TOWARDS AN OPTIMIZATION THEORY FOR DEFORMING DENSE GRANULAR MATERIALS: MINIMUM COST MAXIMUM FLOW SOLUTIONS

Qun Lin

Department of Mathematics and Statistics Curtin University, Perth, Australia 6845, Australia

Antoinette Tordesillas

Department of Mathematics and Statistics University of Melbourne, Melbourne, Australia 3010, Australia

Abstract. We use concepts and techniques of network optimization theory to gain a better understanding of force transmission in dense granular materials. Specifically, we represent a deforming granular material over the different stages of a quasi-static biaxial compression test as a series of representative flow networks, and analyze force transmission through these networks. The forces in such a material are transmitted through the contacts between the constituent grains. As the sample deforms during the various stages of the biaxial test, these grains rearrange: while many contacts are preserved in this rearrangement process, some new contacts form and some old contacts break. We consider the maximum flow problem and the minimum cost maximum flow (MCMF) problem for the flow networks constructed from this evolving network of grain contacts. We identify the flow network bottleneck and establish the sufficient and necessary conditions for a minimum cut of the maximum flow problem to be unique. We also develop an algorithm to determine the MCMF pathway, i.e. a set of edges that always transmit non-zero flow in every solution of the MCMF problem. The bottlenecks of the flow networks develop in the locality of the persistent shear band, an intensively-studied phenomenon that has long been regarded as the signature failure microstructure for dense granular materials. The cooperative evolution of the most important structural building blocks for force transmission, i.e. the force chains and 3-cycles, is examined with respect to the MCMF pathways. We find that the majority of the particles in the major load-bearing columnar force chains and 3-cycles consistently participate in the MCMF pathways.

1. **Introduction.** Flow networks are ubiquitous in everyday life. Typically, these are characterized by some quantity being passed from one node to another in the network: current through electrical networks, product components through assembly lines, energy flow through food webs in an ecosystem, fluid or gas through pipelines, information through communication networks, vehicles through roadways, etc. [9]. A recurring question of great practical interest is: how can we ensure optimal efficiency of flow in these networks subject to certain constraints? In this study, we focus on force transmission through the *contact network*, i.e. the network of

²⁰¹⁰ Mathematics Subject Classification. Primary: 90B10, 90C35, 90C27; Secondary: 76T25. Key words and phrases. Network optimization, maximum flow, minimum cut, minimum cost, granular materials, force chains.

The reviewing process of the paper was handled by Ryan Loxton as Guest Editor.

grain-grain contacts, of a deforming granular material (e.g. soil, rocks, powders, grains, etc.). In the absence of grain breakage, deformation in granular materials is governed by grain rearrangements. As constituent grains or particles rearrange in response to applied stresses and strains, some old contacts break while some new contacts form; in turn, this leads to an evolution of the contact network from one configuration to another, as loading proceeds. We analyze these representative complex contact networks across the consecutive equilibrium states of the loading history to determine the extent to which force transmission is optimized in a granular material under quasi-static loading conditions.

Before proceeding with the analysis, a brief background to this study may be instructive and cast light on why we believe optimization theory has much to offer research into the mechanics and physics of granular media. The mechanical response of a granular material under applied stresses and strains has been the subject of numerous studies throughout history (e.g. [5]). This mechanical response, arising from the interactions of many individual units (i.e. the constituent grains), has not only been a favored paradigm for a complex system, but is also of great significance in many technological and industrial settings (e.g. extraction, processing, transport and handling of minerals, road and off-road transport, geotechnical construction, recycling of cans and papers, and processing of food grains, to name a few examples). Granular materials transmit force through a network of inter-grain contacts. This contact network is dual in nature: one is strong, the other is weak [19, 21]. The strong network is embodied in "force chains": these are self-organized columnlike structures which form in the direction of the maximum or most compressive principal stress, each consisting of at least three particles carrying above the global average load [21, 23]. The particles in the complementary weak network form minimal cycles that surround, and are conjoined with the force chains. The strong and weak networks evolve in a highly cooperative manner [26]. Akin to those seen in architectural structures, the strong columnar force chains carry the majority of the applied load, while the surrounding weak network particles in the minimal cycles provide truss-like lateral supports. In particular, the smallest member of the minimal cycles basis, the 3-cycles, are shown to provide dual support to force chains: (i) by propping-up the force chain to restore its alignment, and (ii) by frustrating relative rotation of particles at contacts. These two mechanisms, through which 3-cycles help resist and delay failure by buckling of force chains and, in turn, global failure of the material — have been systematically studied in both simulations and experiments [1, 23, 25, 26]. In particular, considerable evidence suggests that the failure of granular materials via shear bands is due to the localized buckling of force chains, as originally proposed by Oda and co-workers [4, 14, 15, 16, 17, 18, 20].

Though steady progress has been made in unravelling the co-evolution of these self-organized columnar and truss-like structures, the process by which the material "selects" which particles belong to the strong force chains versus those in the weak network remains unclear [25]. In an effort to unravel the details of this selection process, we have undertaken this study as among the first steps toward understanding force transmission and energy flow from the perspective of combinatorial optimization. We wish to uncover what quantity is being optimized (if any), as the material deforms or as its constituent particles rearrange, a process that is reflected in the evolution of the material's contact network. In particular, are these rearrangements taking place so as to optimize force transmission in the

direction of the major (most compressive) principal stress — the direction in which the self-organized, load-bearing columns of force chains form?

Our strategy involves the construction of an appropriate flow network or directed graph for each equilibrium state of the material, starting from the contact network. Somewhat surprisingly, even though granular materials lend themselves naturally to a network representation, the use of complex networks and related fields (e.g. combinatorial optimization) has only been recently explored in their characterization and modeling (e.g. [23, 24, 26, 27]). Here we wish to continue this line of investigation and use network optimization techniques to establish whether the forces propagate optimally through the contacts of the deforming medium. To the best of our knowledge, this study is among the first attempts to use network optimization principles and techniques to the study of granular materials, in particular, with respect to force transmission [11, 22]. Our analysis proceeds in three steps. First, we construct a flow network at each strain state of the material in static equilibrium: we assign a source, a sink and intermediate nodes in the contact network and assign costs to the edges based on the contact types. Second, we determine the maximum flow that can be transmitted through this resulting flow network, from source to sink, at minimum cost. Third, we relate quantitatively the MCMF solutions to the defining features of force transmission in the deforming material, as expounded above.

The paper is organized as follows. In Section 2, we introduce the basic concepts of flow networks. In Sections 3 and 4, we present the maximum flow problem and the MCMF problem, respectively. In Section 5, we explore the MCMF solutions to examine force transmission in a discrete element simulation, previously shown to exhibit the defining features of a dense granular material under quasi-static biaxial compression. Attention here will be paid to the most important structural building blocks of self-organization, i.e. the force chains and the 3-cycles. We conclude in Section 6.

2. Flow networks. In graph theory, a directed graph (or digraph) consists of a finite set or collection of elements called nodes (or vertices) together with a subset of ordered pairs of the nodes called edges (or arcs). A flow network (or network) is a digraph with two distinguished nodes, called a source and a sink, together with a non-negative real-valued function defined on its edge set, called a capacity function (see [2, 3, 10]).

Let $(\mathcal{N}, \mathcal{A})$ be a finite connected digraph, where \mathcal{N} is the set of nodes such that $s \in \mathcal{N}$ and $t \in \mathcal{N}$ and \mathcal{A} is the set of directed edges. Furthermore, let u_{ij} denote the capacity of the directed edge $(i, j) \in \mathcal{A}$. Then $(\mathcal{N}, \mathcal{A}, u, s, t)$ is called a *flow network* with source s, $sink\ t$ and $capacity\ function\ u$.

An ordered set of values $f \triangleq \{f_{ij} : (i,j) \in \mathcal{A}\}$ is a feasible flow vector on $(\mathcal{N}, \mathcal{A}, u, s, t)$ if it satisfies the following constraints:

$$0 \le f_{ij} \le u_{ij}, \qquad (i,j) \in \mathcal{A}, \tag{1}$$

and

$$\sum_{j:(i,j)\in\mathcal{A}} f_{ij} - \sum_{j:(j,i)\in\mathcal{A}} f_{ji} = 0, \qquad i \in \mathcal{N} \setminus \{s,t\}.$$
 (2)

The capacity constraint (1) requires that each edge carries a non-negative amount of flow which cannot exceed the capacity of the edge, and the conservation constraint (2) means that flows are preserved: at each node, except for the source and

the sink, the amount that flows in also flows out. Let \mathcal{F} denote the set of all feasible flow vectors on $(\mathcal{N}, \mathcal{A}, u, s, t)$. For each $f \in \mathcal{F}$, define

$$val(f) \triangleq \sum_{j: (s,j) \in \mathcal{A}} f_{sj} - \sum_{j: (j,s) \in \mathcal{A}} f_{js}.$$

Then val(f) is the amount of flow that is transmitted through the network from source to sink.

For any two subsets \mathcal{X}_1 and \mathcal{X}_2 of \mathcal{N} , let

$$(\mathcal{X}_1, \mathcal{X}_2) \triangleq \{ (i, j) \in \mathcal{A} : i \in \mathcal{X}_1, j \in \mathcal{X}_2 \}.$$

Then $(\mathcal{X}_1, \mathcal{X}_2)$ is the set of edges from \mathcal{X}_1 to \mathcal{X}_2 . For any $\mathcal{X} \subseteq \mathcal{N}$, let $\overline{\mathcal{X}} \triangleq \mathcal{N} \setminus \mathcal{X}$. A $cut(\mathcal{X}, \overline{\mathcal{X}})$ of $(\mathcal{N}, \mathcal{A}, u, s, t)$ is a partition of \mathcal{N} into two subsets with $s \in \mathcal{X}$ and $t \in \overline{\mathcal{X}}$. The capacity of $(\mathcal{X}, \overline{\mathcal{X}})$ is defined as

$$U(\mathcal{X}, \overline{\mathcal{X}}) \triangleq \sum_{(i,j)\in(\mathcal{X}, \overline{\mathcal{X}})} u_{ij}.$$
 (3)

3. **The maximum flow problem.** For the flow network formulated above, one can solve the maximum flow problem which we define as follows.

Problem 1. Find a feasible flow vector $f^* \in \mathcal{F}$ such that

$$\operatorname{val}(f^*) = \max_{f \in \mathcal{F}} \operatorname{val}(f).$$

Problem 1 can be solved by the Ford-Fulkerson method [2]. Let

$$F_{\max} \triangleq \max_{f \in \mathcal{F}} \operatorname{val}(f).$$

Note that the solution to Problem 1 may not be unique. Let \mathcal{M} denote the set of solutions, namely,

$$\mathcal{M} \triangleq \{ f \in \mathcal{F} : \operatorname{val}(f) = F_{\max} \}.$$

3.1. **Minimum cut.** A dual of the maximum flow problem is the minimum cut. In a flow network, a cut $(\mathcal{X}, \overline{\mathcal{X}})$ is called a *minimum cut* if its capacity defined by (3) is minimal. The maximum flow minimum cut theorem shows that the maximal value of a flow in a flow network is equal to the capacity of the minimum cut of the flow network (e.g. [3]). Thus, the maximal value of a flow is limited by the capacity of the minimum cut.

Let $f \in \mathcal{F}$ and $(\mathcal{X}, \overline{\mathcal{X}})$ be a cut of $(\mathcal{N}, \mathcal{A}, u, s, t)$. According to [3], we have the following results:

(I) $val(f) \leq U(\mathcal{X}, \overline{\mathcal{X}})$, where equality holds if and only if

$$f_{ij} = \begin{cases} u_{ij}, & (i,j) \in (\mathcal{X}, \overline{\mathcal{X}}), \\ 0, & (i,j) \in (\overline{\mathcal{X}}, \mathcal{X}). \end{cases}$$
(4)

- (II) If $\operatorname{val}(f) = U(\mathcal{X}, \overline{\mathcal{X}})$, then $f \in \mathcal{M}$ and $\operatorname{cut}(\mathcal{X}, \overline{\mathcal{X}})$ is minimal.
- (III) If $f \in \mathcal{M}$ and $(\mathcal{X}, \overline{\mathcal{X}})$ is minimal, then $val(f) = U(\mathcal{X}, \overline{\mathcal{X}})$.

Note that (I), (II) and (III) are Theorem 7.3, Corollary 7.4 and Theorem 7.7 of [3], respectively. Using these results, we can prove the following lemma.

Lemma 3.1. Let $(\mathcal{X}, \overline{\mathcal{X}})$ be a cut of $(\mathcal{N}, \mathcal{A}, u, s, t)$. If there exists $f \in \mathcal{M}$ satisfying (4), then $(\mathcal{X}, \overline{\mathcal{X}})$ is a minimum cut. Conversely, if cut $(\mathcal{X}, \overline{\mathcal{X}})$ is minimal, then (4) holds for all $f \in \mathcal{M}$.

Proof. Suppose that $f \in \mathcal{M}$ satisfies equation (4). Then, by (I),

$$U(\mathcal{X}, \overline{\mathcal{X}}) = \operatorname{val}(f).$$

It thus follows from (II) that cut $(\mathcal{X}, \overline{\mathcal{X}})$ is minimal.

Next, we will show that if $(\mathcal{X}, \overline{\mathcal{X}})$ is a minimum cut, equation (4) holds for all $f \in \mathcal{M}$. By (III), we know that for each $f \in \mathcal{M}$,

$$U(\mathcal{X}, \overline{\mathcal{X}}) = \operatorname{val}(f).$$

Suppose there exists an $f \in \mathcal{M}$ such that (4) does not hold. It thus follows from (I) that $\operatorname{val}(f) < U(\mathcal{X}, \overline{\mathcal{X}})$. This contradicts the above equality. Therefore, (4) holds for all $f \in \mathcal{M}$.

Note that the minimum cut of $(\mathcal{N}, \mathcal{A}, u, s, t)$ may or may not be unique. What are the sufficient and necessary conditions for the minimum cut to be unique? Next, we will establish the sufficient and necessary conditions.

An (i-j) path \mathcal{P} is a path that links node i and node j. Let \mathcal{P}^+ and \mathcal{P}^- denote the sets of forward edges and backward edges on \mathcal{P} , respectively.

For $f \in \mathcal{F}$ and a path \mathcal{P} , define

$$\delta(f, \mathcal{P}) \triangleq \min \left\{ \min_{(i,j) \in \mathcal{P}^+} \{u_{ij} - f_{ij}\}, \min_{(i,j) \in \mathcal{P}^-} f_{ij} \right\}.$$

The path \mathcal{P} is f-saturated if $\delta(f,\mathcal{P}) = 0$ and f-unsaturated if $\delta(f,\mathcal{P}) > 0$. Note that for an (s-t) path \mathcal{P} , $\delta(f,\mathcal{P})$ is the largest amount of flow by which the flow f can be increased along \mathcal{P} without violating constraints (1)-(2). Thus, $f \in \mathcal{M}$ if and only if there are no f-unsaturated (s-t) paths.

Given $f \in \mathcal{F}$, let \mathcal{S}_f be the set of all nodes reachable from s by f-unsaturated paths. Furthermore, let \mathcal{T}_f be the set of all nodes from which t can be reached by f-unsaturated paths. Clearly, $s \in \mathcal{S}_f$ and $t \in \mathcal{T}_f$. We have the following result.

Theorem 3.2. For each $f \in \mathcal{M}$, $(\mathcal{S}_f, \overline{\mathcal{S}}_f)$ and $(\overline{\mathcal{T}}_f, \mathcal{T}_f)$ are minimum cuts of $(\mathcal{N}, \mathcal{A}, u, s, t)$.

Proof. Clearly, (S_f, \overline{S}_f) is a cut. Otherwise, there exists an f-unsaturated path \mathcal{P} from s to t. This implies that f is not a maximum flow, which contradicts the condition $f \in \mathcal{M}$.

We now prove that (S_f, \overline{S}_f) is a minimum cut. Suppose, to the contrary, that (S_f, \overline{S}_f) is not minimal. It thus follows from Lemma 3.1 that there exists $i \in S_f$ and $j \in \overline{S}_f$ such that

$$(i,j) \in (S_f, \overline{S}_f)$$
 with $f_{ij} < u_{ij}$,

or

$$(j,i) \in (\overline{\mathcal{S}}_f, \mathcal{S}_f)$$
 with $f_{ji} > 0$.

Since $i \in \mathcal{S}_f$, there exists an f-unsaturated (s-i) path. Combining this with the above inequalities, there exists an f-unsaturated (s-j) path. Thus, $j \in \mathcal{S}_f$, which contradicts the initial assumption that $j \notin \mathcal{S}_f$. Therefore, $(\mathcal{S}_f, \overline{\mathcal{S}}_f)$ is minimal.

Similarly, we can show that cut $(\overline{\mathcal{T}}_f, \mathcal{T}_f)$ is also minimal.

According to Theorem 3.2, both (S_f, \overline{S}_f) and (\overline{T}_f, T_f) are minimal. Moreover, applying Lemma 3.1, we can show that for each minimum cut $(\mathcal{X}, \overline{\mathcal{X}})$, S_f and \overline{T}_f are the lower and upper bounds on \mathcal{X} , respectively.

Theorem 3.3. Let $f \in \mathcal{M}$ and let $(\mathcal{X}, \overline{\mathcal{X}})$ be a minimum cut of $(\mathcal{N}, \mathcal{A}, u, s, t)$. Then $S_f \subseteq \mathcal{X} \subseteq \overline{\mathcal{T}}_f$.

Proof. Using proof by contradiction, we can show that $S_f \subseteq \mathcal{X}$. Indeed, if $S_f \nsubseteq \mathcal{X}$, then there exists a $k \in S_f$ such that $k \notin \mathcal{X}$. Since $k \in S_f$, there exists an f-unsaturated path \mathcal{P} from s to k. Note that $s \in \mathcal{X}$ and $k \in \overline{\mathcal{X}}$. There must be an edge (i, j) on \mathcal{P} such that

$$(i,j) \in (\mathcal{X}, \overline{\mathcal{X}})$$
 with $f_{ij} < u_{ij}$,

or

$$(i,j) \in (\overline{\mathcal{X}}, \mathcal{X})$$
 with $f_{ij} > 0$.

Recall that f is a maximum flow. It thus follows from Lemma 3.1 that cut $(\mathcal{X}, \overline{\mathcal{X}})$ is not minimal. This leads to a contradiction, since $(\mathcal{X}, \overline{\mathcal{X}})$ is minimal.

In a similar manner, we can prove $\mathcal{X} \subseteq \overline{\mathcal{T}}_f$. Hence, $\mathcal{S}_f \subseteq \mathcal{X} \subseteq \overline{\mathcal{T}}_f$, as required.

Combining Theorems 3.2 and 3.3, the minimum cut is unique if and only if there exists a maximum flow f such that $S_f = \overline{\mathcal{T}_f}$.

3.2. The flow network bottleneck. Perhaps the most intuitive application of minimum cuts is in the study of fluid flow through a complex network of pipelines through which some fluid flows. Pumping the fluid through the pipes cannot exceed some maximum flow because of the bottleneck, i.e. a subset of pipes that transfer the fluid at their maximum capacity. In a flow network, if the minimum cut $(\mathcal{X}, \overline{\mathcal{X}})$ is unique, this bottleneck will be the minimum cut, i.e. the set of edges that block the flow. Thus, $(\mathcal{X}, \overline{\mathcal{X}})$ was defined as the bottleneck of the flow network in [22]. However, a minimum cut of a general flow network may or may not be unique and the bottleneck of the flow network may still exist, regardless of the uniqueness of minimum cut. In this subsection, we will define the bottleneck for a general flow network, including the one in [22] as a special case.

Let \mathcal{B} be the intersection of all minimum cuts, namely,

$$\mathcal{B} \triangleq \bigcap_{(\mathcal{X}, \overline{\mathcal{X}}) \text{ is a minimum cut}} (\mathcal{X}, \overline{\mathcal{X}}).$$

Then any sufficiently small change in the capacity of \mathcal{B} , i.e.

$$U(\mathcal{B}) = \sum_{(i,j)\in\mathcal{B}} u_{ij},$$

will change the value of maximum flow. In other words, there exists an $\bar{\epsilon} > 0$ such that for any $|\epsilon| < \bar{\epsilon}$, if the new capacity of \mathcal{B} becomes $U(\mathcal{B}) + \epsilon$ and the capacities on all edges of $\mathcal{A} \setminus \mathcal{B}$ are unchanged, then the value of maximum flow of the flow network with the new capacity will be $\operatorname{val}(f) + \epsilon$. In this sense, we regard \mathcal{B} as the bottleneck of the flow network $(\mathcal{N}, \mathcal{A}, u, s, t)$. Furthermore, we have the following result.

Theorem 3.4. $\mathcal{B} = (\mathcal{S}_f, \mathcal{T}_f)$, where f is a maximum flow of $(\mathcal{N}, \mathcal{A}, u, s, t)$.

Proof. Note from Theorem 3.3 that $S_f \subseteq \overline{\mathcal{T}_f}$ and $\mathcal{T}_f \subseteq \overline{\mathcal{S}_f}$. Then it is easy to see that

$$(\mathcal{S}_f, \overline{\mathcal{S}_f}) \cap (\overline{\mathcal{T}_f}, \mathcal{T}_f) = (\mathcal{S}_f \cap \overline{\mathcal{T}_f}, \overline{\mathcal{S}_f} \cap \mathcal{T}_f) = (\mathcal{S}_f, \mathcal{T}_f),$$

from which it follows that

$$\mathcal{B} \subseteq (\mathcal{S}_f, \overline{\mathcal{S}_f}) \cap (\overline{\mathcal{T}_f}, \mathcal{T}_f) = (\mathcal{S}_f, \mathcal{T}_f). \tag{5}$$

Now, let minimum cut $(\mathcal{X}, \overline{\mathcal{X}})$ be arbitrary but fixed. Applying Theorem 3.3 once again, $S_f \subseteq \mathcal{X}$ and $\mathcal{T}_f \subseteq \overline{\mathcal{X}}$. Hence,

$$(\mathcal{S}_f, \mathcal{T}_f) \subseteq (\mathcal{X}, \overline{\mathcal{X}}).$$

Since $(\mathcal{X}, \overline{\mathcal{X}})$ was chosen arbitrarily, we obtain

$$(\mathcal{S}_f, \mathcal{T}_f) \subseteq \mathcal{B}. \tag{6}$$

Combining (5) and (6) gives $\mathcal{B} = (\mathcal{S}_f, \mathcal{T}_f)$. This completes the proof.

Suppose that the minimum cut is unique, i.e. $S_f = \overline{\mathcal{T}_f}$. Then

$$\mathcal{B} = (\mathcal{S}_f, \overline{\mathcal{S}_f}) = (\overline{\mathcal{T}_f}, \mathcal{T}_f).$$

This is the bottleneck of the flow network defined in [22].

4. The minimum cost maximum flow problem. In most real world networks, there is often a cost associated with the flow. Thus one of the most fundamental problems in flow networks concerns finding the cheapest route of sending the maximum flow through the network [6, 7]. In a deforming granular material, the transmission of force at grain contacts invariably involves some energy dissipation that may be envisaged to be an associated cost. Thus, a natural question to ask is: where are the pathways through the contact network along which force transmission is maximized while the associated energy dissipation is minimized?

For each feasible flow vector $f \in \mathcal{F}$, define the associated cost function as follows:

$$E(f) \triangleq \sum_{(i,j)\in\mathcal{A}} c_{ij} f_{ij},$$

where c_{ij} denotes the cost per unit of flow through edge (i, j).

We are interested in the flow that can transmit the maximal value at the minimum cost. This leads to the following minimum cost maximum flow (MCMF) problem.

Problem 2. Find a feasible flow vector $f^* \in \mathcal{M}$ such that

$$E(f^*) = \min_{f \in \mathcal{M}} E(f).$$

Note that every solution of Problem 1 is feasible for Problem 2 and that Problem 2 involves finding a solution of Problem 1 of minimum cost. We can first obtain a solution of Problem 1 using the Ford-Fulkerson method [2] and then use this solution as an initial guess and apply the network simplex method [2] to solve Problem 2. Alternatively, Problems 1 and 2 together can be solved using the algorithm proposed by Edmonds and Karp [6]. For convenience of the reader, we present the algorithm below. Note that Steps 2-7 are used to find an unsaturated path from s to t of minimum cost.

Algorithm 4.1. Input the flow network $(\mathcal{N}, \mathcal{A}, u, s, t)$ and cost function c.

- 1. Set f = 0.
- 2. Set $p_{\iota} = 0$, $\iota \in \mathcal{N}$, $\epsilon_{\iota} = +\infty$, $\iota \in \mathcal{N}$ and

$$\sigma_{\iota} = \begin{cases} 0, & \iota = s, \\ +\infty, & \iota \in \mathcal{N} \setminus \{s\}. \end{cases}$$

3. Set $S = \{s\}$.

- 4. If $S = \emptyset$, go to Step 8. Otherwise, choose $\kappa \in S$ such that $\sigma_{\kappa} = \min_{n \in S} \sigma_{\ell}$.
- 5. For each $i \in \{\iota \in \mathcal{N} : (\kappa, \iota) \in \mathcal{A} \text{ and } f_{\kappa\iota} < u_{\kappa\iota}\}, \text{ if } \sigma_i > \sigma_{\kappa} + c_{\kappa i}, \text{ let } \mathcal{S} = \mathcal{S} \cup \{i\},$

$$p_i = \kappa$$
, $\sigma_i = \sigma_{\kappa} + c_{\kappa i}$ and $\epsilon_i = \min\{u_{\kappa i} - f_{\kappa i}, \epsilon_{\kappa}\}.$

- 6. For each $i \in \{ \iota \in \mathcal{N} : (\iota, \kappa) \in \mathcal{A} \text{ and } f_{\iota\kappa} > 0 \}$, if $\sigma_i > \sigma_{\kappa} c_{i\kappa}$, let $\mathcal{S} = \mathcal{S} \cup \{i\}$, $p_i = -\kappa$, $\sigma_i = \sigma_{\kappa} c_{i\kappa}$ and $\epsilon_i = \min\{f_{i\kappa}, \epsilon_{\kappa}\}$.
- 7. Set $S = S \setminus \{\kappa\}$ and go to Step 4.
- 8. If $\epsilon_t < +\infty$, let \mathcal{P} be the corresponding unsaturated path from s to t at minimum cost (\mathcal{P} can be obtained from p_t , $t \in \mathcal{N}$). Set

$$f_{ij} = \begin{cases} f_{ij} + \epsilon_t, & (i,j) \in \mathcal{P}^+, \\ f_{ij} - \epsilon_t, & (i,j) \in \mathcal{P}^-, \\ f_{ij}, & (i,j) \in \mathcal{A} \setminus \mathcal{P}, \end{cases}$$

and then go to Step 2. Otherwise, STOP and f is an optimal solution of Problem 2 and $F_{\text{max}} = \text{val}(f)$.

Note that, like Problem 1, the solution of Problem 2 may also not be unique. Let \mathcal{C} denote the set of solutions of Problem 2. Then,

$$E(f) \le E(g), \qquad f \in \mathcal{C}, \quad g \in \mathcal{M}.$$
 (7)

For each $f \in \mathcal{C}$, define

$$\mathcal{D}_f \triangleq \{ (i,j) \in \mathcal{A} : f_{ij} > 0 \}.$$

Note that \mathcal{D}_f is the set of edges that are used to transmit non-zero flow for flow vector f. Note also that any edge from $\mathcal{A} \setminus \mathcal{D}_f$ is removable. In other words, removing edges from $\mathcal{A} \setminus \mathcal{D}_f$ will not increase the minimum cost or decrease the value of maximum flow. Since a solution to Problem 2 may not be unique, there is no point to investigate \mathcal{D}_f for one particular solution f. Let \mathcal{R} be the set of edges that always transmit non-zero flows for every solution of Problem 2, namely,

$$\mathcal{R} \triangleq \bigcap_{f \in \mathcal{C}} \mathcal{D}_f.$$

We call \mathcal{R} the minimum cost maximum flow pathway (or MCMF pathway). Clearly, MCMF pathway is a subset of edges that are used to transmit non-zero flow for a solution of Problem 2. It is easy to see that $\mathcal{R} \subseteq \mathcal{D}_f$ if $f \in \mathcal{C}$ and $\mathcal{R} = \mathcal{D}_f$ if Problem 2 has a unique solution. Furthermore, it follows from Lemma 3.1 that $(\mathcal{S}_f, \overline{\mathcal{S}_f}) \cup (\overline{\mathcal{T}_f}, \mathcal{T}_f) \subseteq \mathcal{R}$. In particular, $\mathcal{B} \subseteq \mathcal{R}$. This means that the bottleneck of the flow network always transmits non-zero flow for maximum flow at the minimum cost.

For the remainder of this section, let $f \in \mathcal{C}$ be an optimal solution of Problem 2 and let $(\varsigma, \kappa) \in \mathcal{D}_f$ be arbitrary but fixed. Note that $(\varsigma, \kappa) \notin \mathcal{R}$ if and only if there exists a $g \in \mathcal{M}$ such that $g_{\varsigma\kappa} = 0$ and E(g) = E(f). Also note that removing any edge from MCMF pathway will increase the associated cost or decrease the value of maximum flow from source to sink. This is stated formally in the following result.

Theorem 4.1. Let $\mathcal{G} = \{g \in \mathcal{F} : g_{\varsigma\kappa} = 0\}$ and let $g^* \in \mathcal{F}$ be a solution of the following MCMF problem:

$$\min \left\{ E(g) : g \in \mathcal{G} \text{ and } \operatorname{val}(g) = \max_{g' \in \mathcal{G}} \operatorname{val}(g') \right\}.$$

Then $(\varsigma, \kappa) \in \mathcal{R}$ if and only if one of the following conditions holds:

- (a) $val(g^*) < val(f)$;
- (b) $val(g^*) = val(f) \text{ and } E(g^*) > E(f).$

Proof. Note that $\operatorname{val}(g^*) \leq \operatorname{val}(f)$ and that if $\operatorname{val}(g^*) = \operatorname{val}(f)$, then $E(g^*) \geq E(f)$. Hence, it is equivalent to prove that $(\varsigma, \kappa) \notin \mathcal{R}$ if and only if $\operatorname{val}(g^*) = \operatorname{val}(f)$ and $E(g^*) = E(f)$. Clearly, if $\operatorname{val}(g^*) = \operatorname{val}(f)$ and $E(g^*) = E(f)$, then $g^* \in \mathcal{C}$ and $g^*_{\varsigma\kappa} = 0$. Thus, $(\varsigma, \kappa) \notin \mathcal{R}$.

We now suppose that $(\varsigma, \kappa) \notin \mathcal{R}$. Then there exists $f' \in \mathcal{C}$ such that $f'_{\varsigma\kappa} = 0$. It thus follows that $f' \in \mathcal{G}$, E(f') = E(f), and val(f') = val(f). Consequently, we have

$$\operatorname{val}(f) = \operatorname{val}(f') \le \operatorname{val}(g^*) = \max_{g \in \mathcal{G}} \operatorname{val}(g) \le \max_{g \in \mathcal{F}} \operatorname{val}(g) = \operatorname{val}(f).$$

Hence, Problem 2 can be rewritten as follows:

$$\min \{ E(g) : g \in \mathcal{F} \text{ and } val(g) = val(f) \}.$$

Recall that f is a solution of Problem 2. Since $f' \in \mathcal{G}$,

$$E(f) = \min \{ E(g) : g \in \mathcal{F} \text{ and } val(g) = val(f) \}$$

$$\leq \min \{ E(g) : g \in \mathcal{G} \text{ and } val(g) = val(f) \}$$

$$= E(g^*) \leq E(f') = E(f).$$

Hence, we obtain $\operatorname{val}(g^*) = \operatorname{val}(f)$ and $E(g^*) = E(f)$. The proof is complete. \square

We can apply Algorithm 4.1 to solve the MCMF problem in Theorem 4.1 by replacing u with v, where

$$v_{ij} = \begin{cases} 0, & (i,j) = (\varsigma, \kappa), \\ u_{ij}, & (i,j) \in \mathcal{A} \setminus \{(\varsigma, \kappa)\}. \end{cases}$$

Note that for a large digraph, $|\mathcal{D}_f \setminus ((\mathcal{S}_f, \overline{\mathcal{S}_f}) \cup (\overline{\mathcal{T}_f}, \mathcal{T}_f))|$ is quite large and it takes a long time to solve the MCMF problem for all edges in $\mathcal{D}_f \setminus ((\mathcal{S}_f, \overline{\mathcal{S}_f}) \cup (\overline{\mathcal{T}_f}, \mathcal{T}_f))$. Hence, we should avoid applying Algorithm 4.1 if we can check multiple edges at the same time. To this end, we need the following result.

Theorem 4.2. Let \mathcal{P} and \mathcal{Q} be directed paths from i to j such that

$$\sum_{(i',j')\in\mathcal{P}} c_{i'j'} = \sum_{(i',j')\in\mathcal{Q}} c_{i'j'}$$

and

$$0 < f_{\iota\tau} = \min \{ f_{i'j'} : (i',j') \in \mathcal{P} \} \le \min \{ u_{i'j'} - f_{i'j'} : (i',j') \in \mathcal{Q} \}.$$

Then $(\iota, \tau) \notin \mathcal{R}$.

Proof. Let

$$f'_{i'j'} = \begin{cases} f_{i'j'} - f_{\iota\tau}, & (i',j') \in \mathcal{P}, \\ f_{i'j'} + f_{\iota,\tau}, & (i',j') \in \mathcal{Q}, \\ f_{i'j'}, & (i',j') \in \mathcal{A} \setminus (\mathcal{P} \cup \mathcal{Q}). \end{cases}$$
(8)

Then, it is easy to check that $f' \in \mathcal{C}$. Moreover, we have $f'_{\iota\tau} = 0$ and hence $(\iota, \tau) \notin \mathcal{R}$.

Using the shortest path algorithm, we can obtain \mathcal{P} and \mathcal{Q} satisfies conditions given in Theorem 4.2.

Let $i, j \in \mathcal{N}$ be arbitrary but fixed. Furthermore, let \mathcal{P}_{ij} denote the shortest directed path from i to j in digraph $(\mathcal{N}, \mathcal{D}_f)$, where the distance is measured by the cost function c, i.e.

$$C(\mathcal{P}_{ij}) = \sum_{(i',j') \in \mathcal{P}_{ij}} c_{i'j'}.$$

Similarly, let Q_{ij} denote the shortest directed path from i to j in digraph $(\mathcal{N}, \mathcal{A}_f)$, where

$$A_f \triangleq \{ (i', j') \in A : u_{i'j'} - f_{i'j'} > 0 \}.$$

Clearly, $C(\mathcal{P}_{ij}) \leq C(\mathcal{Q}_{ij})$. Otherwise, f' defined in (8) is feasible for Problem 2 but with less cost. This contradicts (7). Hence, for any $i, j \in \mathcal{N}$, if $C(\mathcal{P}_{ij}) = C(\mathcal{Q}_{ij})$ and

$$\min \{ f_{i'j'} : (i',j') \in \mathcal{P}_{ij} \} \leq \min \{ u_{i'j'} - f_{i'j'} : (i',j') \in \mathcal{Q}_{ij} \},$$

then $(\iota,\tau) \notin \mathcal{R}$ whenever $(\iota,\tau) \in \mathcal{P}_{ij}$ satisfies $f_{\iota\tau} = \min \{ f_{i'j'} : (i',j') \in \mathcal{P}_{ij} \}.$

5. Results. In order to demonstrate the efficacy of the network optimization analysis for characterizing force transmission in a dense granular material, we use the data from a discrete element simulation (DEM) developed and described in detail elsewhere [21] and further analyzed with respect to other simulation and experimental tests in [23, 26]. In this simulation, a densely packed polydisperse assembly of 5098 spherical particles are constrained to move along a plane throughout the loading history. The assembly is compressed quasi-statically, under a constant strain rate in the vertical direction and allowed to expand under a constant confining pressure in the horizontal direction. Under this loading condition, the evolution of the material can be characterized over a sequence of states, each one of which is in static equilibrium. As loading proceeds, constituent particles rearrange (i.e. some old contacts break as new contacts form), resulting in a contact network that evolves from state to state. Space limitations prevent a full coverage of this simulation here (see [21, 23] for additional details); hence, only the most relevant aspects are recapitulated below.

The resistance to the relative motion at the particle-particle and particle-wall contacts is governed by combinations of a linear spring, a dashpot and a friction slider, and is designed to mimic the response of assemblies of noncircular particles. Between contacting particles, the resistive force is defined as follows:

$$f_n \triangleq k^n \Delta u_n + b^n \Delta v_n, \tag{9}$$

$$f_t \triangleq \begin{cases} k^t \Delta u_t + b^t \Delta v_t, & \text{if } k^t |\Delta u_t| < \mu^t |f_n|, \\ \operatorname{sign}(\Delta u_t) \mu^t |f_n|, & \text{if } k^t |\Delta u_t| \ge \mu^t |f_n|, \end{cases}$$
(10a)

where f_n and f_t are the normal and tangential components of contact force; k^n and k^t are the spring stiffness coefficients; b^n and b^t are the viscous damping coefficients; and μ^t is the Coulomb friction coefficient. Similarly, the rolling resistance or contact moment I is expressed as

$$I \triangleq \begin{cases} k^r \Delta \alpha + b^r \Delta \dot{\alpha}, & \text{if } k^r |\Delta \alpha| < \mu^r R_{\min} |f_n|, \\ \operatorname{sign}(\Delta \alpha) \mu^r R_{\min} |f_n|, & \text{if } k^r |\Delta \alpha| \ge \mu^r R_{\min} |f_n|, \end{cases}$$
(11a)

where R_{\min} denotes the smaller radius of the two contacting particles, k^r and b^r are the spring stiffness and viscous damping coefficients, respectively; and μ^r is the

Value Parameter $-8 \times 10^{-3}/s$ Applied strain rate $\dot{\varepsilon}_{yy}$ $7.035 \times 10^{2} N/m$ Confining pressure σ_{xx} $6.81 \times 10^{-7} s$ Timestep increment Initial height: width ratio 1:1 5098 Number of particles $2.65 \times 10^3 kg/m^3$ Particle density $0.76 \times 10^{-3} m$ Smallest radius $1.52 \times 10^{-3} \ m$ Largest radius $1.14 \times 10^{-3} \ m$ Average radius (uniform distribution) 0.858Initial packing density Interparticle friction μ 0.7Particle-wall friction μ (top, bottom) 0.7

0.0

0.02

 $1.05 \times 10^{5} N/m$

 $5.25 \times 10^4 N/m$

 $6.835 \times 10^{-2} \ Nm/rad$

Particle-wall friction μ (sides)

Normal spring stiffness k^n

Tangential spring stiffness k^t

Rotational spring stiffness k^r

Rolling friction μ^r

Table 1. DEM parameters and material properties used.

friction coefficient. The remaining quantities in (9)-(11) are: the relative normal and tangential displacements and relative rotation denoted, respectively, by Δu_n , Δu_t and $\Delta \alpha$, and the relative normal and tangential translational and rotational velocities denoted, respectively, by Δv_n , Δv_t and $\Delta \dot{\alpha}$.

A summary of the simulation and material parameters used in the model is given in Table 1. The vertical walls are frictionless, so that particles can slide and roll along them without any resistance; all other material properties are identical to those of the particles. The top and bottom walls are assumed to have the same material properties as the particles.

Depending on the magnitude of the contact forces and moments, four contact types can be identified:

- (1) full-stick contact is the contact in which both the tangential force and contact moment are elastic, i.e. (10a) and (11a) are satisfied;
- (2) sliding contact is the contact in which only the tangential force is at the Coulomb plastic threshold, i.e. (10b) and (11a) are satisfied;
- (3) rolling contact is the contact in which only the contact moment is at the Coulomb plastic threshold, i.e. (10a) and (11b) are satisfied;
- (4) slide-and-roll contact is the contact in which both the tangential force and the contact moment are at their respective Coulomb plastic thresholds, i.e. (10b) and (11b) are satisfied.

Figure 1 presents two key properties of the granular material with respect to the axial strain, the time-like quantity in this simulation. These properties are the stress ratio and the global average of the local non-affine strain, both of which display evolutionary trends which are representative of the behavior of dense granular materials under this loading condition. With respect to the stress ratio, a state quantity that is indicative of the load-carrying capacity of the sample, we observe

three distinct regimes. The first is known as the strain-hardening regime: an initial rise to the peak value, during which the sample is globally stable. This is then followed by the strain-softening regime: a brief period when instabilities inside the material intensify leading to a decrease in load-carrying capacity of the sample as indicated by the decrease in stress ratio under increasing strain. Here a single shear band can be observed to form along the forward diagonal of the sample. Shear bands are regarded as the signature failure microstructure of granular materials and for this reason have attracted significant attention in the literature (see [21] and references cited therein). Shear bands exhibit a clearly defined characteristic thickness of around 8-10 particle diameters [13, 21], which is of the order observed for the shear band thickness in this simulation. This band is a region of localized deformation which, when fully developed, splits the material into distinct parts that can then move relatively to each other, causing the material to lose its load-carrying capacity. In the third and final regime, the so-called critical state regime, the stress ratio undergoes marked fluctuations about a near constant value, reflecting the dynamics occurring inside the now fully formed shear band.

Below the stress ratio in Figure 1 is the global average of the local non-affine strain. The local non-affine strain is a particle property that is computed over a strain interval as opposed to, for example, the stress ratio, which is a sample property computed at each strain state. Specifically, the local non-affine strain is a measure of the deviation from a uniform or affine strain and is a function of the relative motions within a small particle cluster comprising the particle in question and its first ring of neighbors [21]. This property is computed for every particle away from the walls to exclude boundary effects, and then averaged over all these particles: in Figure 1 the global average is assigned to the final state of a strain interval comprising two consecutive states of a total of 299 states in the simulation.

The non-affine strain has been proven to correlate strongly with energy dissipation and be a robust local measure of the evolving internal instabilities and failure developing in the mesoscopic domain of the material [21]. This measure reveals three distinct stages of deformation, as indicated by the black dashed lines in Figure 1. In each of these stages, distinct mesoscopic structures can be observed to form inside the deforming material. In Stage 1, the assembly deforms almost uniformly (i.e. affinely) with respect to the macroscopic scale or sample domain, as dictated by the motion of the sample boundaries. Here the only instability that can be observed inside the sample are rattlers, a particle with one or no contacts, "rattling" inside a cage of neighbors, as shown in Figure 2(a). In Stage 2, an increase in the global average non-affine strain can be observed albeit this remains small prior to peak stress ratio. Inside the sample, however, we see the emergence of a secondary displacement field that is characterized by the presence of intermittent instabilities known as microbands, in addition to rattlers. These microbands organize themselves into thin obliquely trending bands, between one and four particle diameters in width, in which particles undergo large relative tangential motion or slip, as shown in Figure 2(b). The local non-affine strain in microbands progressively intensify, before localizing along the developing shear band in the strain-softening regime [21]. Returning now to the macroscopic scale, we see in Figure 1 that the strain evolution of the global average non-affine strain after peak stress ratio is in marked contrast to that seen prior to peak. Post peak, we observe sudden bursts or sharp peaks in the non-affine strain coinciding with the precipitous drops in stress ratio. The failure of force chains by buckling — localized in the developing shear band —

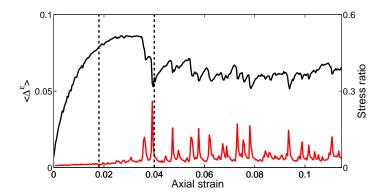


FIGURE 1. Strain evolution of the load-carrying capacity of the sample as measured by the stress ratio (black) and the energy dissipation as measured by the global average of the local non-affine strain $\langle \Delta^{\varepsilon} \rangle$ (red).

has been shown to be the mechanism behind these peaks in non-affine deformation and, in turn, the large energy dissipation and loss of load-carrying capacity of the sample [21]. This trend continues to dominate into Stage 3, the fully developed failure regime or the critical state, during which the sample deforms in the presence of a single fully formed *persistent shear band*, inclined at approximately 45° to the horizontal (i.e. normal to the applied compression), as shown in Figure 2(c).

Having demonstrated the different stages of the loading history from past measures, we now solve the MCMF problem for flow networks constructed for a total of 299 strain states spanning the three stages of the full loading history of the sample. Of interest is whether the solutions to the MCMF problem can similarly capture these three distinct stages of deformation — in addition to uncovering any evidence that may suggest force transmission in the deforming material is optimized.

5.1. Construction of the flow networks. We are interested in force transmission in the direction of the major principal stress along which the primary load-bearing pathways known as force chains develop [21, 23, 26]. We show an example of these self-organized structures taken at peak stress ratio in Figure 3 (recall the behavior of the sample, as earlier discussed in relation to Figure 1). As force chains align themselves in the direction of the maximum (most compressive) principal stress, we will focus on those paths through the contact network that optimize force transmission along this direction, here the vertical. Thus, we choose the top wall as the source node and the bottom wall as the sink node. Each particle in the system then represents a single intermediate node in the network.

The directed edges are used to represent the direction of force transmission in the flow network. Hence, we define the directed set of edges \mathcal{A} as follows. If particles i and j are in contact, let (i,j), $(j,i) \in \mathcal{A}$; if particle i contacts the source (top wall), let $(s,i) \in \mathcal{A}$; and if particle i contacts the sink (bottom wall), let $(i,t) \in \mathcal{A}$. Furthermore, we adopt the simplest possible capacity function for each directed edge, namely, the capacity function in which the contacts are uniformly weighted:

$$u_{ij} \triangleq \begin{cases} \infty, & \text{if } i = s \text{ or } j = t, \\ 1, & \text{otherwise.} \end{cases}$$
 (12)

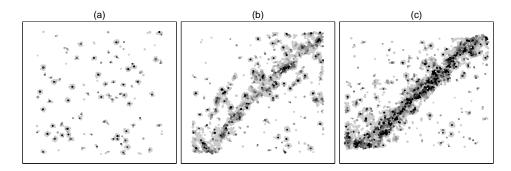


FIGURE 2. Spatial distribution of the normalized local non-affine strain over various axial strain intervals in the three stages of the loading history: (a) [0,0.018] of Stage 1, (b) [0.018,0.04] of Stage 2 and (c) [0.04,0.115] of Stage 3 showing the fully formed shear band. A particle is plotted at the location of the final strain state of the chosen strain interval and colored according to the highest magnitude of its normalized non-affine strain: black (highest) to white (zero).

This means that the maximum flow problem uses information on contact topology alone. We emphasize here that what makes a difference is the relative values of u_{ij} , not the actual values. Indeed, if f is a solution of Problem 2 with capacities u_{ij} , then γf is a solution of Problem 2 with new capacities γu_{ij} , where $\gamma > 0$ is a constant.

Although we define two oppositely directed edges for each pair of contacting particles, the flow value at that contact is simply defined as the absolute value of the difference between the flow values on the corresponding two oppositely directed edges. Consequently, reversing the flow direction, i.e. top wall as the sink and bottom wall as the source, will not alter the results.

In assigning a cost per unit of flow through the directed edges of the network, we make use of only the information on contact type, as defined earlier in equations (10) and (11). The contacts that dissipate energy are the plastic contacts; recall, these can be further classified as sliding, rolling, and slide-and-roll. In this study, we propose a very simple cost function in which the associated cost c_{ij}^w is based on the corresponding contact type as follows:

$$c_{ij}^{w} \triangleq \begin{cases} 1, & \text{if } i = s \text{ or } j = t, \\ 1, & \text{if the contact between particles } i \text{ and } j \text{ is elastic full-stick,} \\ 2, & \text{if the contact between particles } i \text{ and } j \text{ is plastic sliding,} \\ 2, & \text{if the contact between particles } i \text{ and } j \text{ is plastic rolling,} \\ 3, & \text{if the contact between particles } i \text{ and } j \text{ is plastic slide-and-roll.} \end{cases}$$

For comparison, we also consider the case in which all the contacts are of the same type, i.e.

$$c_{ij}^u \triangleq 1, \qquad (i,j) \in \mathcal{A}.$$
 (14)

 $c_{ij}^{u} \triangleq 1, \qquad (i,j) \in \mathcal{A}. \tag{14}$ The cost per unit of flow through the directed edge (i,j) can be defined as c_{ij}^{u} or c_{ij}^{w} . Note that, on average, only 15% of contacts are plastic. This means that about 85%of the edges in the flow network with $c_{ij} = c_{ij}^w$ and flow network with $c_{ij} = c_{ij}^u$ have the same costs on average. Note also that if we replace c_{ij} by γc_{ij} , where $\gamma > 0$ is a given constant, the solution of Problem 2 with costs c_{ij} is still a solution of Problem 2 with the new costs γc_{ij} .

In summary, we have deliberately constructed our flow networks for the simplest capacity and cost functions, making use solely of the contact topology and the contact types. In doing so, we establish a reference level or baseline information for future studies of MCMF solutions in which more sophisticated models of capacity and cost functions will be explored.

5.2. Minimum cost maximum flow solution. We consider the minimum cost of transmitting the maximum flow from source to sink, i.e. from the top wall to the bottom wall of the sample (or vice versa). Let f^* denote a solution of Problem 2. Then val (f^*) is the maximal value of a flow transmitted from source to sink. Note that the higher is the maximal value, the more effective is the network as a transmission medium. Given the uniform capacity function in equation (12), $val(f^*)$ may be used to quantify the influence of contact topology on the capacity of the contact network as a transmission medium for force. Using information on contact topology alone, we present in Figure 4(a), the strain evolution of val (f^*) : this suggests that force propagation through the contact network in the direction of maximum principal stress degrades as the system deforms. Note that the large fluctuations evident in the evolution throughout loading is indicative of the continual grain rearrangements occurring in the material which, in turn, leads to a continual reconfiguration of the contact network. Although large fluctuations persist, the maximum flow fluctuates about a near constant value in Stage 1, which is the highest average value for the three stages of the loading history. The sudden drop in the beginning of Stage 2 in Figure 4(a) is indicative of the onset of instabilities that will eventually precipitate the formation of the shear band in Stage 3 during which the lowest value of val (f^*) can be observed. These trends over the three stages of loading are entirely consistent with a recent study which show a progressive loss of stability of the sample in Stages 1 and 2 down to a near steady minimum in Stage 3 (see Figure 5 of [23]).

We next turn to $E(f^*)$, the minimum cost to transmit $\operatorname{val}(f^*)$ from source to sink. In particular, $E(f^*)/\operatorname{val}(f^*)$ represents the average minimum cost for one unit of maximum flow propagating from source to sink. Note that $E(f^*)/\operatorname{val}(f^*)$ depends on the path length from source to sink in the flow network (the longer the higher). Also note that the Euclidean distance between the top wall and the bottom wall is decreasing as the material deforms. Hence, we consider the average minimum cost for one unit of maximum flow propagating from source to sink per contact. Let

$$\langle c_{\min} \rangle \triangleq \frac{E(f^*)}{\text{val}(f^*)|\mathcal{P}|},$$
 (15)

where \mathcal{P} is the shortest path of the contact network (no weight on the edges) from source to sink and $|\cdot|$ denotes the number of elements in the set. Then, $\langle c_{\min} \rangle$ represents the average minimum cost for transmitting one unit of maximum flow on one contact. We call this the average minimum cost of maximum flow. Note that $\langle c_{\min} \rangle$ is unique even if the solution to Problem 2 is non-unique.

The evolution of $\langle c_{\min} \rangle$ with respect to the axial strain is shown in Figure 4(b). Keep in mind that both the capacity and cost function contribute to this quantity. We wish to test the sensitivity of the MCMF solutions to the assumed cost function and thus examine two flow networks associated with the two cost functions: $c_{ij} = c_{ij}^u$

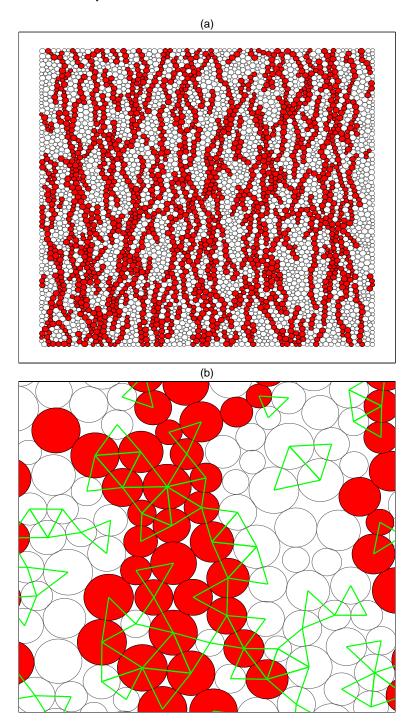


FIGURE 3. Spatial distribution of (a) major load-bearing particles known as force chains (red) and (b) force chains and 3-cycles (green lines) at peak stress ratio. Force chains self-organize into columns aligned in the direction of the maximum principal stress, here the vertical direction.

and $c_{ij} = c_{ij}^w$. The difference between these functions is small in the sense that all but a mere 15%, on average, of the edges in the flow network are plastic and have higher cost values (i.e. 2 and 3 as opposed to 1 assumed for the majority of contacts which are elastic). Apart from Stage 1, the flow network with $c_{ij} = c_{ij}^u$ is essentially constant for most of the loading history. By contrast, that with $c_{ij} = c_{ij}^w$ results in a strain evolution of $\langle c_{\min} \rangle$ which captures the three distinct stages of deformation, as identified earlier from the evolution of the global average of the local non-affine strain. During the initial stages of the strain-hardening regime, $\langle c_{\min} \rangle$ is observed to be increasing until reaching its highest throughout loading history. Then $\langle c_{\min} \rangle$ decreases rapidly as non-affine deformation and dissipation develop and intensify. In the large strain critical state regime, $\langle c_{\min} \rangle$ fluctuates about an essentially constant value. In this regime, recall that the sample has fully failed and deforms in the presence of a persistent shear band [21, 26]. Furthermore, it appears that the average minimum cost of maximum flow reveals the onset of instability as the curve in Figure 4(b) decreases as soon as Stage 2 commences. The difference between the average minimum cost for the two flow networks is as much as 0.35, around 23% of the highest value of the average minimum cost for $c_{ij} = c_{ij}^w$.

It is interesting to note that the weighted flow network with $c_{ij} = c^w_{ij}$ captures the three distinct stages of deformation while the unweighted $c_{ij} = c^u_{ij}$ fails to reproduce the observed evolution of the sample. This marked contrast in the evolution of the predicted minimum cost $\langle c_{\min} \rangle$ despite a small difference between the assumed cost functions, warrants further consideration. Due to the low percentage of slide-and-roll contacts (no more than 3% of the edges have cost values $c_{ij} = 3$), we consider a simpler version of the $c_{ij} = c^w_{ij}$ defined in equation (13), a binary function c^p_{ij} defined as follows:

$$c_{ij}^p \triangleq \begin{cases} x, & \text{if the contact between particles } i \text{ and } j \text{ is elastic,} \\ y, & \text{if the contact between particles } i \text{ and } j \text{ is plastic,} \end{cases}$$
 (16)

where x and y are given two distinct constants, not necessarily the integers in $c_{ij} = c_{ij}^w$. Then, $\langle c \rangle$, the average cost of transmitting one unit of flow through each contact, i.e.

$$\langle c \rangle \triangleq \frac{\sum\limits_{(i,j)\in\mathcal{A}} c_{ij}}{|\mathcal{A}|},$$
 (17)

can be written as

$$\langle c \rangle = \frac{xn_e + yn_p}{n_e + n_p} = \frac{x(n_e + n_p - n_p) + yn_p}{n_e + n_p} = x + (y - x)\frac{n_p}{n_e + n_p} = x + (y - x)R_p,$$

where n_e and n_p denote the number of elastic and plastic contacts, respectively, and R_p is the ratio of plastic contacts to the total number of contacts. Since x and y are given constants, the average of c_{ij}^p over all inter-particle contacts is a linear function of R_p (Pearsons product moment coefficient equals to 1). Note that we excluded the contacts between particle and wall, which is reasonable given these constitute no more than 0.7% of the total contacts over the entire loading history.

If $c_{ij} = c_{ij}^u$, we have a strain invariant $\langle c \rangle = 1$: apart from a relatively small deviation in Stage 1, the corresponding average minimum cost of maximum flow similarly remains essentially invariant for most of loading history (Figure 4(b)). If $c_{ij} = c_{ij}^w$, the average minimum cost $\langle c_{\min} \rangle$ also displays a strong correlation with the average cost $\langle c \rangle$: the Pearsons product moment coefficient is 0.88 for the entire loading history and 0.96 over Stages 2 and 3 (compare Figure 4(b)-(c)). These

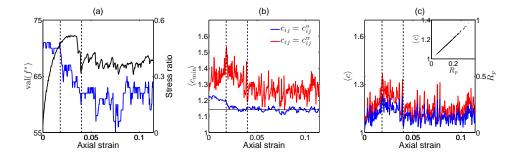


FIGURE 4. (a) Strain evolution of maximum flow val (f^*) . (b) Strain evolution of the average minimum cost of maximum flow $\langle c_{\min} \rangle$, defined by (15). (c) Strain evolution of the average cost $\langle c \rangle$, defined by (17), for $c_{ij} = c^w_{ij}$ (red) and ratio of plastic contacts to total R_p (blue). Inset in (c) shows the relationship between the average cost for $c_{ij} = c^w_{ij}$ and the ratio of plastic contacts to total contacts.

trends highlight the strong influence of the assumed cost function on the predicted average minimum cost $\langle c_{\min} \rangle$, in particular, the strong correlation between the ratio of plastic contacts to the total number of contacts, R_p , and the predicted $\langle c_{\min} \rangle$, for simple cost functions (i.e. equations (13) and (16)).

We can gain some insight into the nature of force transmission by examining the participation of plastic contacts in the minimum cost maximum flow pathway (or MCMF pathway). As we see from Figure 5(a), the transmission of flow through edges (contacts) with higher costs (i.e. the plastic contacts) is kept to a small fraction throughout loading: no more than 22% for $c_{ij} = c_{ij}^u$ and no more than 20% for $c_{ij} = c_{ij}^w$. The lower percentage of plastic contacts in the MCMF pathways for $c_{ij} = c_{ij}^w$ compared to those in $c_{ij} = c_{ij}^u$ suggests these plastic contacts are avoided as much as possible in the MCMF pathways of flow networks where they cost more. A similar trend can be seen in the percentage of plastic contacts in MCMF pathways (see Figure 5(b)): note here also the strong correlation with the corresponding average minimum cost of maximum flow (Pearsons product moment coefficient is 0.85 for the entire loading history and 0.97 over Stages 2 and 3). The transmission of maximum flow from the top wall to the bottom wall at minimum cost cannot, however, completely avoid the more costly plastic edges (contacts). As is evident in the spatial distribution of plastic contacts in MCMF pathway (Figure 6(a)), these contacts cut across the middle of the sample and appear to be strongly biased and concentrated in the region where the persistent shear band ultimately develops (compare Figure 6(a)-(b) in reference to the shear band in Figure 2(c)).

5.3. Bottleneck in the flow network. We now explore the bottlenecks from the minimum cuts of the flow networks in the critical state regime, during which the sample deforms in the presence of a fully developed shear band. Note that the minimum cuts of the contact networks are non-unique. The persistent shear band can be observed from the spatial distribution of the local non-affine strain, as shown earlier in Figure 2(c). Here in Figure 7 we plot the particles having contacts in bottleneck, accumulated over the critical state regime. Although we used only simple capacity function, the bottlenecks can be observed to lie in the locality of

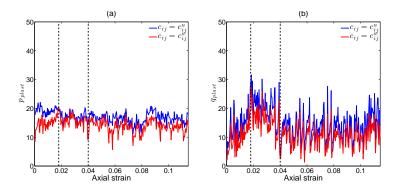


FIGURE 5. Strain evolution of (a) p_{plast} , percentage of plastic contacts in the system that are in MCMF pathways and (b) q_{plast} , percentage of contacts in MCMF pathways that are plastic.

the shear band. This result is consistent with the one found in [22] where the bottleneck requires the uniqueness of the minimum cut and, as such, the capacity of each contact there was defined in terms of the relative displacement between the contacting particles. However, in the method here, we explore the case in which all grain-grain contacts are given equal weights, allowing the maximum flow problem to be associated with multiple minimum cuts. Out of these minimum cuts we then examined the contacts or pathways in the contact network which are common to all the possible solutions of the MCMF problem, as we now discuss below.

5.4. Minimum cost maximum flow pathway. Here, we turn our attention to the minimum cost maximum flow pathway or MCMF pathway, a subset of edges that are used to transmit non-zero flow, and result in the maximum flow transmitted from source to sink at minimum cost. We are interested in this optimization problem for the transmission of force through the material in the direction of the maximum principal stress along which the force chains are known to align. Recall that the strong columnar force chains carry the majority of the applied load and, in turn, these are supported laterally by truss-like structures of 3-cycles. Thus we will focus on the relationship between the MCMF pathways (Figure 8(a)) and these self-organized structural building blocks (Figure 8(b)) that are chiefly responsible for the material's ability to resist and support applied forces.

While the load-bearing force chain particles stay essentially constant in population, the 3-cycles degenerate in Stages 1 and 2, reflecting the progressive loss of connectivity in the contact network typically seen for dense samples prior to peak stress ratio (Figure 8(b)). This defining feature of dense granular materials is due to dilatancy, the increase in void volume of the material in the stages of loading preceding the critical state regime (Stage 3). In [26] we showed that the loss of 3-cycles prior to Stage 3 is accompanied by an increase in the number of large cycles (e.g. 5-cycle, 6-cycle, etc.): 3-cycles open up and join with other cycles to form big cycles encircling the large voids. This loss in laterally supporting 3-cycles occurs under increasing axial load borne by the force chains thereby compromising their stability. Eventually the force chains become overloaded and buckle. But the failure of force chains by buckling is accompanied by the growth of new force chains. That the number of force chain particles in the system remains essentially constant

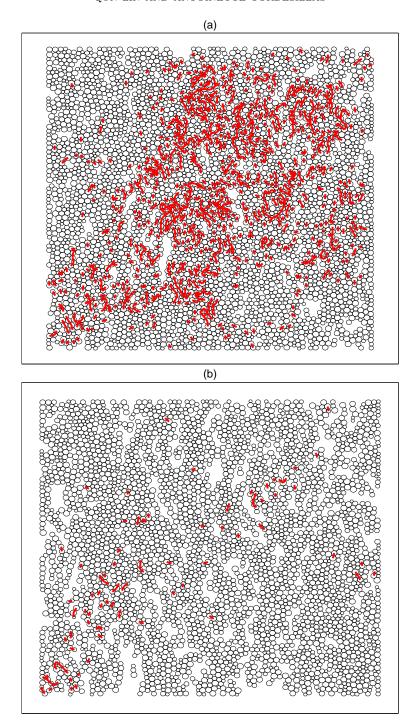


FIGURE 6. Spatial distribution of plastic contacts in MCMF pathway of the flow network with $c_{ij} = c^w_{ij}$ at (a) the beginning of Stage 2 and (b) the beginning of Stage 3 when the shear band in Figure 2(c) is fully formed. Red lines represent the plastic contacts and black outlined particles are the particles having contacts in MCMF pathway.

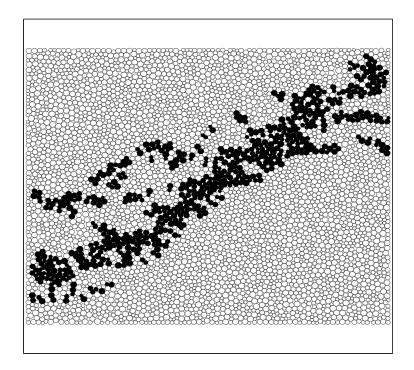


FIGURE 7. Spatial distribution of the bottleneck of flow network of each strain state, accumulated over the entire critical state regime (i.e. axial strains in the range [0.04, 0.115]).

(Figure 8(b)) throughout loading highlights the continual rearrangements of particles in the system: this in turn leads to the constant reconfiguration of the contact network and its MCMF pathway, as evident in the large fluctuations (Figure 8(a)).

As shown in Figure 9(a), and the spatial distributions of the participating versus nonparticipating force chain particles in Figure 10(a) — the pathways through which the maximum flow is transmitted at minimum cost, in the direction of the major principal stress — involve the great majority of the force chain particles. In general, more force chain particles have contacts in the MCMF pathways of the flow networks with the cost function defined according to contact types $c_{ij} = c_{ij}^w$, than the simplest conceivable (i.e. uniform) cost function $c_{ij} = c_{ij}^u$.

The highest participation of force chains in MCMF pathways can be seen in Stage 1, the initial states of the strain-hardening regime where deformation is globally affine and the sample is stable (recall Figures 1 and 2): on average around 93% for $c_{ij} = c^u_{ij}$ and 95% for $c_{ij} = c^w_{ij}$ are in MCMF pathways. It also appears that MCMF pathway is revealing of the onset of instability: Figure 9(a) (like the cost functions in Figure 4(b)) decreases upon commencement of Stage 2.

We now turn our attention to the 3-cycles which support the force chains. These too participate in and have contacts that are part of the MCMF pathway, as shown by Figure 10(b). Again, more of the 3-cycle particles are captured in the flow networks with the cost function decided according to contact types $c_{ij} = c_{ij}^w$ (Figure 9(b)). Similar trends can be seen in the force chains and 3-cycles combined (Figure 9(c)). It is instructive, as a final consideration, to examine the extent to

which the particles having contacts in MCMF pathways participate in the force chain and 3-cycle networks. Here we find most of the particles in MCMF pathway belong to force chains or 3-cycles, with the flow networks exhibiting almost identical percentages throughout loading (Figure 9(d)). This is in contrast to the earlier observed differences and sensitivity to the small difference in the cost function: from Figures 4 and 9 (a)-(c), the predicted average minimum cost, force chain membership in MCMF pathway, 3-cycle membership in MCMF pathway and force chain and 3-cycle membership in MCMF pathway are by as much as around 23% of the maximum value for $c_{ij} = c_{ij}^w$, 14%, 15%, 14%, respectively.

Altogether the results shown here suggest that a cost function which more accurately reflects the energy dissipated at the contact instead of just integer indices that classify contacts (i.e. 1 for elastic stick contact, 2 for sliding or rolling contact, and 3 for slide-and-roll contact) may possibly lead to an even higher if not total participation of force chains and 3-cycles in the MCMF pathways and vice versa. That said, we highlight the need for caution and the importance of undertaking a comprehensive analysis of the influence and sensitivity of the capacity and cost functions.

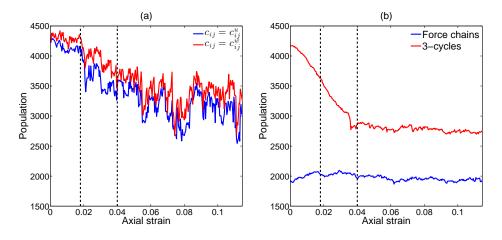


FIGURE 8. Strain evolution of population of (a) particles having contacts in MCMF pathways and (b) particles in force chains and 3-cycles.

6. Conclusions. In this paper, we considered two fundamental combinatorial optimization problems: the maximum flow problem and the minimum cost maximum flow (MCMF) problem. In particular, we identified the force bottlenecks and the MCMF pathway for a general flow network, constructed from the contact network of a deforming, densely-packed granular material. Our objective was to understand the extent to which force transmission is optimized, as the material responds to an applied load through grain rearrangements. These grain rearrangements, reflected in the evolving contact network, are responsible for the material's deformation. Understanding the relationship between deformation and force transmission is crucial to robust predictions, and ultimately control, of the mechanical response of granular materials to applied stresses and strains. In this study, we sought to answer the

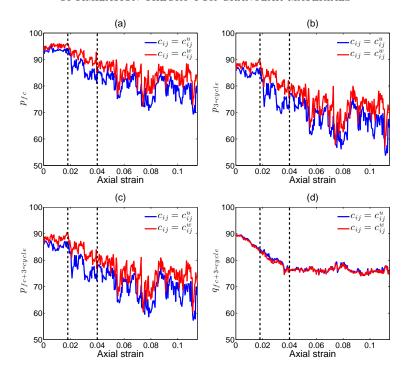


FIGURE 9. Strain evolution of participating force chains and 3-cycles in MCMF pathways: (a) p_{fc} , percentage of force chain particles that have contacts in MCMF pathway; (b) $p_{3\text{-cycle}}$, percentage of particles in 3-cycles that have contacts in MCMF pathway; (c) $p_{fc+3\text{-cycle}}$, percentage of particles in 3-cycles and force chains that have contacts in MCMF pathway; and (d) $q_{fc+3\text{-cycle}}$, percentage of particles with contacts in MCMF pathway that are also in force chains and 3-cycles. In (c) and (d), particles which are part of both a force chain and a 3-cycle are only counted once.

question: are these rearrangements taking place so as to optimize force transmission in the direction of the major (most compressive) principal stress — the direction in which the self-organized, load-bearing columns of force chains form?

We found that the great majority of the force chain particles are indeed transmitting maximum flow in the major principal stress direction at minimum cost. We observed as much as 93% for $c_{ij} = c^u_{ij}$ and 95% for $c_{ij} = c^w_{ij}$ in the stable, initial stages of the strain-hardening regime down to 79% for $c_{ij} = c^u_{ij}$ and 84% for $c_{ij} = c^w_{ij}$ during the failure regime. Inclusion of the particles in the supporting 3-cycles suggests that the MCMF pathway mainly lies in the particulate network comprising the force chains and their laterally supporting 3-cycles. We also examined the minimum cut using a capacity function that solely contains information on the contact topology: the edges of the flow networks, here representing the grain-grain contacts, are given equal weights. The resulting minimum cut uncovered bottlenecks in the locality of the persistent shear band — the mechanism by which the material fails.

In an attempt to establish a point of reference for future studies, we deliberately chose the capacity and cost functions to depend solely on contact topology and

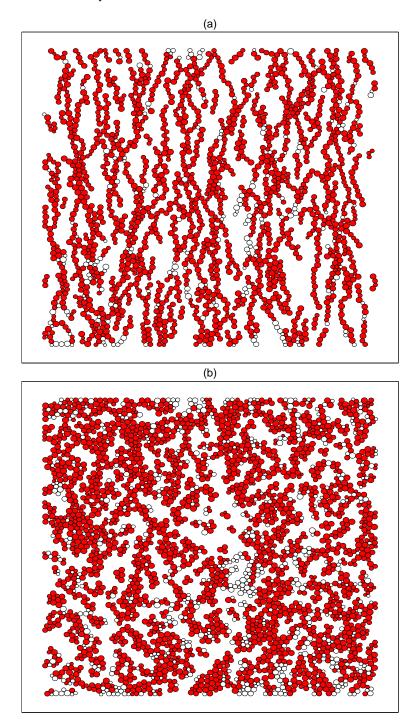


FIGURE 10. The great majority of force chains and 3-cycles participate in the transmission of maximum flow at minimum cost. Spatial distribution of particles in (a) force chains and (b) 3-cycles, at the peak stress ratio for $c_{ij} = c_{ij}^w$. Red (white) shaded particles are particles having contacts (no contacts) in MCMF pathway.

contact types. Despite these functions having the simplest possible form, the MCMF solutions captured the defining features of force transmission, i.e. force chains, 3cycles and the failure mechanism of shear banding. We probed the sensitivity of the minimum cost solutions, by comparing two flow networks of the same capacity function (i.e. uniform) but with two distinct cost functions, albeit the difference is small and affects a mere 15% of the edges in the graph on average. The differences in the predicted average minimum cost and membership of force chains and 3-cycles in MCMF pathways, suggest that the choice of capacity and cost functions are important and warrant further study. The strong influence of energy dissipation on force transmission is evident in the high correlation between the fraction of plastic contacts and the average minimum cost of transmitting the maximum flow through the network. Altogether the findings of this study lay bare the potential of optimization theory to contribute to the field of granular media research, and a strong impetus to explore, in particular, network-based optimization techniques [2, 3] and optimal control techniques [8, 12, 28, 29] in the rheological modeling of dense granular materials. Work is underway in establishing capacity and cost functions that more accurately reflect the salient aspects of force transmission and energy dissipation beyond contact topology — for a variety of loading conditions — not just in simulations but in experiments of synthetic as well as natural granular media.

Acknowledgments. This study was supported from grants to AT through the Australian Research Council Discovery Project DP120104759, the Melbourne Energy Institute seed fund, and the US Army Research Office Single Investigator Award W911NF-11-1-0175.

REFERENCES

- R. Arévalo, I. Zuriguel and D. Maza, Topology of the force network in the jamming transition of an isotropically compressed granular packing, Physical Review E, 81 (2010), 041302.
- [2] D. P. Bertsekas, "Network Optimization: Continuous and Discrete Models (Optimization, Computation, and Control)," Athena Scientific, 1998.
- [3] J. A. Bondy and U. S. R. Murty, "Graph Theory," Graduate Texts in Mathematics, 244. Springer, New York, 2008.
- [4] I. Cavarretta and C. O'Sullivan, The mechanics of rigid irregular particles subject to uniaxial compression, Géotechnique, 62 (2012), 681–692.
- [5] J. Duran, "Sands, Powders, and Grains: An Introduction to the Physics of Granular Materials," Springer-Verlag, New York, 2000.
- [6] J. Edmonds and R. M. Karp, Theoretical improvements in algorithmic efficiency for network flow problem, Journal of the Association for Computing Machinery, 19 (1972), 248–264.
- [7] A. Garg and R. Tamassia, A new minimum cost flow algorithm with applications to graph drawing, Graph Drawing, 1190 (1997), 201–216.
- [8] M. Gerdts and M. Kunkel, A nonsmooth Newton's method for discretized optimal control problems with state and control constraints, Journal of Industrial and Management Optimization, 4 (2008), 247–270.
- [9] F. S. Hillier and G. J. Lieberman, "Introduction to Operations Research," McGraw-Hill, 2005.
- [10] D. Jungnickel, "Graphs, Networks and Algorithms," Third edition. Algorithms and Computation in Mathematics, 5. Springer, Berlin, 2008.
- [11] Q. Lin and A. Tordesillas, Granular rheology: Fine tuned for optimal efficiency? Proceedings of the 23rd International Congress of Theoretical and Applied Mechanics, (2012).
- [12] R. C. Loxton, K. L. Teo, V. Rehbock and K. F. C. Yiu, Optimal control problems with a continuous inequality constraint on the state and the control, Automatica J. IFAC, 45 (2009), 2250–2257.
- [13] H. B. Mühlhaus and I. Vardoulakis, The thickness of shear bands in granular materials, Géotechnique, 37 (1987), 271–283.

- [14] M. Oda and H. Kazama, Microstructure of shear bands and its relation to the mechanisms of dilatancy and failure of dense granular soils, Géotechnique, 48 (1998), 465–481.
- [15] M. Oda, J. Konishi and S. Nemat-Nasser, Experimental micromechanical evaluation of strength of granular materials: Effects of particle rolling, Mechanics of Materials, 1 (1982), 269–283.
- [16] A. Ord and B. E. Hobbs, Fracture pattern formation in frictional, cohesive, granular material, Philosophical Transactions of the Royal Society A, 368 (2010), 95–118.
- [17] J. Paavilainen and J. Tuhkuri, Pressure distributions and force chains during simulated ice rubbling against sloped structures, Cold Regions Science and Technology, 85 (2013), 157–174.
- [18] J. M. Padbidri, C. M. Hansen, S. D. Mesarovic and B. Muhunthan, Length scale for transmission of rotations in dense granular materials, Journal of Applied Mechanics, 79 (2012), 031011.
- [19] F. Radjai, D. E. Wolf, M. Jean and J. J. Moreau, Bimodal character of stress transmission in granular packings, Physical Review Letters, 80 (1998), 61–64.
- [20] A. L. Rechenmacher, S. Abedi, O. Chupin and A. D. Orlando, Characterization of mesoscale instabilities in localized granular shear using digital image correlation, Acta Geotechnica, 6 (2011), 205–217.
- [21] A. Tordesillas, Force chain buckling, unjamming transitions and shear banding in dense granular assemblies, Philosophical Magazine, 87 (2007), 4987–5016.
- [22] A. Tordesillas, A. Cramer and D. M. Walker, Minimum cut and shear bands, Powders & Grains AIP Conference Proceedings 1542 (2013), 507–510.
- [23] A. Tordesillas, Q. Lin, J. Zhang, R. P. Behringer and J. Shi, Structural stability and jamming of self-organized cluster conformations in dense granular materials, Journal of the Mechanics and Physics of Solids, 59 (2011), 265–296.
- [24] A. Tordesillas, D. M. Walker, E. Andò and G. Viggiani, Revisiting localised deformation in sand with complex systems, Proceedings of the Royal Society of London Series A, (2013).
- [25] A. Tordesillas, D. M. Walker, G. Froyland, J. Zhang and R. P. Behringer, Transition dynamics and magic-number-like behavior of frictional granular clusters, Physical Review E, 86 (2012), 011306.
- [26] A. Tordesillas, D. M. Walker and Q. Lin, Force cycles and force chains, Physical Review E, 81 (2010), 011302.
- [27] D. M. Walker, A. Tordesillas, S. Pucilowski, Q. Lin, A. L. Rechenmacher and S. Abedi, Analysis of grain-scale measurements of sand using kinematical complex networks, International Journal of Bifurcation and Chaos, 22 (2012), 1230042.
- [28] L. Y. Wang, W. H. Gui, K. L. Teo, R. Loxton and C. H. Yang, Time delayed optimal control problems with multiple characteristic time points: Computation and industrial applications, Journal of Industrial and Management Optimization, 5 (2009), 705–718.
- [29] Y. Zhao and M. A. Stadtherr, Rigorous global optimization for dynamic systems subject to inequality path constraints, Industrial and Engineering Chemistry Research, 50 (2011), 12678– 12603

Received February 2013; 1st revision March 2013; final revision July 2013.

E-mail address: q.lin@curtin.edu.au
E-mail address: atordesi@ms.unimelb.edu.au