \mathbb{R})$ *denote the set of monotone functions on* CN *equipped with the sup norm* $\| \cdot \|_\infty$. *Then*

PD: $M(CN, [0, \infty)) \rightarrow \mathcal{PD}$

defined by (4) *is a Lipschitz continuous map.*

Corollary 7.3. *The map* PD: $M(CN_I, [0, \infty)) \rightarrow \mathcal{PD}$ *is Lipschitz continuous.*

Corollary 7.3 implies that a small change in the forces, either through perturbation of the system or experimental error, results in a small change in the associated persistence diagrams. This is the long promised stability result. The failure of $CN_P^\spadesuit$ and CN_D to be stable with respect to perturbations follows from the fact that small changes of particle positions can result in changes of the underlying complex and thus Theorem 7.2 is not applicable. Fig. 10 demonstrates that it is possible, using $CN_P^\spadesuit$, for an arbitrarily small change in the position of the particles to lead to an order one change in the bottleneck distance.

8. Application of persistent homology to the results of discrete element simulations

The discussions of the previous sections provide a mathematical framework for studying force networks associated with DGM. In this section we apply these concepts to simulated data. We begin with a brief review of the numerical simulations and the computational tools employed in Section 8.1. We then analyze persistent homology on a variety of levels. First, in Section 8.2, we consider the stability of the persistence diagrams obtained from the digital, position, or interaction networks with respect to the numerical error. Then, in Section 8.3 we discuss how the results depend on the choice of complex used, and therefore on the quality of input data. Section 8.4 presents few examples that outline how the information contained in individual persistence diagrams can be related to physically observable properties of the DGM network considered. We note that while the focus of our discussion is on the networks defined using normal force between the particles, we also briefly discuss persistence diagrams obtained using tangential forces (in frictional systems). Finally, in Section 8.5 we briefly illustrate the application of the concept of distance between persistence diagrams to DGM networks.

8.1. Simulations used

We perform a series of discrete element simulations (DES) similar to our previous works [18,19]. For the present paper, we consider a set of about 2000 circular particles contained in a square domain with rough walls composed of monodisperse particles. The system is slowly compressed allowing for a change of packing fraction, ρ, between 0.6 and 0.9. Initially the particles are given

random velocities and are placed on a square lattice. The equations of motion are integrated using a fourth order predictor–corrector scheme. We implement the Cundall–Strack model for static friction which includes normal and tangential forces at the contact [27]. For frictionless system, the contact force reduces to a normal force with a spring and viscous damping term. In general, we use polydisperse particles where the particle sizes are chosen from a uniform distribution with width $r_p = (r_{max} - r_{min})/r_{ave}$, where r_{ave} is the mean particle radius. The coefficient of restitution measuring energy loss is given the value of $e_n = 0.5$, and the coefficient of static friction is either $\mu = 0.5$ for the frictional case, or $\mu = 0$ for the frictionless one. See [18,19] for more details.

We focus on a system of particles with $\rho \approx 0.86$, except if specified differently. This ρ is beyond ρ_c, at which jamming transition occurs. (Note that the $r_p = 0$, $\mu = 0$ system has the highest ρ_c, which is under the implemented protocol, and for the considered realization at $\rho \approx 0.85$. More extensive discussion of transition through jamming, as well as averaged (over initial conditions) values for ρ_c, are given in [19].) Therefore, the particles are packed close enough so that most of the particles belong to the same connected component of the position network CN_P.

For the ρ's of interest, we extract the magnitude $\psi_{i,j}$ of the normal force between any two particles p_i and p_j. The values $\psi_{i,j}$ completely determine the interaction network FN_I. To construct FN_P and FN_D the positions of the particles need to be extracted as well. The value ψ_i assigned to the particle p_i is the total force experienced by this particle, i.e.,

$$\psi_i := \sum_{\{j \mid \langle i,j \rangle \in CN_P^{(1)}\}} \psi_{i,j} = \sum_j \psi_{i,j}.$$

For consistency with our previous works [18,19], we normalize the function $f : CN \to \mathbb{R}$, defined in Section 5, by dividing it by the average force $\hat{f}$ defined as follows: for the interaction force network

$$\hat{f}_I = \frac{1}{M} \sum_{i,j=1}^N \psi_{i,j}, \tag{6}$$

and for the position and digital force networks

$$\hat{f}_P = \frac{1}{N} \sum_{i=1}^N \psi_i = \frac{1}{N} \sum_{i,j=1}^N \psi_{i,j}, \tag{7}$$

where M is the number of non zero force interactions, $\psi_{i,j}$, and N is the number of particles. Note that the average number of contacts is $Z = M/N$ and hence

$$\hat{f}_P = Z\hat{f}_I. \tag{8}$$

We have produced open source software [24] that is used to encode this procedure and produce a force networks filtration $\{FN(f, \theta_k) \mid \theta_k \in \Theta\}$. The persistent homology of each filtration is computed using the open source software Perseus [28,25]. We note that the size of the digital complex, CN_D, is considerably larger than the size of the position or interaction complexes, implying that the computational cost of analyzing CN_D is much larger as well. To give a sense of the time needed to perform the types of computations we remark that using a 2.53 GHz processor to compute the persistence diagrams for the position, digital, and interaction force network required 25, 97, and 43 s, respectively. The worse case complexity of computing the bottleneck distance between the persistence diagrams PD_1 and PD_2 is $O((n_1 + n_2)^2)$ where n_i is the number of generators in PD_i. For the Wasserstein distance the complexity is even higher $O((n_1 + n_2)^3)$. Therefore the time required for computing the distances strongly depends on the number of generators in the persistence diagrams. In practice the number of generators is rather small for CN_I and CN_P. In our case there are typically

a few hundred persistence generators. Thus both distances can be computed in a couple of seconds. However the digital networks contain a large amount of artificial loops shown in Fig. 3(a) and (c) (this number typically increases with the resolution) and the runtimes are much longer. For the resolution 1000×1000 runtime required to compute the bottleneck distance is 10 min and for the resolution 2000×2000 it is 30 min. Finally we needed three weeks to compute the Wasserstein distance between two persistence diagrams for digital networks with the resolution 1000×1000. We stopped computation of this distance for the digital network with higher resolution after it had not terminate within three weeks.

8.2. Stability of persistence diagrams

Fig. 11 shows the three networks for $r_p = 0$, $\mu = 0$ system. The associated persistence diagrams are depicted in Fig. 12. We will discuss some features of these networks and diagrams in what follows; to start with, we ask the following question: how stable is the information contained in the persistence diagrams with respect to an error in the input data?

The numerical simulations and the extraction of particle positions and normal forces are done using double precision floating point numerics. We then compute d_B and d_{W^1} distances between the original and perturbed persistence diagrams. The results are given in Table 1. The relatively small values associated with the interaction network are predicted by Corollary 7.3. The fact that the values for d_{W^1} are significantly larger than d_B for each type of network is not surprising since d_B distance measures the single largest change in the network while d_{W^1} is sensitive to all local perturbations that may be occurring.

Behavior of the position network with respect to small perturbations lacks the stability of the interaction network. A measurement error at the third decimal place radically changes the network, with the d_B distance three orders of magnitude larger than the error introduced. This large difference is caused by the phenomena shown in Fig. 10. To show this, note that d_{W^1} is several orders of magnitude larger than d_B, implying that there must be many locations in the network at which the local maxima and minima of the forces change due to introduced error.

We have no theoretical results that explain the relative differences in perturbations of distances between the persistence diagrams associated with the position and the digital force networks. The digital force network was constructed using resolutions of 1000×1000 and 2000×2000 pixels. Given the size of the domain, each pixel in 1000×1000 case represents a measurement to approximately three significant figures. We hypothesize that this explains the relatively small (as compared with the position network) d_B distance. The d_{W^1} distance for truncated data is larger for the digital (1000) force network than for the position network. This seems to be connected to a significant drop in the number of β_1 generators corresponding to the artificial loops (see Fig. 3) in the perturbed digital (1000) force network. The original network contains 6344 β_1 generators while the network obtained by truncation to three significant digits only 4487. In contrast, the difference between the number of generators for the position networks is 340. Thus the difference is much smaller and so is the d_{W^1} distance.

The sensitivity of digital complexes to small perturbations demonstrated in Fig. 3 suggests that the larger distance value for the 2000×2000 digital complex with truncation at three significant digits should not come as a surprise. The different distance values for different digital complexes raises another issue; how sensitive is the persistence diagram to the resolution of the digital network? We consider this issue using the system shown in Fig. 11. Computing with the original numerical data at a resolution of 2000×2000 pixels, we find that the β_0 persistence diagrams for the digital networks are almost identical. Comparison of the β_1 persistence diagrams reveals that the number of loops is around
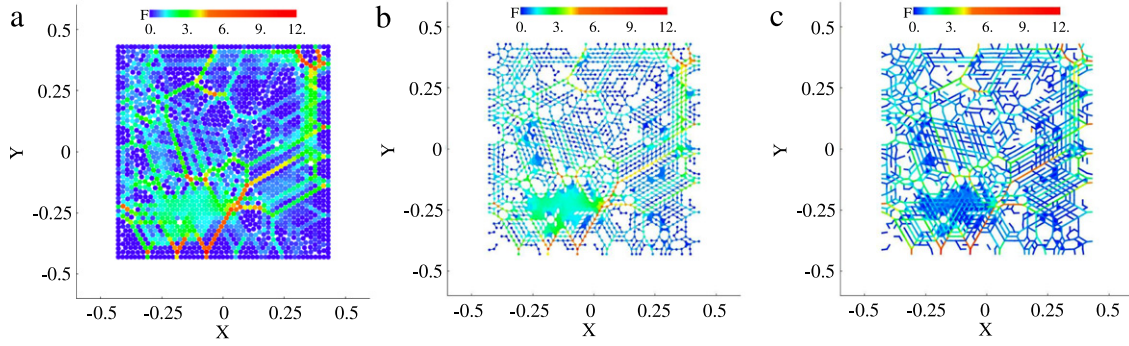
Fig. 11. Different force networks for monodisperse, $r_p = 0$, frictionless, $\mu = 0$, system at the packing fraction, $\rho = 0.86$. (a) Digital force network, (b) position force network and (c) non zero simplices of the interaction force network.

Table 1
The distance between the persistence diagram of original networks shown in Fig. 12, and the persistence diagram after truncation of numerical data to 2 or 3 significant digits. The bottleneck distance d_B measures the single largest difference between the persistence diagrams while the Wasserstein d_{W^1} distance is a sum of all differences between the diagrams (see Definition 7.1). Computation of d_{W^1} for the digital (2000) network did not terminate within three weeks using 2.53 GHz processor.

Metric	Truncation	Network type			
		Interaction	Position	Digital (1000)	Digital (2000)
d_B	3	0.0004	2.6151	1.2946	2.3339
	2	0.0035	2.6387	3.4672	3.4672
d_{W^1}	3	0.1806	439.01	659.4	–
	2	1.8447	689.79	2226.1	–

30% larger for the higher resolution. We have verified that this increase is caused by the formation of extra loops at the places where the particles are close to each other; essentially the phenomenon indicated in Fig. 3(a).

8.3. Force networks as a function of complex type

Figs. 11 and 12 demonstrate that the digital, position and interaction force networks of a single system of particles can be quite different. The idea behind the construction of the digital and position force networks is the same, the difference arises from the fact that they are based on different complexes that provide different approximations of the geometry of the system of particles. Thus, to focus on the essential differences we mostly restrict our discussion to a comparison of the position and interaction force networks. Fig. 13 provides an enlarged view for three different subregions of the position and interaction force networks of Fig. 11(b) and (c). The position force network is defined in terms of the vertices and thus the corresponding figures include the magnitude of the force there. The vertices are not highlighted in the interaction force network since the value of the force on the edges is used to define the values on the vertices.

Fig. 13(a) and (b) are typical of a region in which we see crystalline structure or equivalently a region over which there are no defects. Observe that in this crystalline region the normal forces for the position force network are significantly larger than those of the interaction force network. This has to do with large number of contacts (6), so the sum of the forces on each particle is high. Note that the forces are rather uniform in the crystalline zone and $\psi_i \approx 6 \max_j \{\psi_{i,j}\}$. Let f_I and f_P denote the forces in the interaction and particle force networks, respectively. Then

$$f_I(i) = \frac{\max_j \{\psi_{i,j}\}}{\hat{f}_I} = \frac{Z \max_j \{\psi_{i,j}\}}{\hat{f}_P} \approx \frac{Z \psi_i}{6\hat{f}_P} = \frac{Z}{6} f_P(i).$$

Except for the perfect crystal the value of Z is less than 6. For the network shown in Fig. 11 we computed that $Z \approx 3$. Therefore $f_I(i) \approx \frac{1}{2} f_P(i)$.

We now consider a part of the domain where we find sets of particles interacting by large forces, resembling a 'force chain'.

In this case, as can be seen along the orange chain in Fig. 13(d), the position force network tends to report a lower magnitude of force than the interaction force network, compare with the red chain in Fig. 13(c). Observe that along the red chain of particles in Fig. 13(c) each particle typically has contact with 2 or 3 other particles. Therefore $\psi_i \leq Z \max_j \{\psi_{i,j}\}$. By Eq. (8) and the inequalities stated in this paragraph we obtain

$$f_I(i) = \frac{\max_j \{\psi_{i,j}\}}{\hat{f}_I} = \frac{Z \max_j \{\psi_{i,j}\}}{\hat{f}_P} \geq \frac{\psi_i}{\hat{f}_P} = f_P(i).$$

An added effect is that a single continuous chain of strong force interactions in the interaction force network is reported to be a collection of shorter chains in the position force network (see Fig. 13(e) and (f)). An immediate consequence is that we expect to see more points with relatively large birth values in the β_0 persistence diagram of the position force network than in the β_0 persistence diagram of the interaction force network. This is confirmed by counting the number of points in the β_0 persistence diagrams of Fig. 12(b) and (c) with birth value greater than a given value.

Fig. 13(e) and (f) demonstrate another important difference between the position and interaction force networks. In Fig. 13(f) there is a strong branching chain that forms a (red) loop. The values at the edges forming the loop are stronger in the position force network. For the interaction force network, value at the edge next to the crystalline region is small. A larger value in the position force network is caused by the presence of particles in the crystalline region. This difference implies that loops are formed at lower force levels in the interaction force networks as compared to the position force networks which, in turn, implies that there should be fewer points with relatively large birth values in the β_1 persistence diagram of the interaction force network as compared to the position force network. This is corroborated by Fig. 12(b) and (c).

In the rest of this paper we use the following convention. If the feature persists until the zero threshold then we set the death coordinate to minus one. This only impacts the persistence diagram for the interaction force network, allowing for simple visual identification of the defects. This convention is solely for visualization proposes and is not used for distance computations.
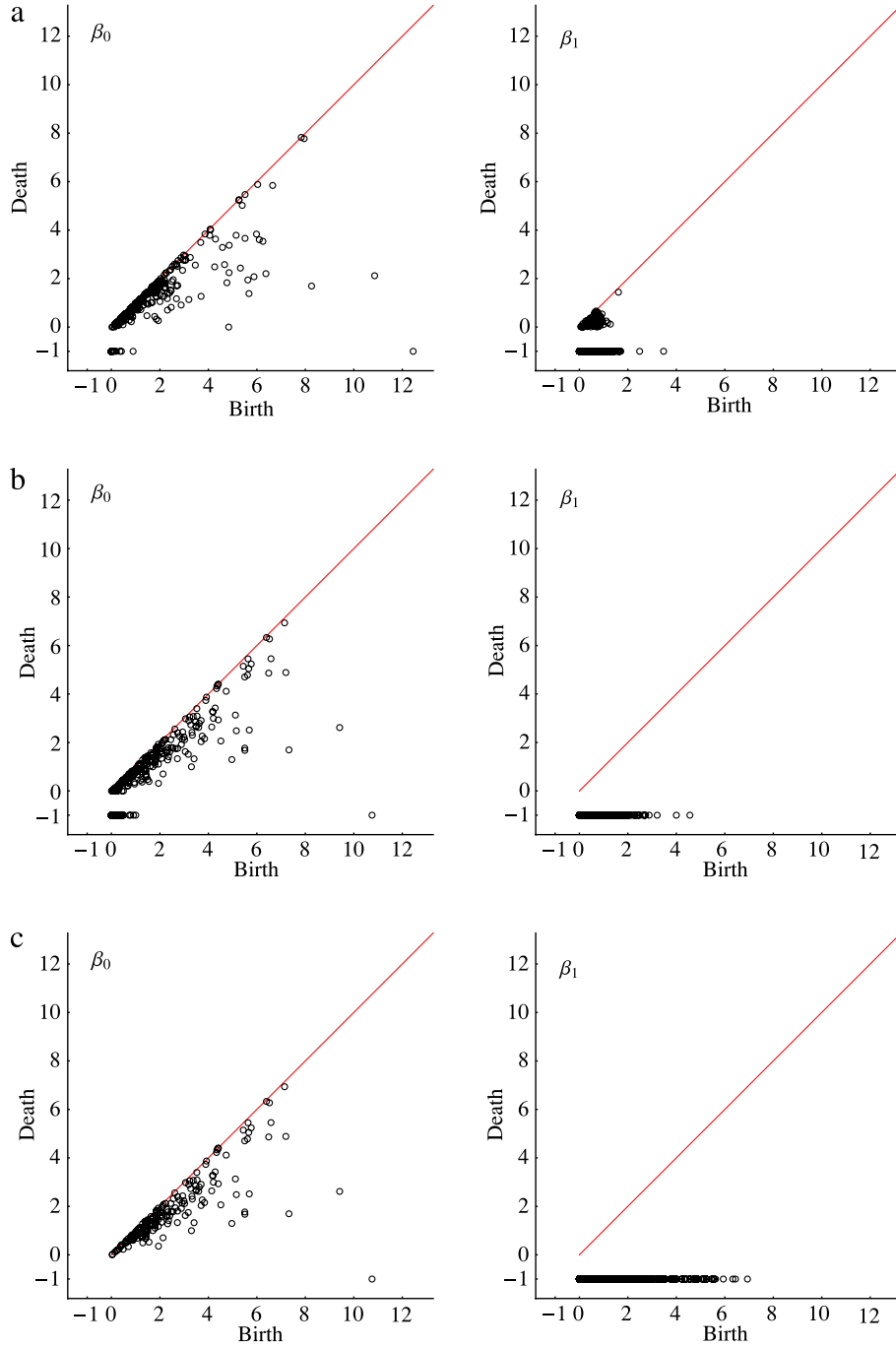
Fig. 12. Persistence diagrams for $r_p = 0$, $\mu = 0$ system at $\rho = 0.86$ shown in Fig. 11. Persistence diagrams for (a) digital force network based on 1000×1000 pixels, (b) position force network, and (c) non zero simplices of the interaction force network.

Another striking difference between the position and the interaction force networks is that in the β_1 persistence diagram of the position force network the death value for all points is -1, i.e., once a loop is formed it never dies. This is not the case for the β_1 persistence diagram of the interaction network. There is no reason that loops in the position force network cannot die, but a possible explanation is as follows. The death of loops is associated with the appearance of 2-simplices (triangles) that is indicative of crystalline structure. Consider a single 2-simplex $\langle v_i, v_j, v_k \rangle$ and assume that

$$f(\langle v_i \rangle) > f(\langle v_j \rangle) > f(\langle v_k \rangle).$$

Given the definition of the position force network, the vertex $\langle v_i \rangle$ appears first, followed by the vertex $\langle v_j \rangle$ and the edge $\langle v_i, v_j \rangle$. Finally, the vertex $\langle v_k \rangle$, the edges $\langle v_i, v_k \rangle$ and $\langle v_j, v_k \rangle$ and the

2-simplex $\langle v_i, v_j, v_k \rangle$ are all included at the same step. Thus there is no opportunity for a loop consisting of three edges to be generated. A similar argument can be made for the interaction network and hence loops that appear in the persistence diagrams must involve multiple edges. If we think of this sequence of edges as a 'force chain', then the previous argument suggests that for the position force network this chain is more likely to contain edges of lower magnitude than in the interaction force network. At the same time, we have observed that in crystalline regions the force magnitudes at the particles are larger in the position than in the interaction force network. These two observations suggest that in the position force network it is difficult to construct a loop in a crystalline region that surrounds vertices with lower forces.
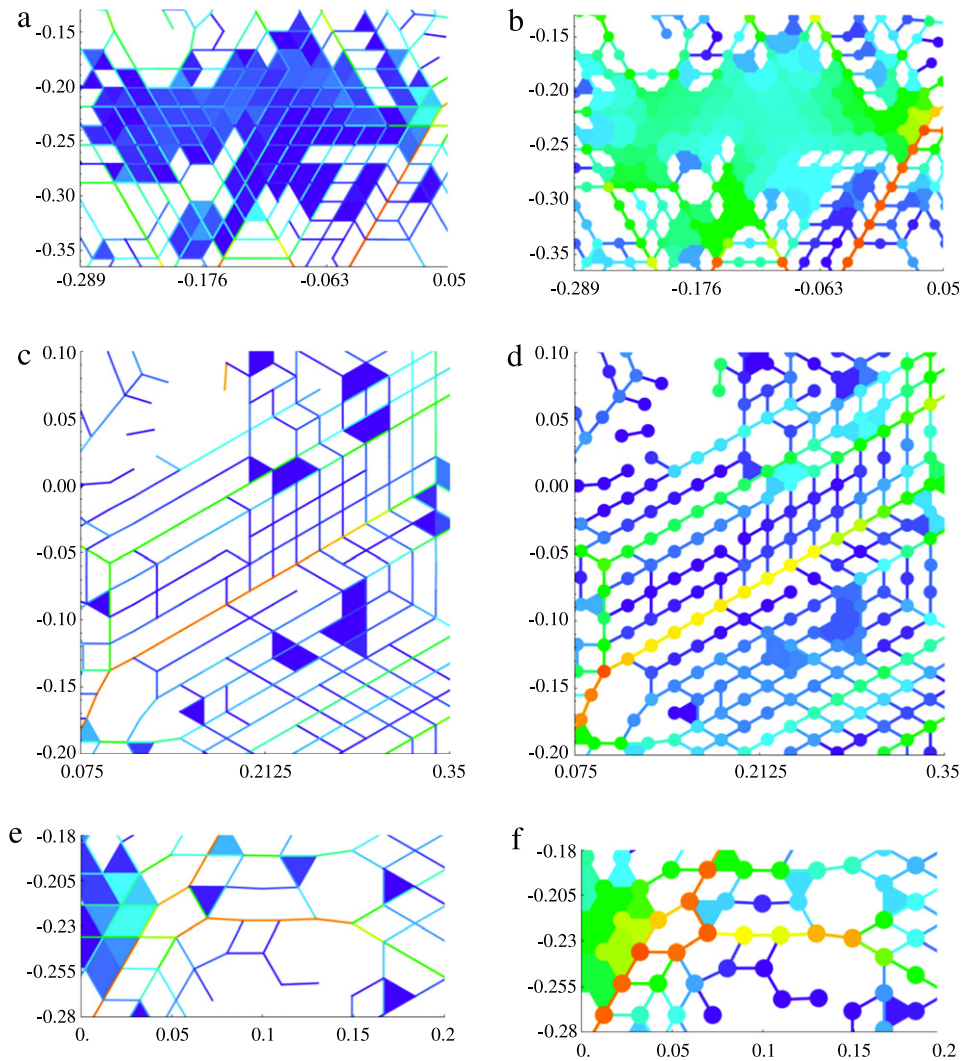
Fig. 13. Enlarged views, using the same color scheme, of three different subregions of the interaction (a), (c), (e) and position (b), (d), (f) force networks of Fig. 11(b) and (c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Let us now consider briefly the digital force networks. Though not obvious from Fig. 12 the β_0 persistence diagrams indicate that the digital force network have less points than those for the position or interaction network. This is due to the fact that, as is discussed in Section 3, our construction of the digital network artificially inflates area associated with each particle and hence it is possible for distinct components in the position or interaction force network to form a single component in the digital force network. This effect is particularly relevant in the context of rattlers, the particles that do not experience any force. In contrast, the number of persistence generators in β_1 persistence diagrams is larger for digital force networks due to the formation of the artificial loops, again as described in Section 3.

In this section we have shown that the information that is available about a granular system influences the detected properties of the force networks. The interaction force networks, that are based on the information about the forces between the interacting particles, provide most precise and reliable information and we concentrate on these networks from now on.

8.4. Comparison of different systems via persistence diagrams

The most direct means of applying persistence diagrams is to use them to distinguish and/or interpret the global force structures of DGM composed of particles characterized by different physical properties. This approach was followed in our recent work [19], where we considered the number of generators in different parts of the persistence diagrams as the systems that differed by their physical properties were compressed. Here we describe how persistence diagrams can be used to extract a significant amount of information about interaction force networks by considering only two snapshots, at $\rho = 0.86$, of the following two systems: a monodisperse frictionless ($r_p = 0$, $\mu = 0$) and a polydisperse frictional ($r_p = 0.4$, $\mu = 0.5$) system. We already know that these two systems behave differently under compression [19], and we use persistence diagrams to illustrate these differences. The reader should note that the focus here is on the illustration of the technique and on the interpretation of a limited set of results: more general (although less detailed) discussion that concentrates on the physical interpretation can be found in [19]. In particular, note that the jamming points for these two systems differ, so that by considering the same ρ, we consider two systems which are at different distances from the jamming transition.

We begin by assigning physical meaning to the location of persistence points in the persistence diagrams. Fig. 14 shows a persistence diagram divided into five regions. With the exception of the region labeled defects, the location of the division lines is intended to be either system specific or conceptual.

Roughness. The geometric features corresponding to the points in the region labeled roughness persist only over a small range of
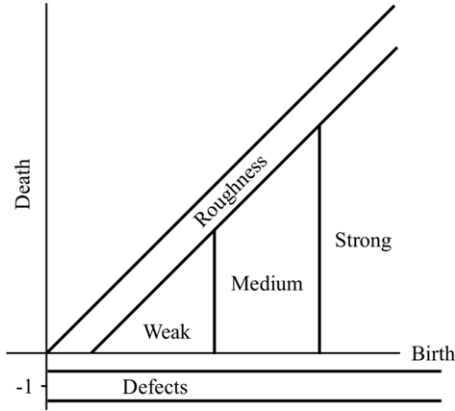
Fig. 14. Persistence diagram divided into regions. Explanation of the regions is provided in the text.

force values. For instance the point (θ_3, θ_2) in the β_0 persistence diagram shown in Fig. 7 corresponds to a feature that persists over a relatively short range compared with the other features in the same figure. There are at least two different interpretations of the points in the β_0 persistence diagram that lie in this region. The first is to treat these points as noise, i.e. a byproduct of the imperfect measurements of the normal forces between the particles. While this may be appropriate for many experimental settings, the data

represented in Figs. 11, 12 and 15 come from simulations, and the errors are small. This fact leads to the second interpretation, which we adopt, that this region of the persistence diagram provides information about small fluctuations of the forces. These fluctuations can be interpreted as a measurement of how rough or bumpy the normal force landscape is, e.g. should we view the surface of the landscape as being made of glass or sandpaper? Therefore by comparing Figs. 12(c) and 15 we conclude that for the considered ρ, $r_p = 0.4$, $\mu = 0.5$ system is rougher than the $r_p = 0$, $\mu = 0$.

Strong. To understand the region labeled as strong, observe that the image in Fig. 15(a) of the forces for the $r_p = 0.4$, $\mu = 0.5$ system does not contain any red simplices, implying that there are no strong force interactions. In contrast, such red simplices are present in the $r_p = 0$, $\mu = 0$ system displayed in Fig. 11(c). This difference can be inferred from the β_0 persistence diagrams shown in Figs. 12 and 15. For the $r_p = 0.4$, $\mu = 0.5$ system there are no persistence points with the birth value larger than 3 (in terms of average interaction force) and only a few points with birth value larger than 2.5. Thus, depending on the exact cut-off there are no or at most few points in the region marked strong for the $r_p = 0.4$, $\mu = 0.5$ system, in clear contrast to the $r_p = 0$, $\mu = 0$ system, for the considered value of ρ.

Medium. If we take the left division marker for the medium regime in Fig. 14 to be 1, then the persistence points in the medium and strong regions provide information about the geometry of strong contacts. For the $r_p = 0.4$, $\mu = 0.5$ system, we see a
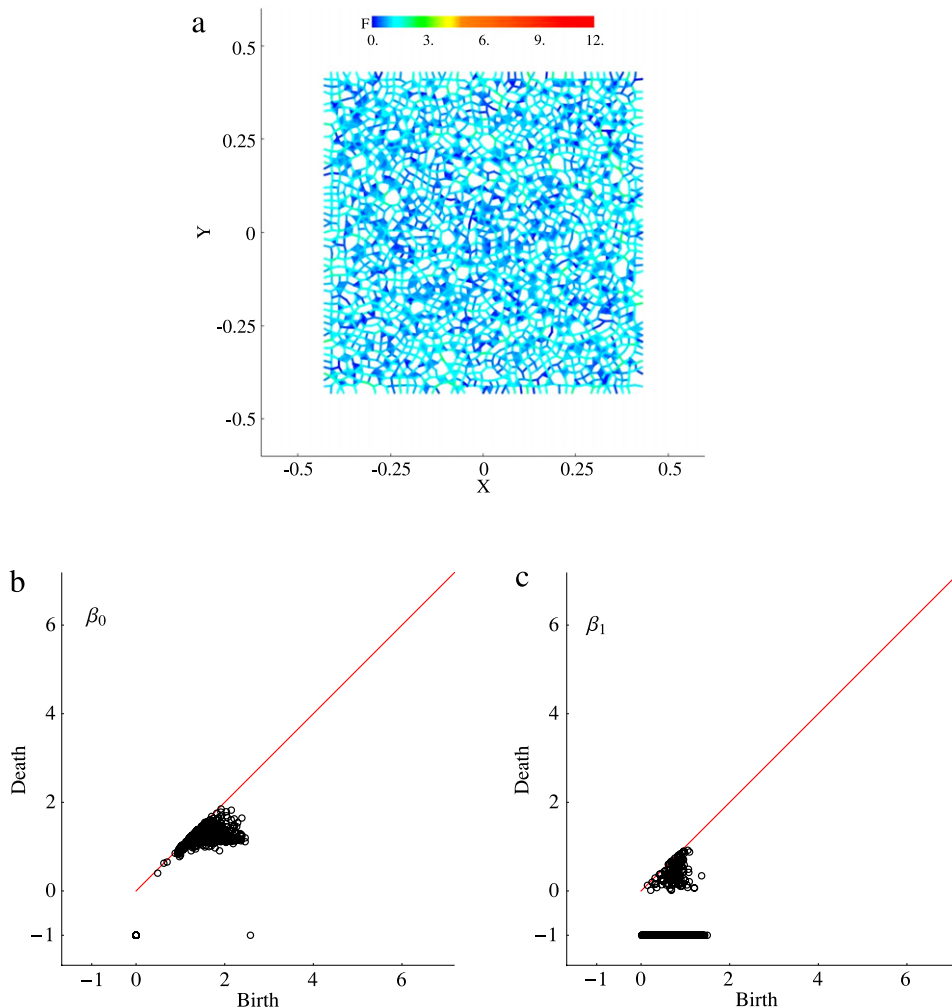

Fig. 15. (a) Non-zero simplices of the interaction force network based on normal force. (b) β_0 (c) β_1 persistence diagrams for the $r_p = 0.4$, $\mu = 0.5$ system at $\rho = 0.86$.

large number of β_0 persistence points that are born between 1 and 2.5 and die before 0.8. This suggests a landscape consisting of moderately high peaks separated by moderately high valleys. To continue the geographic metaphor, the $r_p = 0.4$, $\mu = 0.5$ force chain network takes place on a high plateau. In contrast, the $r_p = 0$, $\mu = 0$ system has fewer moderately high peaks, but they are separated by much deeper valleys since there are points with death values below 0.6. Therefore, we conjecture that, for the considered ρ, landscape for the $r_p = 0$, $\mu = 0$ system has fewer peaks (but some of them are strong) than that of the $r_p = 0.4$, $\mu = 0.5$ landscape, and these peaks are in general much more isolated and more likely to be separated by valleys of much weaker forces.

Let us summarize the previous paragraphs in terms of 'force chains'. There are no force chains with force value exceeding three times the average value for the considered $r_p = 0.4$, $\mu = 0.5$ system. These kind of force chains are present only in the $r_p = 0$, $\mu = 0$ system. On the other hand the number of 'force chains' with strongest link exceeding the average force is much larger for $r_p = 0.4$, $\mu = 0.5$ system. Fact that $\theta_d > 0.8$ for most of the points in the β_0 diagram implies that in the $r_p = 0.4$, $\mu = 0.5$ system there are links connecting the 'force chains' with forces larger than 0.8 times the average force. The connections between the 'force chains' in the $r_p = 0$, $\mu = 0$ system tend to be much weaker.

Defects. In a β_0 persistence diagram, each point in this region corresponds to a distinct connected component of contact network. In the context of the $r_p = 0.4$, $\mu = 0.5$ system, these points mostly correspond to the rattlers. This conclusion is obtained by observing that aside from the single persistence point corresponding to a large birth force, that corresponds to the component containing most of the particles, the persistence points in the defects region have a birth value of 0, indicating that they are not experiencing any normal force. This is quite different from the $r_p = 0$, $\mu = 0$ system. In this case we have persistence points in the defects region with non-zero birth forces. This implies the existence of small clusters of particles (a single separated particle does not experience a force) that are not interacting with the dominant particle cluster. Close inspection of the interaction force network in Fig. 11(c) reveals these small components.

Weak. Finally, the points in the region labeled weak represent small clusters of particles, interacting weakly with the dominant particle cluster. Inspection of the β_0 persistence diagram reveals existence of these clusters for the $r_p = 0$, $\mu = 0$ system while they are virtually absent for the $r_p = 0.4$, $\mu = 0.5$ system, for the considered ρ.

The defects region of the β_1 persistence diagrams provides additional information. As indicated in Section 8.3, β_1 persistence points lie in the defects region if and only if they correspond to the loops that enclose non-crystalline regions. There are about twice as many persistence points in the defects region in the $r_p = 0.4$, $\mu = 0.5$ system as compared to the $r_p = 0$, $\mu = 0$ one. This suggests that the $r_p = 0.4$, $\mu = 0.5$ system is more likely to support defects, for the considered value of ρ. At the same time there are 50% more points in β_1 persistence diagram that are not in the defects region for the $r_p = 0.4$, $\mu = 0.5$ system. These persistence points correspond to the loops that are filled in by 2-dimensional simplices and thus must be contained within crystallized regions. Thus this difference in the number of persistence points suggests that the $r_p = 0.4$, $\mu = 0.5$ system contains a multitude of small crystalline regions as opposed to the $r_p = 0$, $\mu = 0$ system. This is corroborated by a careful examination of the force networks in Figs. 11(c) and 15.

The finding above may appear surprising since it is known that frictionless monodisperse systems may crystallize, as it was also shown in our earlier work [18]. The resolution of this apparent contradiction is that polydisperse systems considered appear to form large number of small (involving only a few particles)

crystalline regions, in contrast to monodisperse frictionless ones that are expected to form large crystalline zones involving many particles [18].

Tangential forces. Analogous analysis can be done for the force network based on tangential forces. Fig. 16 shows the tangential force network together with the persistence diagrams for the same particle configuration as in Fig. 15; note that the tangential forces are normalized by the average tangential force. While one could argue that a visual comparison of Figs. 15(a) and 16(a) is not particularly insightful, examining the persistence diagrams gives much more information. The significant and crucial difference is that the tangential forces go to much larger values than the normal ones and consequently the generators for connected components as well as loops cover much larger range. This is particularly visible when considering β_0 generators shown in Fig. 16(b). Therefore, the analysis based on the persistence diagrams suggests that the landscape defined by tangential forces is significantly more 'mountainous' with much higher mountain tops. Here, we give only an example of tangential force results; future work should analyze how tangential forces evolve as system is exposed to shear or compression, as done for normal forces in [19].

8.5. Comparison of different systems via distances

Up to this point we have focused on the interpretation of individual persistence diagrams. However, we believe that the most significant value of this technique will come through the analysis of large sets of persistence diagrams. We present a simple example of this idea here, leaving more detailed investigations for future work.

A metric in the space of persistence diagrams measures the level of similarity between different states of DGM. If the states are similar the distance is small. By using different metrics we can access different notions of similarity. Bottleneck distance informs us about a single dominant change. Notion of similarity induced by this distance has a local character since it measures only the largest difference and ignores all the other changes. Overall similarity of two states is measured by Wasserstein d_{W^1} distance which sums up all the differences between the states. Since large number of small differences can result in large d_{W^1}, this metric is sensitive to noise. This sensitivity can be mitigated by using d_{W^q} for $q > 1$.

We use these metrics to analyze and compare the time evolution of the two considered DGM systems; to facilitate the comparison, we chose two different ρ's, that are at the similar distance (in packing fraction) from the jamming points. In the present work, the jamming point is loosely defined as a packing fraction, ρ_J, at which $Z \approx 3$. We use $\rho = 0.87$ for $r_p = 0$, $\mu = 0$ with $\rho_J = 0.853$ and $\rho = 0.786$ for $r_p = 0.4$, $\mu = 0.5$ with $\rho_J = 0.786$. The difference between ρ's of the consecutive states is fixed to $\Delta \rho = 8 \times 10^{-5}$ in both cases.

Figs. 17 and 18 show the distances between all these states (100 of them). Let us focus first on the parts (a) of these two figures, showing the results for $r_p = 0$, $\mu = 0$ system. Figs. 17(a) and 18(a) are the color coded distance matrices D for the d_B and d_{W^1} distances. The entry $D(i, j)$ is the distance between the state i and j. Thus this is a symmetric matrix, $D(i, j) = D(j, i)$ and $D(i, i) = 0$ (black) for all i, j.

Figs. 17(a) and 18(a), upper left corners, show that for the first (17) states considered, the states are similar suggesting that the system is evolving slowly. Though less pronounced this is also the case for the last 20 or so states (bottom right corner). The fact that the upper right corner is yellow indicates that the structure of the forces has evolved over the range of ρ's (10^{-3} in the considered case). Therefore, the persistent homology can be used to capture the evolution of the force networks.

These comments are applicable to both the d_B and d_{W^1} distance matrices. However, the metrics carry different information. The
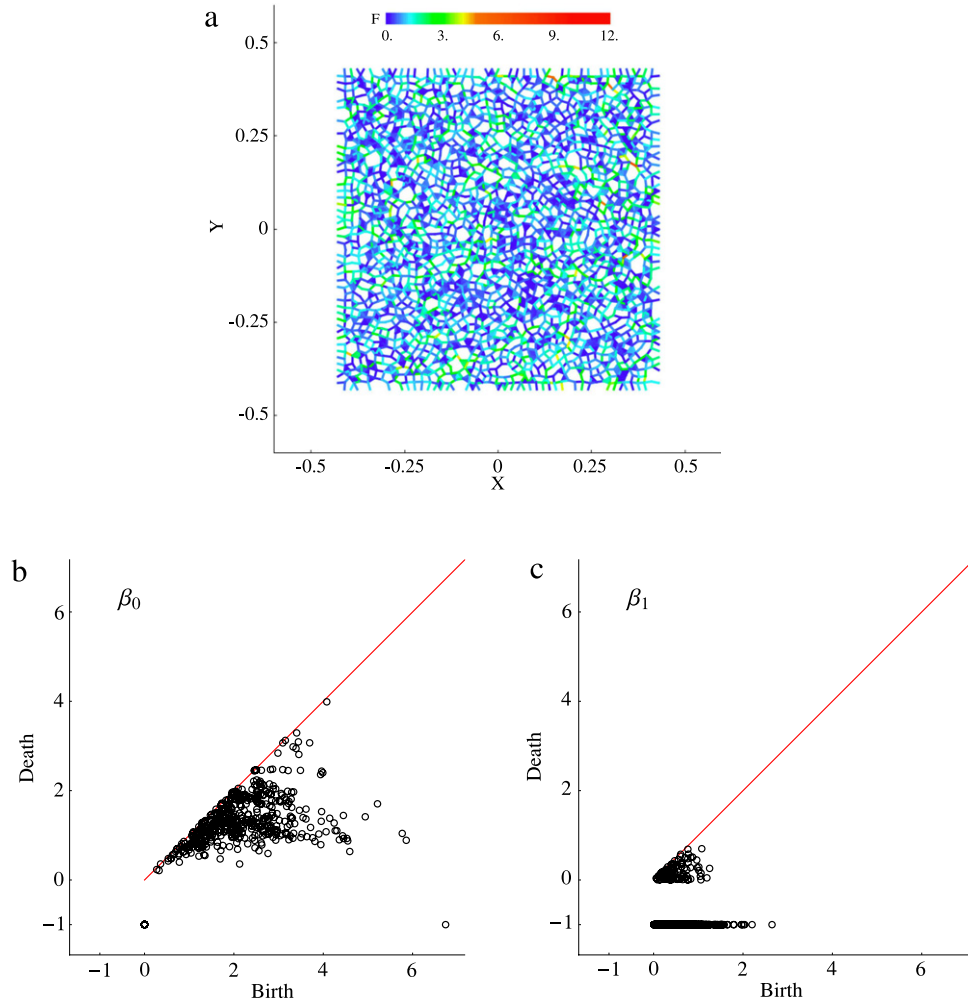
Fig. 16. (a) Non-zero simplices of the interaction force network based on tangential force. (b) β_0 (c) β_1 persistence diagrams for the $r_p = 0.4$, $\mu = 0.5$ system at $\rho = 0.86$ (the same particle configuration as in Fig. 15). Note different range in (b) compared to Fig. 15(b).
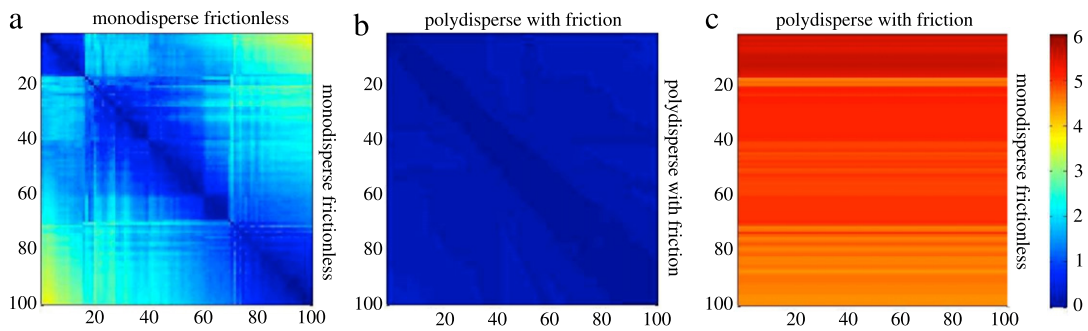

Fig. 17. Distance matrices for the bottleneck distance. The value $D(i, j)$ is the d_B distance between the state i and j for the (a) $r_p = 0$, $\mu = 0$ system (b) (a) $r_p = 0.4$, $\mu = 0.5$ system. (Note that by construction D is a symmetric matrix.) (c) Comparison of the $r_p = 0$, $\mu = 0$ system with the $r_p = 0.4$, $\mu = 0.5$ system. The states of the $r_p = 0.4$, $\mu = 0.5$ system change along the horizontal axis and the states of the $r_p = 0$, $\mu = 0.5$ along the vertical axis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

d_B matrix implies that for the first 17 or last 20 steps there is no single location at which a large change in the force structure occurs. The corresponding d_{W^1} matrix entries, which are approximately 50 times larger, suggests that (i) many small changes in the force structure are taking place from step to step, and (ii) that from step to step roughly the same number of small changes are occurring. To provide an analogy; there are ripples on the pond, but the number of ripples are not changing with time.

The abrupt color change (increase in distance) at $D(17, 18)$ indicates that there has been a significant structural change of the forces of the DGM from state 17 to 18. Another significant transition can be seen between 70 and 71. The evolution in the region between the states 20 and 70 has a similar character. Parts with slow evolution are separated by sudden transitions. The fact that we see these transitions more clearly in the d_B matrix as opposed to the d_{W^1} matrix indicates that when significant changes occur, they are
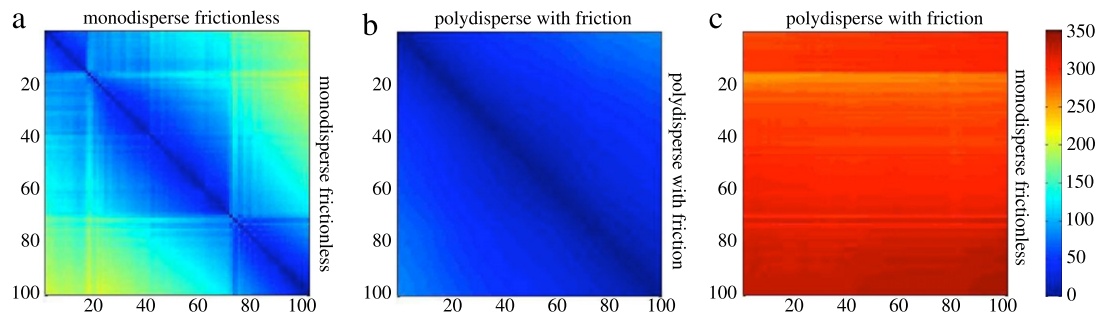
Fig. 18. Distance matrices for the degree-1 Wasserstein distance, considering the same systems as in Fig. 17. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

localized in space. Conversely, since the two dominant transitions at 17 and 70 are clearly seen in the d_{W1} matrix suggests that they are more global in nature.

The distance matrices for $r_p = 0.4$, $\mu = 0.5$ system are shown in Figs. 17(b) and 18(b). The fact that the distances between all the states are small and the relatively broad band of dark blue along the diagonal implies that the evolution of this system is much slower and without any abrupt changes. This is profoundly different from the behavior of the $r_p = 0$, $\mu = 0$ system, for the considered range of ρ's.

We give a brief physical interpretation of the results shown in Figs. 17 and 18; more complete and detailed discussing will be given elsewhere. The results shown in these figures suggest that at least for the considered narrow range of ρ's, the evolution of $r_p = 0$, $\mu = 0$ system is abrupt, evolving through large scale rearrangements, while the evolution of $r_p = 0.4$, $\mu = 0.5$ is comparably smooth. This is consistent with our earlier results [18,19], we show that the $r_p = 0$, $\mu = 0$ system tends to form polycrystalline zones, while the $r_p = 0.4$, $\mu = 0.5$ system is much more disordered. The large distances between the states (17, 18) and (70, 71) shown in Figs. 17 and 18 are related to the breaking up of these polycrystalline zones.

We end this section by directly comparing the two systems. The entry $D(i, j)$ of the distance matrices shown in Figs. 17(c) and 18(c) is the d_B and d_{W1} distance, respectively, between the state i of the system $r_p = 0$, $\mu = 0$ and state j of the system $r_p = 0.4$, $\mu = 0.5$. Since the differences between the systems are much larger than the differences within a single system, the two systems can be clearly distinguished.

9. Conclusion

Based on different methods for collecting data we have defined three different chain complexes and used them to construct force networks for particulate systems. Using the force networks we compute persistence diagrams and discuss how one can use persistent homology to extract information about the geometric structure of the force distributions between the particles. We provide both theoretical and numerical arguments to show that the persistence diagrams obtained from interaction force networks are the most robust with respect to the numerical errors in the input data. Using numerical data obtained from discrete element simulations of a system of slowly compressed particles, we show that the persistence diagrams associated to the different force networks can differ significantly. This in turn implies that the geometry of the force distributions observed depends upon the methods by which the system is sampled. We provide some intuition concerning how in general the sampling method affects the geometry. We also demonstrate that using persistent homology of any of the three force networks allows one to draw meaningful distinctions between the properties of the force distributions of different systems.

These results now provide a solid mathematical background for the analysis of the force network in more complex configurations, including systems going through jamming transitions, systems consisting of particles of different shapes, and particulate systems in 3D. Our research in these directions is currently in progress.

References

[1] J. Brujic, S. Edwards, I. Hopkinson, H. Makse, Measuring the distribution of interdroplet forces in a compressed emulsion system, Physica A 327 (2003) 201.
[2] M.E. Cates, J.P. Wittmer, J.-P. Bouchaud, P. Claudin, Jamming, force chains and fragile matter, Phys. Rev. Lett. 81 (1998) 1841.
[3] C. Liu, S.R. Nagel, D.A. Schecter, S.N. Coppersmith, S. Majumdar, O. Narayan, T.A. Witten, Force fluctuations in bead packs, Science 269 (1995) 513.
[4] T.S. Majmudar, R.P. Behringer, Contact force measurements and stress-induced anisotropy in granular materials, Nature 435 (2005) 1079.
[5] S. Alexander, Amorphous solids: their structure, lattice dynamics and elasticity, Phys. Rep. 296 (1998) 65.
[6] F. Radjai, M. Jean, J.J. Moreau, S. Roux, Force distribution in dense two-dimensional granular systems, Phys. Rev. Lett. 77 (1996) 274.
[7] F. Radjai, D.E. Wolf, M. Jean, J.-J. Moreau, Bimodal character of stress transmission in granular packings, Phys. Rev. Lett. 80 (1998) 61.
[8] S. Ostojic, E. Somfai, B. Nienhuis, Scale invariance and universality of force networks in static granular matter, Nature 439 (2006) 828.
[9] B.P. Tighe, J.H. Snoeijer, T.J.H. Vlugt, M. van Hecke, The force network ensemble for granular packings, Soft Matter 6 (2010) 2908–2917.
[10] J. Peters, M. Muthuswamy, J. Wibowo, A. Tordesillas, Characterization of force chains in granular material, Phys. Rev. E 72 (2005) 041307.
[11] A. Tordesillas, D.M. Walker, G. Froyland, J. Zhang, R. Behringer, Transition dynamics and magic-number-like behavior of frictional granular clusters, Phys. Rev. E 86 (2012) 011306.
[12] A. Tordesillas, D.M. Walker, Q. Lin, Force cycles and force chains, Phys. Rev. E 81 (2010) 011302.
[13] D. Bassett, E. Owens, K. Daniels, M. Porter, Influence of network topology on sound propagation in granular materials, Phys. Rev. E 86 (2012) 041306.
[14] M. Herrera, S. McCarthy, S. Sotterbeck, E. Cephas, W. Losert, M. Girvan, Path to fracture in granular flows: dynamics of contact networks, Phys. Rev. E 83 (2011) 061303.
[15] D. Walker, A. Tordesillas, Taxonomy of granular rheology from grain property networks, Phys. Rev. E 85 (2012) 011304.
[16] R. Arévalo, I. Zuriguel, D. Maza, Topology of the force network in jamming transition of an isotropically compressed granular packing, Phys. Rev. E 81 (2010) 041302.
[17] R. Arévalo, L. Pugnaloni, I. Zuriguel, D. Maza, Contact network topology in tapped granular media, Phys. Rev. E 87 (2013) 022203.
[18] L. Kondic, A. Goullet, C. O'Hern, M. Kramar, K. Mischaikow, R. Behringer, Topology of force networks in compressed granular media, Europhys. Lett. 97 (2012) 54001.
[19] M. Kramar, A. Goullet, L. Kondic, K. Mischaikow, Persistence of force networks in compressed granular media, Phys. Rev. E 87 (2013) 042207.
[20] G. Carlsson, Topology and data, Bull. Amer. Math. Soc. (NS) 46 (2009) 255.
[21] H. Edelsbrunner, J.L. Harer, Computational Topology, AMS, Providence, RI, 2010.
[22] S. Ardanza-Trevijano, I. Zuriguel, R. Arevalo, D. Maza, Topological analysis of tapped granular media using persistent homology, Phys. Rev. E 89 (5) (2014) 052212.

[23] R.R. Hartley, R.P. Behringer, Logarithmic rate dependence of force networks in sheared granular materials, Nature 421 (2003) 928.

[24] M. Kramar, "Chomp", 2013. http://Chomp.rutgers.edu/Archives/Topological_Data_Analysis/Granular/QuantifyingForceNetwork/SupplementalMaterials.html.

[25] V. Nanda, Perseus, 2012. http://www.math.rutgers.edu/~vidit/perseus.html.

[26] T. Kaczynski, K. Mischaikow, M. Mrozek, Computational Homology, in: Applied Mathematical Sciences, vol. 157, Springer-Verlag, New York, 2004.

[27] P.A. Cundall, O.D.L. Strack, A discrete numerical model for granular assemblies, Géotechnique 29 (1979) 47.

[28] K. Mischaikow, V. Nanda, Morse theory for filtrations and efficient computation of persistent homology, Discrete Comput. Geom. 50 (2) (2013) 330–353.