A stochastic approach to shortcut bridging in programmable matter

Marta Andrés Arroyo, Sarah Cannon, Joshua J. Daymude, Dana Randall & Andréa W. Richa

Springer

CrossMark

A stochastic approach to shortcut bridging in programmable matter

Marta Andrés Arroyo[1] · Sarah Cannon[2] · Joshua J. Daymude[3] · Dana Randall[4] · Andréa W. Richa[3]

Abstract

In a *self-organizing particle system*, an abstraction of programmable matter, simple computational elements called *particles* with limited memory and communication self-organize to solve system-wide problems of movement, coordination, and configuration. In this paper, we consider a stochastic, distributed, local, asynchronous algorithm for "shortcut bridging", in which particles self-assemble bridges over gaps that simultaneously balance minimizing the length and cost of the bridge. Army ants of the genus *Eciton* have been observed exhibiting a similar behavior in their foraging trails, dynamically adjusting their bridges to satisfy an efficiency trade-off using local interactions. Using techniques from Markov chain analysis, we rigorously analyze our algorithm, show it achieves a near-optimal balance between the competing factors of path length and bridge cost, and prove that it exhibits a dependence on the angle of the gap being "shortcut" similar to that of the ant bridges. We also present simulation results that qualitatively compare our algorithm with the army ant bridging behavior. Our work gives a plausible explanation of how convergence to globally optimal configurations can be achieved via local interactions by simple organisms (e.g., ants) with some limited computational power and access to random bits. The proposed algorithm also demonstrates the robustness of the stochastic approach to algorithms for programmable matter, as it is a surprisingly simple extension of our previous stochastic algorithm for compression.

Keywords Programmable matter · Self-organizing particle systems · Shortcut bridging · Distributed algorithms · Markov chains · Bio-inspired algorithms

1 Introduction

To develop a system of *programmable matter*, one endeavors to create a material or substance that utilizes user input or stimuli from its environment to change its physical properties in a programmable fashion. Many such systems have been proposed (e.g., DNA tiles, synthetic cells, and reconfigurable modular robots) and each attempts to perform tasks subject to domain-specific capabilities and constraints. In our work on *self-organizing particle systems*, we abstract away from specific settings and envision a system of computationally limited devices (which we call *particles*) that can actively move and individually execute distributed, local, asynchronous algorithms to cooperatively achieve macro-scale tasks of movement and coordination.

✉ Joshua J. Daymude
jdaymude@asu.edu

Marta Andrés Arroyo
martaandres@correo.ugr.es

Sarah Cannon
sarah.cannon@berkeley.edu

Dana Randall
randall@cc.gatech.edu

Andréa W. Richa
aricha@asu.edu

[1] University of Granada, Granada, Spain

[2] Computer Science Division, University of California, Berkeley, Berkeley, CA 94709, USA

[3] School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ 85281, USA

[4] College of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0765, USA

The phenomenon of local interactions yielding emergent, collective behavior is often found in natural systems; for example, honey bees choose hive locations based on decentralized recruitment (Camazine et al. 1999) and cockroach larvae perform self-organizing aggregation using pheromones with limited range (Jeanson et al. 2005). In this paper, we present an algorithm inspired by the work of Reid et al. (2015), who found that army ants continuously modify the shape and position of foraging bridges—constructed and maintained by their own bodies—across holes and uneven surfaces in the forest floor. These bridges appear to stabilize in a structural formation that balances the "benefit of increased foraging trail efficiency" with the "cost of removing workers from the foraging pool to form the structure" (Reid et al. 2015).

We attempt to capture this inherent trade-off in our algorithm for "shortcut bridging" in self-organizing particle systems (formally defined in Sect. 1.3). Our algorithm is an extension of the stochastic, distributed algorithm for *compression* introduced in Cannon et al. (2016), demonstrating that many fundamental elements of our stochastic approach can be generalized to applications beyond the specific context of compression, in which a particle system gathers together as tightly as possible. In particular, this stochastic approach may be of future interest in the molecular programming domain, where simpler variations of bridging have been studied. Groundbreaking works in this area, such as that of Mohammed et al. (2017), focus on forming molecular structures that connect some fixed points; our work may offer insights on further optimizing the quality and/or cost of the resulting bridges.

Shortcut bridging is an attractive goal for programmable matter systems, as many application domains envision deploying programmable matter on surfaces with structural irregularities or dynamic topologies. For example, one commonly imagined application of smart sensor networks is to detect and span small cracks in infrastructure such as roads or bridges; dynamic bridging behavior would enable the system to remain connected and shift position as cracks form.

1.1 Related work

When considering recently proposed and realized systems of programmable matter, one can distinguish between *passive* and *active* systems. In passive systems, computational units cannot control their movements and have (at most) very limited computational abilities, relying instead on their physical structure and interactions with the environment to achieve locomotion (e.g., Woods 2015; Angluin et al. 2006; Reid and Latty 2016). A large body of research in molecular self-assembly falls under this category, which has mainly focused on shape formation (e.g., Douglas et al.

2009; Cheung et al. 2011; Wei et al. 2012). In contrast, our work examines building dynamic bridges whose exact shape is not predetermined. Mohammed et al. studied a similar problem of connecting DNA origami landmarks with DNA nanotubes, using a carefully designed process of nanotube nucleation, growth, and diffusion to achieve and maintain the desired connections (Mohammed et al. 2017). Significant differences between their approach and ours are: (1) the bridges we consider already connect their endpoints at the start and we focus on the specific goal of optimizing their shape with respect to a parameterized objective function, and (2) our system is active as opposed to passive.

Active systems are composed of computational units that can control their actions to solve a specific task. Examples include *swarm robotics*, various other models of modular robotics, and the *amoebot model*, which is our computational framework (detailed in Sect. 1.2).

Swarm robotic systems usually involve collections of autonomous robots moving freely in space with limited sensing and communication ranges. These systems can perform a variety of tasks including gathering (Cieliebak et al. 2012), shape formation (Flocchini et al. 2008; Rubenstein et al. 2014), and imitating the collective behavior of natural systems (Chazelle 2009); however, the individual robots typically have more powerful communication and processing capabilities than those we consider. *Modular self-reconfigurable robotic systems* focus on the motion planning and control of kinematic robots to achieve dynamic morphology (Yim et al. 2007), and *metamorphic robots* form a subclass of self-reconfiguring robots (Chirikjian 1994) that share some characteristics with our geometric amoebot model. Walter et al. have conducted some algorithmic research on these systems (e.g., Walter et al. 2004a, b), but focus on problems disjoint from those we consider.

In the context of molecular programming, our model most closely relates to the *nubot* model by Woods et al. (2013), Chen et al. (2015), which seeks to provide a framework for rigorous algorithmic research on self-assembly systems composed of active molecular components, emphasizing the interactions between molecular structure and active dynamics. This model shares many characteristics with our amoebot model (e.g., space is modeled as the triangular lattice, nubot monomers have limited computational abilities, and there is no global orientation) but differs in that nubot monomers can replicate or die and can perform coordinated rigid body movements. These additional capabilities prohibit direct translation of results under the nubot model to our amoebot model.

1.2 The amoebot model

Our computational framework is the *amoebot model* (Daymude et al. 2017b), originally proposed in Derakhshandeh et al. (2014) as an abstract model for programmable matter designed to enable rigorous algorithmic research on nano-scale systems. We envision programmable matter as a collection of individual, homogeneous computational elements called *particles*. The structure of a particle system is represented as a connected subgraph of the infinite, undirected graph $G = (V, E)$, where V is the set of all locations a particle can occupy relative to its structure and E is the set of all atomic movements between locations in V. Each location in V can be occupied by at most one particle at a time. For shortcut bridging (and many other problems), we further assume the *geometric* amoebot model, in which $G = G_\Delta$, the *triangular lattice*[1] (Fig. 1a).

Each particle is either *contracted*, occupying a single location, or *expanded*, occupying a pair of adjacent locations in G_Δ (Fig. 1b). Particles move via a series of *expansions* and *contractions*: a contracted particle may expand into an adjacent unoccupied location, and completes its movement by contracting to once again occupy a single location. An expanded particle's *head* is the location it last expanded into and the other location it occupies is its *tail*; a contracted particle's head and tail are the same location.

Two particles occupying adjacent locations in G_Δ are said to be *neighbors*. Each particle is *anonymous*, lacking a unique identifier, but can locally identify each of its neighboring locations and can determine which of those locations are occupied by particles. Each particle has a constant-size, local memory that its neighbors can directly read from for communication. A particle's memory stores whether it is contracted or expanded and identifies if neighboring locations are incident to its head or tail. Particles do not have access to any global information such as a global compass or an estimate of the size of the system.

We assume the standard asynchronous model from distributed computing (see, e.g., Lynch 1996), where a system progresses through *atomic actions*. A classical result under this model states that for any concurrent asynchronous execution of atomic actions, there is a sequential ordering of actions producing the same end result, provided conflicts that arise in the concurrent execution are resolved. In our setting, an atomic action is an activation of a single particle. Once activated, a particle can perform an arbitrary, bounded amount of computation involving its local memory and the memories of its
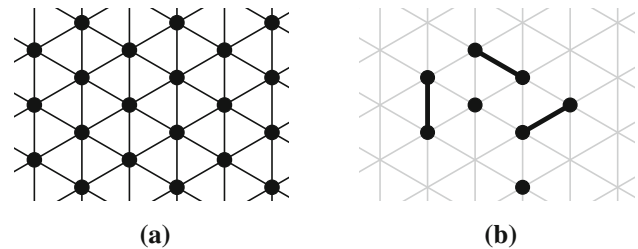

Fig. 1 a A section of the triangular lattice G_Δ; **b** expanded and contracted particles

neighbors, and can perform at most one contraction or expansion. We assume conflicts arising from simultaneous particle expansions into the same unoccupied location are resolved arbitrarily such that at most one particle is expanding into a given location at a time. Thus, while in reality many particles may be active concurrently, it suffices when analyzing our algorithm to consider a sequence of activations where only one particle is active at a time.

Terminology for particle systems In addition to the formal model, we introduce some terminology for our application of shortcut bridging. Just as the uneven surfaces of the forest floor affect the foraging behavior of army ants, the collective behavior of particle systems should change when G_Δ is non-uniform. Here, we focus on system behaviors when the locations of G_Δ are either *gap* (unsupported) or *land* (supported). A particle can tell whether its location is a gap location or a land location. An *object* is a static particle that does not perform computation; these are used to keep the particle system connected to certain fixed sites.

A particle system *configuration* is the finite set of occupied locations of G_Δ. An *edge* of a configuration is an edge of G_Δ where both endpoints are occupied by particle tails.[2] When referring to a *path*, we mean a path of such edges. Two particles are *connected* if there exists a path between them, and a configuration is *connected* if all pairs of particles are. A *hole* in a configuration is a maximal finite component of adjacent unoccupied locations. We specifically consider connected configurations with no holes, and our algorithm—if starting at such a configuration—will maintain these properties, a fact we will prove in Sect. 3.2.

Let σ be a connected configuration with no holes. The (single, external) *boundary* of σ is the walk composed of all edges in σ between particles that are not surrounded (i.e., those with less than 6 neighbors).[3] In order to analyze the strength of the solutions our algorithm produces, we define

---

[1] Our past works refer to G_Δ as the *equilateral triangular grid graph* G_{eqt} and the triangular lattice Γ.

[2] Lattice edges incident to a node occupied by an expanded particle's head are not counted as configuration edges, since these are exploratory and temporary. This is explained further in Sect. 3.1.

[3] Note that an edge may appear twice in the boundary if it is a cut-edge (e.g., the bottom-left most edges in Fig. 3b).

the *weighted perimeter* $\overline{p}(\sigma, c)$ to be the summed weight of the edges on the boundary of σ, where edges between land locations have weight 1, edges between gap locations have weight $c > 1$, and edges with one endpoint on land and one endpoint in the gap have weight $(1 + c)/2$.

1.3 Problem description

In the *shortcut bridging problem*, we consider an instance $(L, O, \sigma_0, c, \alpha)$, where $L \subseteq V$ is the set of land locations, O is the set of (two) objects to bridge between, σ_0 is the initial configuration of the particle system, $c > 1$ is a fixed weight for edges between gap locations, and $\alpha > 1$ is a parameter capturing our error tolerance. An instance is *valid* if (*i*) the objects of O and particles of σ_0 all occupy locations in L, (*ii*) σ_0 connects the objects, and (*iii*) σ_0 is connected. A (distributed) algorithm *solves* a valid instance $(L, O, \sigma_0, c, \alpha)$ if, beginning from σ_0, it reaches and remains in a set of configurations Σ^* such that any $\sigma \in \Sigma^*$ has weighted perimeter $\overline{p}(\sigma, c)$ within an α-factor of its minimum possible value, with high probability.[4]

In analogy to the apparatus used in Reid et al. (2015) (Fig. 3a), we are particularly interested in instances where L forms a V-shape, O has two objects positioned at either base of L, and σ_0 lines the interior sides of L, as in Fig. 2a. However, our algorithm is not limited to this setting; for example, we show simulation results for an N-shaped land mass (Fig. 2b) in Sect. 5.

The weighted perimeter balances the trade-off observed in Reid et al. (2015) between the competing objectives of establishing a short path between the fixed endpoints while not having too many particles in the gap. Although both metrics are amenable to our analysis, we focus on weighted perimeter instead of the number of particles in the gap for two reasons. First, the structure and thickness of bridges produced using weighted perimeter more closely resemble those of ant bridges, while using particles in the gap results in consistently thin, jagged structures (see Fig. 3b vs. c). Second, only particles on the perimeter can move, and thus recognize the potential risk of being in the gap.

2 Approach, techniques, and results

In Cannon et al. (2016), we introduced a stochastic, distributed algorithm for compression in the amoebot model; here we extend that work to show our stochastic approach is more widely applicable.

---

[4] An event occurs *with high probability (w.h.p.)* if the probability of success is at least $1 - 1/\mathrm{poly}(n)$; here, n is the number of particles.

2.1 The stochastic approach to particle systems

In the stochastic approach to self-organizing particle systems, we use concepts from statistical physics to design our algorithms, a process we outline here. At a high level, we define an energy function that captures our objectives for the particle system and then design a Markov chain that, in the long run, favors configurations with desirable energy values. Care is taken to ensure this Markov chain can be executed in a distributed, asynchronous manner by each particle individually. While understanding our approach and motivation is not necessary for understanding our results, it provides further insights into our methodologies.

In statistical physics, ensembles of particles similar to those we consider represent physical systems and demonstrate that local micro-behavior can induce global macro-scale changes to the system (Baxter et al. 1980; Blanca et al. 2018; Restrepo et al. 2013). Like a spring relaxing, physical systems favor configurations that minimize energy. Each configuration σ has energy determined by a *Hamiltonian* $H(\sigma)$, and we then assign each a weight $w(\sigma) = e^{-B \cdot H(\sigma)}$, where $B = 1/T$ is inverse temperature. Markov chains have been well-studied as a tool for sampling configurations of these systems with probability proportional to $w(\sigma)$, that is, with probability $w(\sigma)/Z$, where $Z = \sum_\tau e^{-B \cdot H(\tau)}$ is the normalizing constant known as the *partition function*. The configurations with the lowest values of $H(\sigma)$—those with the least energy—are most likely to be sampled.

For shortcut bridging, we introduce a Hamiltonian over particle system configurations that assigns the lowest energy values to configurations with desirable bridge structures; we then design our algorithm to favor these configurations with small Hamiltonians. We assign each configuration σ a Hamiltonian $H(\sigma) = \overline{p}(\sigma, c)$, its weighted perimeter. Setting $\lambda = e^B$, we get $w(\sigma, c) = \lambda^{-\overline{p}(\sigma, c)}$, where $w(\sigma, c)$ is the likelihood with which we want our algorithm to yield σ. As λ gets larger (by increasing B, effectively lowering temperature), these weights increasingly favor configurations where $H(\sigma) = \overline{p}(\sigma, c)$ is small and the desired bridging behavior is exhibited. Using a Markov chain, we will ensure that the eventual probability with which we are at state σ is $w(\sigma, c)/Z$, where $Z = \sum_\tau w(\tau, c)$ in the necessary normalizing factor.

2.2 Markov chains

We briefly review relevant terminology on Markov chains. A *Markov chain* $\mathcal{M}$ is a memoryless stochastic process defined on a state space Ω. We only consider Ω which are finite and discrete; in particular, the states of Ω will be connected, hole-free configurations with a common land
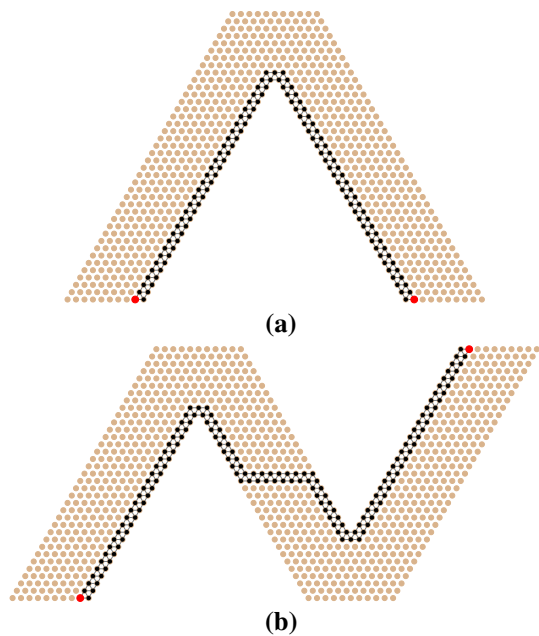
(a)

(b)

Fig. 2 Example initial configurations σ_0 of particles (black) connecting two objects O (large, red) on land masses L (brown and black) for two instances of the shortcut bridging problem for which we present simulation results (Sect. 5). (Color figure online)
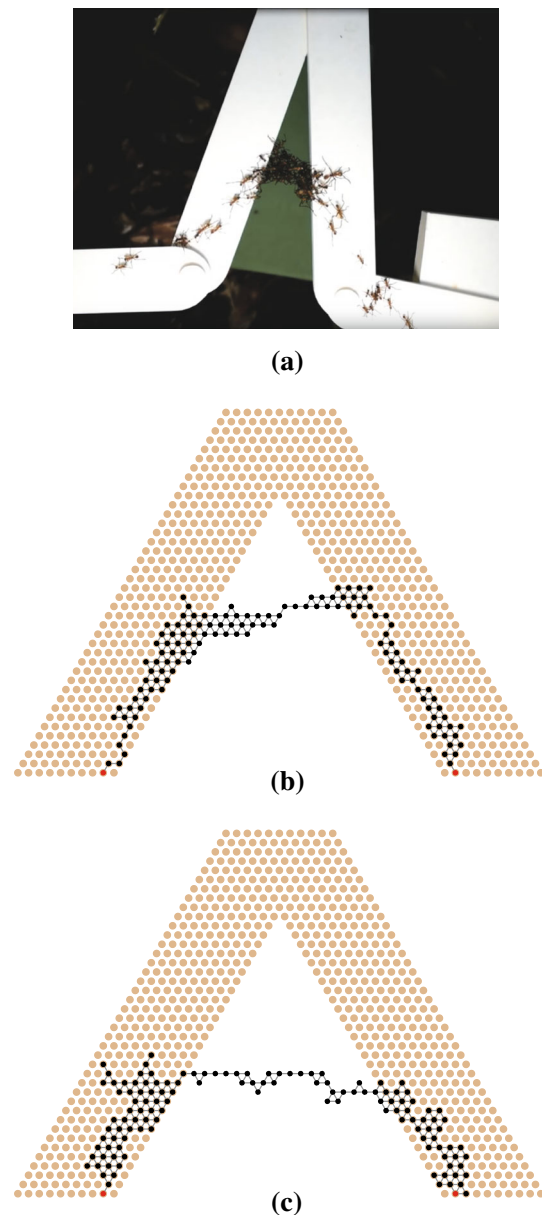

(a)

(b)

(c)

Fig. 3 **a** In this image from Reid et al. (2015), army ants of the genus *Eciton* build a dynamic bridge which balances the benefit of a shortcut path with the cost of committing ants to the structure. **b** Our shortcut bridging algorithm also balances competing objectives and converges to similar configurations. **c** Minimizing the number of particles in the gap instead of the weighted perimeter results in thin bridges with large clusters of particles on land that do not resemble the ant bridges as closely

mass L, objects O, and number of particles n. The transition matrix $Q : \Omega \times \Omega \to [0, 1]$ of a Markov chain $\mathcal{M}$ is defined so that $Q(\sigma, \tau)$ is the probability of moving from state σ to state τ in one step, for any pair of states $\sigma, \tau \in \Omega$. For our Markov chain, transitions will correspond to one particle moving one unit in one direction, and the probabilities of these transitions will be chosen carefully. The t-step transition probability $Q^t(\sigma, \tau)$ is the probability of moving from σ to τ in exactly t steps.

A Markov chain is *irreducible*, or its state space is *connected*, if there is a sequence of valid transitions from any state to any other state, i.e., for all $\sigma, \tau \in \Omega$, there is a t such that $Q^t(\sigma, \tau) > 0$. A Markov chain is *aperiodic* if for all $\sigma, \tau \in \Omega$, $\gcd\{t : Q^t(\sigma, \tau) > 0\} = 1$. A Markov chain is *ergodic* if it is both irreducible and aperiodic. Any finite, ergodic Markov chain converges to a unique *stationary distribution* π given by, for all $\sigma, \tau \in \Omega$, $\lim_{t \to \infty} Q^t(\sigma, \tau) = \pi(\tau)$. Any distribution π' satisfying $\pi'(\sigma)Q(\sigma, \tau) = \pi'(\tau)Q(\tau, \sigma)$ for all $\sigma, \tau \in \Omega$ (the *detailed balance condition*) must be this unique stationary distribution (see, e.g., Feller 1968).

Given a state space Ω, a set of allowable transitions between states, and a desired stationary distribution π on Ω (e.g., $\pi(\sigma) \sim w(\sigma, c)$), the celebrated Metropolis–Hastings algorithm (Hastings 1970) gives a Markov chain on Ω that uses only allowable transitions and has stationary distribution π. This is accomplished by carefully setting the probabilities of the state transitions as follows. Starting at

$\sigma \in \Omega$, pick a neighbor $\tau \in \Omega$ (i.e., a state τ to which σ has an allowable transition) uniformly with probability $1/(2\Delta)$, where Δ is the maximum number of neighbors of any state, and move to τ with probability $\min\{1, \pi(\tau)/\pi(\sigma)\}$; with the remaining probability stay at σ and repeat. If the allowable transitions suffice to go between any two states of Ω, then π must be the stationary distribution by detailed balance. While calculating $\pi(\tau)/\pi(\sigma)$ seems to require

global knowledge, this ratio can often be calculated using only local information when many terms cancel out. For shortcut bridging, because our desired stationary distribution will be $\pi(\sigma) = w(\sigma, c)/Z = \lambda^{-\overline{p}(\sigma,c)}/Z$ where $Z = \sum_\tau w(\tau, c)$, the Metropolis-Hastings probabilities can be written as $\min\{1, \lambda^{\overline{p}(\sigma,c) - \overline{p}(\tau,c)}\}$. Using this probability calculation to decide whether or not to make a transition is a *Metropolis filter*. Importantly, if σ and τ only differ by one particle P, as is the case with all moves of our algorithm, then $\overline{p}(\sigma, c) - \overline{p}(\tau, c)$, the difference in weighted perimeter due to particle P's move, can be calculated using only local information from the neighborhood of P (Lemma 1).

2.3 Results

We present a Markov chain $\mathscr{M}$ for *shortcut bridging* in the geometric amoebot model that translates directly to a fully distributed, local, asynchronous algorithm $\mathscr{A}$. This Markov chain $\mathscr{M}$ uses only local moves and, using a Metropolis filter, eventually reaches a distribution that favors configurations proportional to their weight $w(\sigma, c)$. Thus, configurations with smaller weighted perimeter $\overline{p}(\sigma, c)$ are more likely, as desired. Rather than terminating $\mathscr{M}$ at some point and using the resulting configuration as a random sample (as is often done with Markov chains) we instead run $\mathscr{M}$ indefinitely, moving among different configurations but remaining at the stationary distribution π, which we prove meets our desired objectives with high probability.

We prove that $\mathscr{M}$ (and by extension, $\mathscr{A}$) solves the shortcut bridging problem: for any constant $\alpha > 1$, for appropriately chosen values of parameters, the long run probability that $\mathscr{M}$ is in a configuration σ with $\overline{p}(\sigma, c)$ larger than α times its minimum possible value is exponentially small. The key tool used to establish this is a careful *Peierls argument*, used in statistical physics to study non-uniqueness of limiting Gibbs measures and in computer science to establish slow mixing of Markov chains (see, e.g., Levin et al. 2009, Chapter 15). We then specifically consider V-shaped land masses with an object on each branch of the V, and prove that the resulting bridge structures vary with the interior angle of the V-shaped gap being shortcut—a phenomenon also observed by Reid et al. (2015) in the army ant bridges—and show in simulation that they are qualitatively similar to those of the ants (e.g., Fig. 3).

Our shortcut bridging algorithm and others developed with the stochastic approach (e.g., Cannon et al. 2016) have several advantages over other algorithms for programmable matter and self-organizing particle systems. They are nearly oblivious, only needing to store at most one bit of information between activations. They are also more robust to failures; even if particles crash and stop moving, our algorithm will converge to the best bridge possible with respect to the crashed particles' fixed locations. On the other hand, other algorithms for particle systems (e.g. Daymude et al. 2017a; Derakhshandeh et al. 2017) would fail even with a single particle crash. Finally, our algorithm requires little to no communication between particles. Since these algorithms are derived from stochastic processes, powerful tools developed to analyze Markov chains can be employed to rigorously understand their behavior.

3 A stochastic algorithm for shortcut bridging

Recall that for the shortcut bridging problem, we desire for our algorithm to achieve small weighted perimeter, where boundary edges in the gap cost a factor of $c > 1$ more than those on land. The algorithm must balance the competing objectives of having a short path between the two objects while not forming too large of a bridge. We capture these factors by preferring configurations σ that have both small *perimeter* $p(\sigma)$, the length of the walk around the boundary of the particle system, and small *gap perimeter* $g(\sigma)$, the number of perimeter edges that are in the gap, where edges with one endpoint in the gap and one endpoint on land count as half an edge in the gap. While these objectives may appear to be aligned rather than competing, decreasing the length of the overall perimeter increases the gap perimeter and vice versa in the problem instances we consider (e.g., Fig. 2). We note that $\overline{p}(\sigma, c) = p(\sigma) + (c - 1)g(\sigma)$, and thus minimizing weighted perimeter is equivalent to simultaneously minimizing both perimeter and gap perimeter.

Our Markov chain algorithm incorporates two bias parameters: λ and γ. The value of λ controls the preference for having small perimeter, while γ controls the preference for having small gap perimeter. In this paper, we only consider $\lambda > 1$ and $\gamma > 1$, which correspond to favoring small perimeter and small gap perimeter, respectively. Using a Metropolis filter, we ensure our algorithm converges to stationary distribution π given by $\pi(\sigma) = \lambda^{-p(\sigma)}\gamma^{-g(\sigma)}/Z$ where $Z = \sum_\tau \lambda^{-p(\tau)}\gamma^{-g(\tau)}$ is the normalizing factor necessary to make π a probability distribution. Arithmetic shows:

$$\lambda^{-\overline{p}(\sigma,c)} = \lambda^{-p(\sigma)-(c-1)g(\sigma)} = \lambda^{-p(\sigma)}\left(\lambda^{c-1}\right)^{-g(\sigma)},$$

so setting $\gamma = \lambda^{c-1}$ yields our desired stationary distribution.

We note λ is the same parameter that controlled compression in Cannon et al. (2016), where particle configurations converged to a distribution proportional to $\lambda^{-p(\sigma)}$.

That work showed that $\lambda > 1$ is not sufficient to ensure compression, so we restrict our attention to $\lambda > 2 + \sqrt{2}$, the regime where compression provably occurs.

To ensure our algorithm maintains some desired invariants throughout its execution, we introduce two properties every movement must satisfy. Specifically, these properties maintain system connectivity,[5] prevent holes from forming, and ensure it is possible for our Markov chain to be reversible; more details can be found in Cannon et al. (2016). These last two conditions are necessary for applying established tools from Markov chain analysis.

We use the following notation. For a location ℓ, let $N(\ell)$ denote the set of particles and objects[6] adjacent to ℓ. For adjacent locations ℓ and ℓ', we use $N(\ell \cup \ell')$ to denote the set $N(\ell) \cup N(\ell')$, excluding particles or objects occupying ℓ or ℓ'. Let $\mathbb{S} = N(\ell) \cap N(\ell')$ be the particles and objects adjacent to both locations; we note $|\mathbb{S}| \in \{0, 1, 2\}$. The following properties can be locally checked by an expanded particle occupying ℓ and ℓ' (e.g., as in Step 6 of $\mathscr{M}$, Algorithm 1), and are symmetric with respect to these locations.

Property 1 $|\mathbb{S}| \in \{1, 2\}$ and every particle or object in $N(\ell \cup \ell')$ is connected to a particle or object in $\mathbb{S}$ by a path through $N(\ell \cup \ell')$.

Property 2 $|\mathbb{S}| = 0$, ℓ and ℓ' each have at least one neighbor, all particles and objects in $N(\ell) \setminus \{\ell'\}$ are connected by paths within this set, and all particles and objects in $N(\ell') \setminus \{\ell\}$ are connected by paths within this set.

We can now present our Markov chain $\mathscr{M}$ for an instance $(L, O, \sigma_0, c, \alpha)$ of shortcut bridging. For input parameter $\lambda > 2 + \sqrt{2}$, set $\gamma = \lambda^{c-1}$. Beginning at initial configuration σ_0, which we assume is connected and hole-free[7], repeat the steps of Algorithm 1.

Algorithm 1 Markov Chain $\mathscr{M}$ for Shortcut Bridging

1: Choose a particle P uniformly at random (u.a.r.) from all n particles; let ℓ be its location.
2: Choose a neighboring location ℓ' and $q \in (0, 1)$ u.a.r.
3: **if** ℓ' is unoccupied **then**
4:     P expands to occupy both ℓ and ℓ'.
5:     Let σ (resp., σ') be the configuration with P at ℓ (resp., at ℓ').
6:     **if** (i) ℓ and ℓ' satisfy Property 1 or Property 2, (ii) $|N(\ell)| < 5$, and (iii) $q < \lambda^{p(\sigma)-p(\sigma')}\gamma^{g(\sigma)-g(\sigma')}$ **then** P contracts to ℓ'.
7:     **else** P contracts back to ℓ.

Conditions (i) and (ii) of Step 6 ensure that the particle system remains connected and no new holes are formed during the execution of $\mathscr{M}$. In particular, condition (ii) explicitly disallows a particle with five neighbors from moving into the only unoccupied location in its neighborhood, as doing so would create a hole. Condition (iii) is the Metropolis filter discussed above; the proposed particle move, once confirmed to be valid, only occurs with probability

$$\min\{1, \lambda^{p(\sigma)-p(\sigma')}\gamma^{g(\sigma)-g(\sigma')}\} = \min\{1, \lambda^{\overline{p}(\sigma,c)-\overline{p}(\sigma',c)}\},$$

where σ is the configuration with P at location ℓ and σ' is the configuration with P at location ℓ'. Although $p(\sigma) - p(\sigma')$ and $g(\sigma) - g(\sigma')$ are values defined at system-level scale, we show these differences can be calculated locally.

Lemma 1 *An expanded particle P occupying adjacent locations ℓ and ℓ' in G_Δ can calculate the values of $p(\sigma) - p(\sigma')$ and $g(\sigma) - g(\sigma')$ in Step 6(iii) of $\mathscr{M}$ using only local information involving ℓ, ℓ', and $N(\ell \cup \ell')$.*

Proof Observe that these values need only be calculated if conditions (i) and (ii) of Step 6 holds. By a result of Cannon et al. (2016),

$$p(\sigma) - p(\sigma') = |N(\ell')| - |N(\ell)|,$$

which can be calculated using only local information.

Recall that gap perimeter is defined as the number of boundary edges in the gap, counting edges between gap and land as half an edge; this is equal to the number of particles that are on the perimeter and in the gap, counted with appropriate multiplicity if a particle appears on the perimeter more than once. Given a particle R and a configuration τ, let $G(R, \tau)$ be equal to 1 if R occupies a gap location in τ and 0 otherwise. Let $\delta(R, \tau)$ be the number of times R appears on the perimeter of τ. Then the desired difference is:

$$g(\sigma) - g(\sigma') = \sum_R [G(R, \sigma)\delta(R, \sigma) - G(R, \sigma')\delta(R, \sigma')].$$

Define $\Delta(R) = \delta(R, \sigma) - \delta(R, \sigma')$. For particle P, since conditions (i) and (ii) of Step 6 hold, $\Delta(P) = 0$. For any particle $R \notin \{P\} \cup N(\ell \cup \ell')$, $\Delta(R) = 0$ since its neighborhood is not affected by the movement of P. Moreover,

---

for any particle $R \neq P$, $G(R, \sigma) = G(R, \sigma')$ since it does not move. So:

$$g(\sigma) - g(\sigma') = \delta(P, \sigma)[G(P, \sigma) - G(P, \sigma')] + \sum_{R \in N(\ell \cup \ell')} G(R, \sigma) \Delta(R).$$

The first term is easily calculated locally. For the summation, it remains to show that P can locally calculate $\Delta(R)$ for any $R \in N(\ell \cup \ell')$. First suppose that R is occupies a location adjacent to ℓ but not ℓ'. Then:

$$\Delta(R) = \begin{cases} -1 & \text{if } R \text{ has two neighbors in } N(\ell), \\ 1 & \text{if } R \text{ has no neighbors in } N(\ell), \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

The opposite is true if R occupies a location adjacent to ℓ' but not ℓ. Lastly, suppose R occupies a location adjacent to both ℓ and ℓ'. Then:

$$\Delta(R) = \begin{cases} 0 & \text{if } R \text{ has zero or two neighbors in } N(\ell \cup \ell'), \\ -1 & \text{if } R \text{ shares a neighbor with } \ell \text{ but not } \ell', \text{ and} \\ 1 & \text{if } R \text{ shares a neighbor with } \ell' \text{ but not } \ell. \end{cases}$$

In all cases, P can calculate $\Delta(R)$, and thus also $g(\sigma) - g(\sigma')$, using only local information. $\qquad\square$

The state space Ω of $\mathcal{M}$ is the set of all configurations reachable from σ_0 via valid transitions of $\mathcal{M}$. We conjecture that this includes all connected, hole-free configurations of n particles connected to both objects, but proving all such configurations are reachable from σ_0 is not necessary for our results. (The proof of the corresponding result in Cannon et al. (2016) does not generalize due to the presence of static objects).

3.1 From $\mathcal{M}$ to a distributed, local algorithm $\mathcal{A}$

In order for individual particles to run $\mathcal{M}$, a Markov chain with centralized control, we must translate $\mathcal{M}$ into a distributed, local, asynchronous algorithm $\mathcal{A}$ that fully respects the constraints of the amoebot model (Sect. 1.2). In particular, the uniformly at random particle selection in Step 1 of $\mathcal{M}$ must be translated to individual, asynchronous particle activations and a particle's combined expansion and contraction in Steps 4–7 of $\mathcal{M}$ must be decoupled into two separate activations because a particle can perform at most one movement per activation. The remainder of $\mathcal{M}$ can be executed directly in $\mathcal{A}$: Properties 1 and 2 are locally verifiable as they only involve a particle's immediate neighborhood, and Lemma 1 showed that the differences $p(\sigma) - p(\sigma')$ and $g(\sigma) - g(\sigma')$ used in Step 6 of $\mathcal{M}$ can be calculated locally. Full details of this construction can be found in Cannon et al. (2016).

Under the usual assumptions of the asynchronous model from distributed computing, one cannot assume that the next particle to be activated is equally likely to be any particle, as specified in Step 1 of $\mathcal{M}$. To mimic this uniformly random activation sequence in a local way, we assume each particle has its own Poisson clock with mean 1 and activates after a delay t drawn with probability e^{-t}. After completing its activation, a new delay is drawn to its next activation, and so on. The exponential distribution guarantees that, regardless of which particle has just activated, all particles are equally likely to be the next to activate (see, e.g., Feller 1968). We could even better approximate asynchronous activation sequences by allowing each particle to have its own constant mean for its Poisson clock, allowing for some particles to activate more often than others in expectation. In this setting, the probability that a particle P is the next of the n particles to activate is not $1 / n$, but rather some probability a_P that depends on all particles' Poisson means.[8] This does not change the stationary distribution of $\mathcal{M}$; Lemma 4 still holds with a nearly identical proof that replaces $1 / n$ with a_P, and Lemma 5 and Theorem 1 still follow. Because the same results hold regardless of the rates of particles' Poisson clocks, we assume clocks with mean 1 for simplicity.

Unlike in $\mathcal{M}$, the amoebot model assumes a particle P can perform at most one movement per activation (Sect. 1.2), so we must decouple P's movement in one iteration of $\mathcal{M}$, which includes both an expansion and a contraction, into two activations. However, due to asynchrony, other particles may expand into P's neighborhood after it has expanded but before it contracts. We utilize flag-locking mechanisms to ensure P retains consistent snapshots of its neighborhood regardless of the movements of other particles between its activations. When P expands from location ℓ to also occupy neighboring location ℓ' (Step 4 of $\mathcal{M}$), it sets a Boolean flag f to TRUE if it is the only expanded particle in its neighborhood, and to FALSE otherwise. When P is later activated again, it checks its flag: if f is FALSE, it simply contracts back to its original position ℓ since some other particle in its neighborhood activated and expanded earlier. Otherwise, P checks the conditions of Step 6 of $\mathcal{M}$ (ignoring any expanded heads, see the next paragraph) and decides whether to contract to ℓ or ℓ' accordingly. Particle P then resets f to FALSE and completes its second activation. This ensures that at most one particle per neighborhood moves at a time, mimicking the sequential nature of $\mathcal{M}$.

Some explanation is warranted on how particle P identifies expanded heads in its neighborhood and why it ignores them when checking the conditions of Step 6 of $\mathcal{M}$.

---

[8] Probability a_P only plays a role in the analysis of $\mathcal{A}$ and $\mathcal{M}$, not in their execution. Particle P does not need to know or calculate a_P.

Recall from Sect. 1.2 that a particle stores whether it is expanded or contracted and which neighboring locations are adjacent to its head in memory. Particle P can read this information from its neighbors to identify expanded heads in its neighborhood. Moreover, for particle P to reach Step 6 of $\mathcal{M}$, its flag f must be set to TRUE. Any other particle Q that expands into the neighborhood of P must then set its flag to FALSE, since it observes P is already expanded. Thus, P should ignore the heads of these expanded neighbors, since it is only a matter of time before they are activated again and simply contract their expanded heads.

We have shown our Markov chain $\mathcal{M}$ can be translated into a distributed, local, asynchronous algorithm $\mathcal{A}$, but such an implementation is not always possible in general. Any Markov chain for particle systems that relies on non-local particle moves or has transition probabilities that rely on non-local information cannot be executed by a local, distributed algorithm. Moreover, many algorithms under the amoebot model are not stochastic and thus cannot be meaningfully described as Markov chains; see, e.g. Daymude et al. (2017a), Derakhshandeh et al. (2017).

3.2 Properties of Markov chain $\mathcal{M}$

We now show some useful properties of Markov chain $\mathcal{M}$. Our first two claims follow from work in Cannon et al. (2016) and basic properties of Markov chains and our particle systems.

Lemma 2 *If σ_0 is connected and has no holes, then at every iteration of $\mathcal{M}$, the current configuration is connected and has no holes.*

Proof Cannon et al. (2016) proved that no moves allowed in their compression algorithm could introduce holes or disconnect the particle system. Since the moves allowed by $\mathcal{M}$ are a subset of those in the compression algorithm (since the local properties checked at each iteration are the same), $\mathcal{M}$ cannot introduce holes or disconnect the system. □

Lemma 3 *If σ_0 has no holes, then $\mathcal{M}$ is ergodic.*

Proof Markov chain $\mathcal{M}$ is irreducible because we defined Ω to be precisely those configurations reachable by valid transitions of $\mathcal{M}$ starting from σ_0. $\mathcal{M}$ is aperiodic because at each iteration there is a probability of at least $1/6$ that no move occurs, as each particle has at least one neighbor. Thus, the chain $\mathcal{M}$ is ergodic. □

As $\mathcal{M}$ is finite and ergodic, it converges to a unique stationary distribution, and we can find that distribution using detailed balance.

Lemma 4 *The stationary distribution of $\mathcal{M}$ is*

$$\pi(\sigma) = \lambda^{-p(\sigma)}\gamma^{-g(\sigma)}/Z,$$

where $Z = \sum_{\sigma' \in \Omega} \lambda^{-p(\sigma')}\gamma^{-g(\sigma')}$.

Proof Properties 1 and 2 ensure that particle P moving from location ℓ to location ℓ' is valid if and only if P moving from ℓ' to ℓ is. This implies for any configurations σ and τ, $Q(\sigma, \tau) > 0$ if and only if $Q(\tau, \sigma) > 0$. Using this, we easily verify the lemma via detailed balance.

Let $\sigma, \tau \in \Omega$ be distinct configurations that differ by one valid move of a particle P from location ℓ to neighboring location ℓ', and let n be the number of particles. Then,

$$Q(\sigma, \tau) = \frac{1}{n} \cdot \frac{1}{6} \cdot \min\{\lambda^{p(\sigma)-p(\tau)}\gamma^{g(\sigma)-g(\tau)}, 1\}, \text{ and}$$

$$Q(\tau, \sigma) = \frac{1}{n} \cdot \frac{1}{6} \cdot \min\{\lambda^{p(\tau)-p(\sigma)}\gamma^{g(\tau)-g(\sigma)}, 1\}.$$

Without loss of generality, assume that λ and γ satisfy $\lambda^{p(\sigma)-p(\tau)}\gamma^{g(\sigma)-g(\tau)} \leq 1$. Then,

$$\pi(\sigma)Q(\sigma, \tau) = \frac{\lambda^{-p(\sigma)}\gamma^{-g(\sigma)}}{Z} \cdot \frac{\lambda^{p(\sigma)-p(\tau)}\gamma^{g(\sigma)-g(\tau)}}{6n}$$

$$= \frac{\lambda^{-p(\tau)}\gamma^{-g(\tau)}}{Z} \cdot \frac{1}{6n} = \pi(\tau)Q(\tau, \sigma).$$

The definition of Z implies π satisfies $\sum_{\sigma' \in \Omega} \pi(\sigma') = 1$, so π is a valid probability distribution and we conclude π is the unique stationary distribution of $\mathcal{M}$. □

The stationary distribution can be alternately expressed using weighted perimeter.

Lemma 5 *For $c = 1 + \log_\lambda \gamma$, the stationary distribution of $\mathcal{M}$ is given by*

$$\pi(\sigma) = \lambda^{-\overline{p}(\sigma,c)}/Z,$$

where $Z = \sum_{\sigma' \in \Omega} \lambda^{-\overline{p}(\sigma',c)}$.

Proof This follows from the definition of $\overline{p}(\sigma, c)$. □

Theorem 1 *Consider an execution of Markov chain $\mathcal{M}$ on state space Ω, with $\lambda > 2 + \sqrt{2} =: \nu$ and $\gamma > 1$, where starting configuration σ_0 has n particles. For any constant α satisfying*

$$\alpha > \frac{\log \lambda}{\log \lambda - \log \nu} > 1,$$

the probability that a particle configuration σ drawn at random from $\mathcal{M}$'s stationary distribution π satisfies

$$\overline{p}(\sigma, 1 + \log_\lambda \gamma) > \alpha \cdot \overline{p}_{min}$$

is exponentially small in n for sufficiently large n, where $\overline{p}_{min}$ is the minimum weighted perimeter of a configuration in Ω.

Proof This proof mimics that of α-compression in Cannon et al. (2016), but additional insights and care are necessary

to accommodate the difficulties introduced by considering weighted perimeter instead of perimeter. Throughout we consider weighted perimeter $\overline{p}(\sigma) = \overline{p}(\sigma, 1 + \log_\lambda \gamma)$.

Define the weight of a configuration $\sigma \in \Omega$ to be:

$$w(\sigma) := \pi(\sigma) \cdot Z = \lambda^{-p(\sigma)} \gamma^{-g(\sigma)} = \lambda^{-\overline{p}(\sigma)},$$

where $Z = \sum_{\sigma' \in \Omega} \lambda^{-p(\sigma')} \gamma^{-g(\sigma')}$. For a set of configurations $S \subseteq \Omega$, we define its weight $w(S) = \sum_{\sigma \in S} w(\sigma)$; analogously, let $\pi(S) = \sum_{\sigma \in S} \pi(\sigma) = w(S)/Z$. Let $\sigma_{min} \in \Omega$ be a configuration with minimal weighted perimeter $\overline{p}_{min}$, and let S_α be the set of configurations with weighted perimeter at least $\alpha \cdot \overline{p}_{min}$. We show that for sufficiently large n,

$$\pi(S_\alpha) = \frac{w(S_\alpha)}{Z} < \frac{w(S_\alpha)}{w(\sigma_{min})} \leq \zeta^{\sqrt{n}},$$

where $\zeta < 1$. The first equality and inequality follow directly from the definitions of Z, w, and σ_{min}. We focus on the last inequality.

Stratify S_α into sets of configurations that have the same weighted perimeter; there are at most $O(n^2)$ such sets, as the total perimeter and gap perimeter can each take on at most $O(n)$ values. Label these sets as $A_1, A_2, \ldots, A_m$ in order of increasing weighted perimeter, where m is the total number of distinct weighted perimeters of configurations in S_α. Let $\overline{p}_i$ be the weighted perimeter of all configurations in set A_i; since $A_i \subseteq S_\alpha$, then $\overline{p}_i \geq \alpha \cdot \overline{p}_{min}$.

Note $w(\sigma) = \lambda^{-\overline{p}_i}$ for every $\sigma \in A_i$, so to bound $w(A_i)$ it suffices to bound $|A_i|$. A configuration with weighted perimeter $\overline{p}_i$ has perimeter $p \leq \overline{p}_i$, and a result from Cannon et al. (2016) that exploits a connection to self-avoiding walks in the hexagon lattice (Duminil-Copin and Smirnov 2012) implies the number of connected, hole-free particle configurations with perimeter p is at most $f(p) v^p$, for some subexponential function f. Letting p_{min} denote the minimum possible (unweighted) perimeter of a configuration of n particles, we conclude that:

$$w(A_i) = \lambda^{-\overline{p}_i} |A_i| \leq \lambda^{-\overline{p}_i} \cdot \sum_{p=p_{min}}^{\overline{p}_i} f(p) v^p \leq \lambda^{-\overline{p}_i} f_1(\overline{p}_i) v^{\overline{p}_i},$$

where $f_1(\overline{p}_i) = \sum_{p=p_{min}}^{\overline{p}_i} f(p)$ is necessarily also a subexponential function because it is a sum of at most a linear number of subexponential terms. So,

$$w(S_\alpha) = \sum_{i=1}^{m} w(A_i) \leq \sum_{i=1}^{m} f_1(\overline{p}_i) \left(\frac{v}{\lambda}\right)^{\overline{p}_i} \leq f_2(n) \left(\frac{v}{\lambda}\right)^{\alpha \cdot \overline{p}_{min}},$$

where $f_2(n) = \sum_{i=1}^{m} f_1(\overline{p}_i)$ is a subexponential function because $\overline{p}_i = O(n)$, $m = O(n^2)$, and f_1 is subexponential. The last inequality above holds as $\lambda > v$ and $\overline{p}_i \geq \alpha \cdot \overline{p}_{min}$. Then, since $w(\sigma_{min}) = \lambda^{-\overline{p}_{min}}$,

$$\pi(S_\alpha) < \frac{w(S_\alpha)}{w(\sigma_{min})} \leq f_2(n) \left(\frac{v}{\lambda}\right)^{\alpha \cdot \overline{p}_{min}} \lambda^{\overline{p}_{min}}$$
$$= f_2(n) \left[\lambda \left(\frac{v}{\lambda}\right)^\alpha\right]^{\overline{p}_{min}}.$$

The constant $\lambda(v/\lambda)^\alpha$ is less than one whenever $\alpha > \frac{\log \lambda}{\log \lambda - \log v}$. Since the perimeter of any configuration of n particles is at least $\sqrt{n}$, $\overline{p}_{min} \geq \sqrt{n}$. Because $f_2(n)$ is subexponentially large but $(\lambda(v/\lambda)^\alpha)^{\sqrt{n}}$ is exponentially small, asymptotically the latter term dominates and we conclude there exists $\zeta < 1$ such that for all sufficiently large n,

$$\pi(S_\alpha) < f_2(n)(\lambda(v/\lambda)^\alpha)^{\sqrt{n}} < \zeta^{\sqrt{n}},$$

which proves the theorem.   □

Though Theorem 1 is proved only in the case where the number of particles is sufficiently large, we expect and observe it to hold for much smaller n. However, we are unable to compute an explicit bound on how large n must be for these results to hold because the exact form of the subexponential function $f(p)$ in the above proof is unknown (see Section 4 of Duminil-Copin and Smirnov (2012) and references therein).

The following corollary shows that our algorithm solves any instance $(L, O, \sigma_0, c, \alpha)$ of the shortcut bridging problem when parameters λ and γ are chosen accordingly.

Corollary 1 *The distributed, local algorithm $\mathscr{A}$ associated with Markov chain $\mathscr{M}$ solves any valid instance of the shortcut bridging problem where the number of particles is sufficiently large.*

Proof Given any valid instance $(L, O, \sigma_0, c, \alpha)$ of the shortcut bridging problem, it suffices to run $\mathscr{A}$ starting from configuration σ_0 with parameters $\lambda > (2 + \sqrt{2})^{\frac{\alpha}{\alpha-1}}$ and $\gamma = \lambda^{c-1}$. Then $\alpha > \frac{\log(\lambda)}{\log(\lambda) - \log(2 + \sqrt{2})} > 1$, so by Theorem 1 the system reaches and remains with all but exponentially small probability in a set of configurations with weighted perimeter $\overline{p}(\sigma, c) \leq \alpha \cdot \overline{p}_{min}$, where $\overline{p}_{min}$ is the minimum weighted perimeter of a configuration in Ω. Solving the shortcut bridging problem only requires the weaker condition that this occurs with all but a polynomially small probability, which our algorithm certainly achieves.   □

4 Dependence on gap angle

To understand the relationship between bridging and shape, we consider V-shaped land masses of various angles (e.g., Fig. 2a). We prove our shortcut bridging algorithm has a dependence on the internal angle θ of the gap similar to that of the army ant bridges studied by Reid et al. (2015).
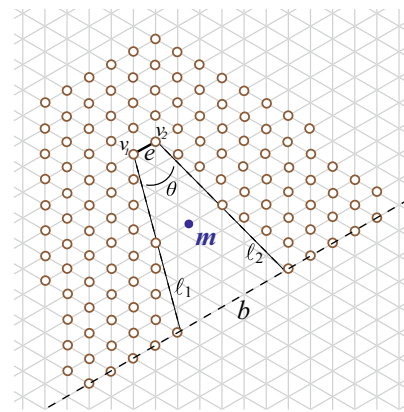
We show that when θ is sufficiently small, with all but exponentially small probability the bridge constructed by the particles stays close to the bottom of the gap (away from the apex of angle θ). On the other hand, we show that for some large values of θ, when λ and γ satisfy certain conditions, with all but exponentially small probability the bridge stays close to the top of the gap. We prove these results with a Peierls argument and careful analysis of the geometry of the gap. Simulations of our shortcut bridging algorithm for varying angles can be found in Sect. 5.

We first give a formal construction for the V-shaped land mass L given any $\theta \in (0, \pi)$ and constant width $w \geq 2$. Let $e \in E$ be any edge of the triangular lattice and label its endpoints as v_1 and v_2. Extend line segment ℓ_1 from v_1 such that it forms an angle of $\pi/2 + \theta/2$ with e. Similarly extend line segment ℓ_2 from v_2, of the same length and on the same side of e as ℓ_1, also forming an angle of $\pi/2 + \theta/2$ with e. Segments ℓ_1 and ℓ_2 then differ in their orientation by angle θ. Without loss of generality, we assume ℓ_1 is clockwise from ℓ_2 around e. Let b be the line through ℓ_1 and ℓ_2's other endpoints (not v_1 and v_2). The land mass consists of v_1, v_2, and all vertices of G_Δ that are outside of ℓ_1 and ℓ_2 and from which there exists a lattice path of length at most w to a vertex strictly between ℓ_1 and ℓ_2. Vertices of G_Δ on the opposite side of b from e are not included in the land mass. For example, Fig. 4a depicts a land mass with $\theta \sim \pi/6$ and Fig. 4b shows another with $\theta \sim \pi/2$; both have width $w = 5$. This careful definition involving edge e is necessary to ensure there are no adjacent land locations on opposite sides of the gap, as could happen for small θ if the land mass is not constructed carefully.
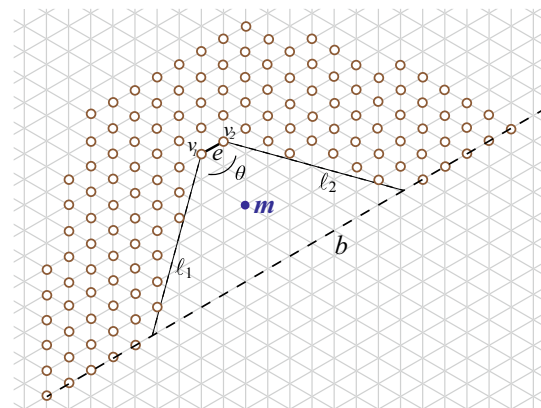
From now on we will, in a slight abuse of notation, refer to the gap locations between ℓ_1 and ℓ_2 as *the gap*. By the *bottom of the gap*, we mean the line b through ℓ_1 and ℓ_2's other endpoints (not v_1 and v_2). We may assume b is a line of the triangular lattice by truncating ℓ_1 and ℓ_2 so that both end on a lattice line; this does not change the land mass L. We also assume $b \cap \ell_1$ and $b \cap \ell_2$ are not vertices of the triangular lattice G_Δ; if they are, we can perturb ℓ_1 and ℓ_2 slightly, without changing the land mass. Note b is always parallel to e.

The *height* of land mass L is the length of a shortest path in G_Δ from v_1 or v_2 to b that only visits land locations; the land mass in Fig. 4a has height 8, while the land mass in Fig. 4b has height 9. Let m be the midpoint of the segment connecting the midpoints of ℓ_1 and ℓ_2; m is in the center of the gap, halfway between e and b.

The initial configuration σ_0 we consider is a path of width 2 lining the interior sides of the land mass L; see Fig. 5. We position the two fixed objects of O in line b at the second vertices outside ℓ_1 and ℓ_2, anchoring the particles on either side of the gap. Note the height of L is

(a)

(b)

Fig. 4 The land mass L of constant width 5 for **a** a small value of $\theta \sim \pi/6$ and height 8 and **b** a large value of $\theta \sim \pi/2$ and height 9. Point m is the midpoint of the segment between the midpoints of ℓ_1 and ℓ_2, and b is shown as a dashed line

exactly the number of particles in σ_0 next to ℓ_1 (or ℓ_2), excluding v_1 and v_2.

Lemma 6 *Let L be a V-shaped land mass of height k and angle θ. The initial configuration σ_0 has $4k + 5$ particles and two objects.*

Proof First, suppose $\theta \leq \pi/3$, as in Fig. 5a. Each lattice line parallel to e and intersecting ℓ_1 and ℓ_2, up to but not including b, contains exactly four particles. There are k such lattice lines. Line b contains two particles. In the lattice line above and parallel to e, there are three particles. In total, this gives $4k + 2 + 3 = 4k + 5$ particles and two objects.

Now, suppose $\theta > \pi/3$, as in Fig. 5b; a different counting approach is necessary. Consider the lattice line through v_1 and the gap location adjacent to v_1 and v_2; this line and all lines parallel to it intersecting ℓ_1 contain exactly two particles, and there are k such lines. The same is true for v_2 and ℓ_2. Uncounted by this approach are five
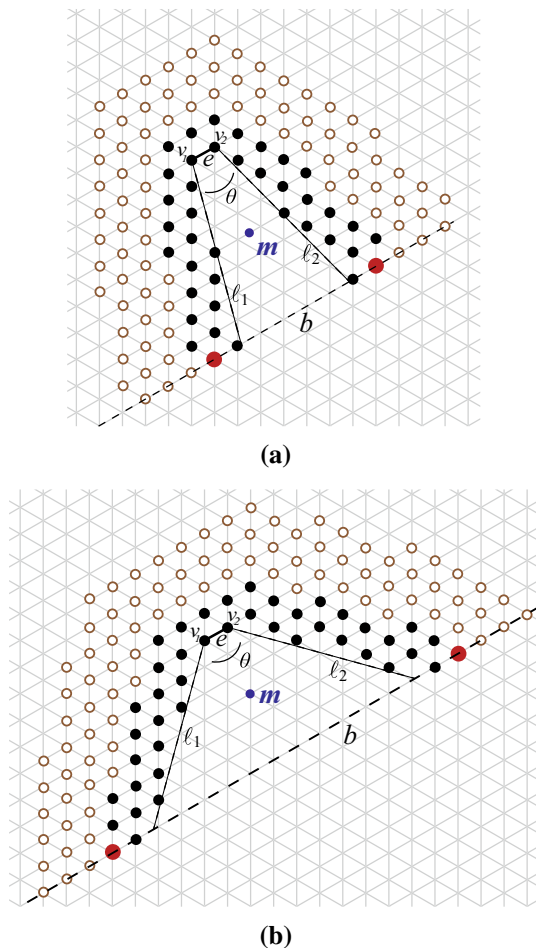
(a)

(b)

Fig. 5 The initial configuration σ_0, with particles shown in black and objects enlarged and red, for **a** a small value of $\theta \sim \pi/6$ and **b** a large value of $\theta \sim \pi/2$. Point m is the midpoint of the segment between the midpoints of ℓ_1 and ℓ_2, and b is shown as a dashed line. (Color figure online)

additional particles: the two particles adjacent to each of the two objects, and the particle adjacent to v_1 and v_2. In total, this gives $2k + 2k + 4 + 1 = 4k + 5$ particles and two objects. □

For a given σ, let x be the particle or object contained in line b farthest outside of ℓ_1, and let y be the particle or object in line b farthest outside of ℓ_2. We will refer to the perimeter of σ traversed counterclockwise from x to y as the *inner perimeter* of σ. We say the inner perimeter is *above a point p* if p is to the right of the inner perimeter traversed from x to y; it is *below a point p* if it is to its left.

We can partition Ω into two sets S_1 and S_2, where S_1 contains all configurations whose inner perimeter is strictly above midpoint m of the gap and S_2 contains all configurations whose inner perimeter goes through or below m. We first prove that for $\lambda > 2 + \sqrt{2}$ (i.e., in the range of compression) and $\gamma > 1$, there is an angle θ_1 such that for

all $\theta < \theta_1$, $\pi(S_1)$ is exponentially small. We then prove that for $\lambda > 2 + \sqrt{2}$ and $\gamma > \lambda^4(2 + \sqrt{2})^4$, there is a θ_2 such that for all $\theta \in (\pi/3, \theta_2)$, $\pi(S_2)$ is exponentially small. We expect much better bounds θ_1 and θ_2 can be obtained with more effort, and that these results generalize to all $\lambda > 2 + \sqrt{2}$ and $\gamma > 1$, but here we simply demonstrate it is possible to give rigorous results about the dependence of the bridge structure on θ.

4.1 Proofs for small θ

We begin with some structural lemmas.

Lemma 7 *Let L be a V-shaped land mass of height k and angle $\theta \leq \pi/3$. Then any path in G_Δ that starts and ends at the bottom of the gap and goes strictly above the midpoint m of the gap has length at least $k + 1$.*

Proof For $\theta \leq \pi/3$, there are $k - 1$ lattice lines parallel to b strictly between b and e. Of these lines exactly $\lceil (k - 1)/2 \rceil$ are below or contain m. Any path from b to a location above m and back to b must contain at least two vertices in each of these lattice lines, two vertices in b, and one vertex strictly above m, giving a total of

$$3 + 2\lceil (k - 1)/2 \rceil \geq 3 + 2((k - 1)/2) = k + 2$$

vertices. As the length of a path is the number of edges it contains, the path must have length at least $k + 1$. □

Lemma 8 *The i-th lattice line below and parallel to e contains $h(i)$ gap locations between ℓ_1 and ℓ_2, where*

$$i\sqrt{3} \tan \frac{\theta}{2} \leq h(i) \leq i\sqrt{3} \tan \frac{\theta}{2} + 2.$$

Proof Let b_i be the i-th lattice line below and parallel to e. We use trigonometry to analyze the length of b_i between ℓ_1 and ℓ_2; see Fig. 6a. Consider the triangle formed by b_i, ℓ_1, and the line perpendicular to e at v_1, which we call ℓ^*. Lines ℓ_1 and ℓ^* form an angle of $\theta/2$, and the distance between e and b_i along ℓ^* is $i\sqrt{3}/2$. It follows that the length of b_i between ℓ_1 and ℓ^* is $i\sqrt{3}\tan(\theta/2)/2$. Altogether, this implies b_i between ℓ_1 and ℓ_2 is of length $i\sqrt{3}\tan(\theta/2) + 1$. As each edge of the triangular lattice has length 1, this means there are between $i\sqrt{3}\tan(\theta/2)$ and $i\sqrt{3}\tan(\theta/2) + 2$ gap locations in b_i, as claimed. □

Lemma 9 *Let L be a V-shaped land mass of height k and angle $\theta \leq \pi/3$. Then the normalizing constant Z of the stationary distribution π of $\mathcal{M}$ satisfies*

$$Z \geq C\left[(\lambda\gamma)^{-2\sqrt{3}\tan\frac{\theta}{2}}\right]^k,$$

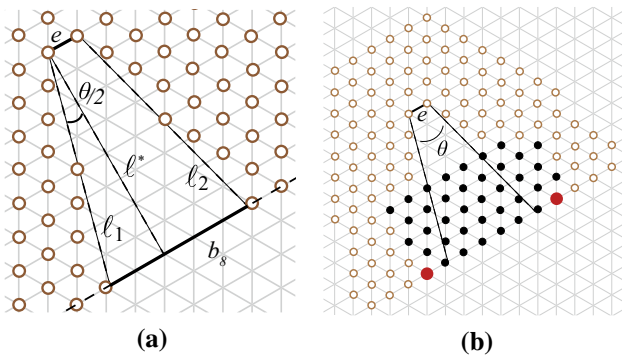*for a constant C that depends on θ, λ, and γ but not on k.*

Fig. 6 Figures from proofs in Sect. 4.1. **a** A depiction of the notation used in the proof of Lemma 8; the intersection of b_8 and the gap is depicted as a solid segment, which is of length $8\sqrt{3}\tan(\theta/2) + 1$ and contains 4 gap locations. **b** The configuration σ^* used in Lemma 9 for $\theta = \pi/6$ and $k = 8$

Proof Observe that $Z = \sum_{\sigma \in \Omega} \lambda^{-p(\sigma)} \gamma^{-g(\sigma)}$ satisfies $Z \geq \lambda^{-p(\sigma^*)} \gamma^{-g(\sigma^*)}$ for any $\sigma^* \in \Omega$. We now construct a particular σ^* (Fig. 6b) and calculate its perimeter and gap perimeter. Let σ^* contain a straight line of particles along b connecting the two objects, and let u be the number of objects and particles in this line. By Lemma 8, since $b = b_k$ and u includes two particles on land as well as two objects,

$$k\sqrt{3}\tan\frac{\theta}{2} + 4 \leq u \leq k\sqrt{3}\tan\frac{\theta}{2} + 6.$$

Continue constructing σ^* by placing rows of u particles above this initial row such that the row starts and ends on opposite sides of the gap. By Lemma 6, there are $4k + 7$ total objects and particles, so there will be $v = \lceil(4k + 7)/u\rceil$ such rows, with the last row possibly incomplete. We note that v satisfies:

$$v = \left\lceil\frac{4k+7}{u}\right\rceil \leq \frac{4k+7}{u} + 1 \leq \frac{4k+7}{k\sqrt{3}\tan\frac{\theta}{2} + 4} + 1$$

$$\leq \frac{4}{\sqrt{3}\tan\frac{\theta}{2}} + \frac{7}{4} + 1 \leq \frac{4}{\sqrt{3}\tan\frac{\theta}{2}} + 3;$$

$$v = \left\lceil\frac{4k+7}{u}\right\rceil \geq \frac{4k+7}{u} \geq \frac{4k+7}{k\sqrt{3}\tan\frac{\theta}{2} + 6}$$

$$\geq \frac{4k}{k\sqrt{3}\tan\frac{\theta}{2} + 6k} \geq \frac{4}{\sqrt{3}\tan\frac{\theta}{2} + 6}.$$

Configuration σ^* has perimeter at most $2u + 2v - 4$ and gap perimeter at most $u - 4 + z$, where z is the number of particles occupying gap locations in the upper perimeter of σ^*. These z remaining particles must be in either the $(k - v + 1)$-th or $(k - v + 2)$-th lattice lines below e, so we can bound z by again applying Lemma 8:

$$z \leq (k - v + 1)\sqrt{3}\tan\frac{\theta}{2} + 2.$$

Altogether, this implies:

$$p(\sigma^*) \leq 2u + 2v - 4$$

$$\leq 2k\sqrt{3}\tan\frac{\theta}{2} + 12 + \frac{8}{\sqrt{3}\tan\frac{\theta}{2}} + 6 - 4$$

$$\leq k\left(2\sqrt{3}\tan\frac{\theta}{2}\right) + \left(\frac{8}{\sqrt{3}\tan\frac{\theta}{2}} + 14\right),$$

and

$$g(\sigma^*) \leq u - 4 + z$$

$$\leq k\sqrt{3}\tan\frac{\theta}{2} + 6 - 4 + (k - v + 1)\sqrt{3}\tan\frac{\theta}{2} + 2$$

$$\leq 2k\sqrt{3}\tan\frac{\theta}{2} + \left(-\frac{4}{\sqrt{3}\tan\frac{\theta}{2} + 6} + 1\right)\sqrt{3}\tan\frac{\theta}{2} + 4$$

$$\leq k\left(2\sqrt{3}\tan\frac{\theta}{2}\right) + \left(\sqrt{3}\tan\frac{\theta}{2} - \frac{4\sqrt{3}\tan\frac{\theta}{2}}{\sqrt{3}\tan\frac{\theta}{2} + 6} + 4\right).$$

We note that the second parentheses in the final bounds above for $p(\sigma^*)$ and $g(\sigma^*)$ are constants that only depend on θ. This implies that there is a constant

$$C = \lambda^{-\left(14 + \frac{8}{\sqrt{3}\tan\frac{\theta}{2}}\right)} \gamma^{-\left(\sqrt{3}\tan\frac{\theta}{2} - \frac{4\sqrt{3}\tan\frac{\theta}{2}}{\sqrt{3}\tan\frac{\theta}{2} + 6} + 4\right)}$$

such that

$$Z \geq \lambda^{-p(\sigma^*)} \gamma^{-g(\sigma^*)} \geq C\left[(\lambda\gamma)^{-2\sqrt{3}\tan\frac{\theta}{2}}\right]^k.$$

As claimed, C depends only on λ, γ, and θ, and is independent of k. $\square$

Theorem 2 *Let $\lambda > 2 + \sqrt{2} =: v$ and $\gamma > 1$. Then there exists a constant θ_1 such that for all V-shaped land masses with angle $\theta < \theta_1$, the probability that the inner perimeter is above midpoint m is exponentially small in k, the height of the gap, provided k is sufficiently large. In particular,*

$$\theta_1 = 2\tan^{-1}\left(\frac{\log_{\lambda\gamma}(\lambda/v)}{\sqrt{3}}\right).$$

Proof Recall that $S_1 \subseteq \Omega$ is the set of configurations for which the inner perimeter is strictly above m. We show that S_1 has exponentially small weight at stationarity; in particular, we show $\pi(S_1)$ is bounded above by $f_2(k)\xi^k$, where $f_2(k)$ is a subexponential function and $\xi < 1$ is a constant.

If $\sigma \in S_1$, then by Lemma 7 we have $p(\sigma) \geq 2k + 2$, as its inner perimeter—and thus the rest of the perimeter as well—must be above m. Furthermore, because the perimeter by definition includes both objects and particles, which number $4k + 7$ by Lemma 6, any configuration $\sigma \in \Omega$ has $p(\sigma) \leq 2(4k + 7) - 2 = 8k + 12$. A result from Cannon et al. (2016) exploits a connection to self-avoiding walks in the hexagon lattice to show the number of connected, hole-free particle configurations with perimeter p is at most
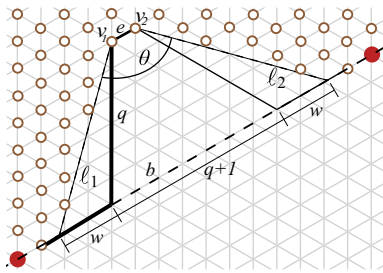
Fig. 7 The path of length k (bold) from vertex v_1 to the first land location in line b considered in the proof of Lemma 10; this path is used to calculate the gap height k in terms of the gap depth q. By also considering the reflection of this path from v_2 (solid line), we can calculate the distance between the two objects to be $q + 2\lceil w \rceil + 3$ (Lemma 11)

$f(p)(v)^p$ for some subexponential function f. This is certainly also an upper bound on the number of configurations in S_1 with perimeter p. Because $\gamma^{-g(\sigma)} < 1$, we have:

$$\pi(S_1) = \sum_{\sigma \in S_1} \frac{\lambda^{-p(\sigma)} \gamma^{-g(\sigma)}}{Z} < \sum_{p=2k+2}^{8k+12} \frac{f(p) v^p \lambda^{-p}}{Z}.$$

Let $f_1(k) = \sum_{p=2k+2}^{8k+12} f(p)$, and note that this function is subexponential in k because its number of summands is linear in k. Because $\lambda > v$ and $p \geq 2k + 2$, we have that:

$$\pi(S_1) \leq \frac{f_1(k) \left(\frac{v}{\lambda} \right)^{2k+2}}{Z}.$$

By Lemma 9, there is a constant $C_1 = v^2 / (\lambda^2 C)$ such that:

$$\pi(S_1) \leq \frac{f_1(k) \left(\frac{v}{\lambda} \right)^{2k+2}}{C \left[(\lambda \gamma)^{-2\sqrt{3}\tan\frac{\theta}{2}} \right]^k} = C_1 f_1(k) \left(\frac{v(\lambda \gamma)^{\sqrt{3}\tan\frac{\theta}{2}}}{\lambda} \right)^{2k}.$$

For all $\theta < 2\tan^{-1}\left(\log_{\lambda\gamma}(\lambda/v)/\sqrt{3} \right)$, the term in parentheses above is less than one:

$$\frac{v(\lambda\gamma)^{\sqrt{3}\tan\frac{\theta}{2}}}{\lambda} < \frac{v(\lambda\gamma)^{\log_{\lambda\gamma}\left(\frac{\lambda}{2+\sqrt{2}} \right)}}{\lambda} = 1.$$

Because $C_1 f_1(k)$ is a subexponential function but the term above, raised to the $2k$ power, is exponentially small, the latter eventually dominates and we conclude there is a constant $\xi < 1$ such that for sufficiently large k, $\pi(S_1) < \xi^k$, proving the theorem. □

Since $n = 4k + 5$ by Lemma 6, the probability that the inner perimeter is above point m is also exponentially small in n, the number of particles.

As an example, for $\lambda = 4$ and $\gamma = 2$ (the parameters of the simulations in Figs. 10 and 9), our methods give $\theta_1 = 0.0879 \sim 5.03°$. However, simulations suggest this bound is far from tight. In general, as λ increases, so does the angle θ_1: a stronger bias towards a shorter perimeter

means the bridge forms closer to the bottom of the gap and at even larger angles the bridge remains below m. Similarly, as γ decreases the bridge moves down towards the bottom of the gap and at even larger angles remains below m.

As with Theorem 1, we are unable to give explicit bounds on the "sufficiently large k" required by the statement of Theorem 2 because determining the exact form of the subexponential function $f(p)$ in the above proof remains an open problem (see Section 4 of Duminil-Copin and Smirnov (2012)). However, we expect and observe that the claims of this theorem hold even for the small k for which our proofs do not apply.

4.2 Proofs for large θ

We now consider the set $S_2 = \Omega \setminus S_1$, which consists of all configurations where the inner perimeter goes through or below m. We will show that for some large angles θ, for all $\lambda > 2 + \sqrt{2}$ and $\gamma > (2 + \sqrt{2})^4 \lambda^4$, $\pi(S_2)$ is exponentially small. While a lower bound on γ is necessary for the proofs presented below, we believe this is an artifact of our proof rather than the problem itself and suspect this requirement can be loosened or removed altogether.

For $\theta \geq \pi/3$, it is no longer true that a V-shaped land mass of height k has exactly $k - 1$ lattice lines between b and e. We define a new quantity q, the *gap depth*, as the length of a shortest path from e to b in G_Δ; unlike in the definition of the height k of a gap, this shortest path is not required to stay on land locations. The Euclidean distance between e and b is then $\sqrt{3}q/2$. Furthermore, q can be expressed as a function of k and θ.

Lemma 10 *For a V-shaped land mass of height k and angle $\theta \geq \pi/3$, the gap depth q satisfies*

$$k = \left\lceil \left(\frac{1}{2} + \frac{\sqrt{3}}{2} \tan\frac{\theta}{2} \right) q \right\rceil.$$

Proof Consider the path from v_1 to line b that leaves v_1 forming an angle of $2\pi/3$ with e, and then proceeds along b until it reaches a land location; see Fig. 7, where this path is shown in bold. The total length of this path is k, and its first segment from v_1 to b is length q. Let w be the length of b between this path's turning point and ℓ_1; then $k = q + \lceil w \rceil$. This path and ℓ_1 form an obtuse triangle where two sides have lengths q and w, respectively. The angle opposite the side of length w is $\theta/2 - \pi/6$, while the angle opposite the side of length q is $\pi - 2\pi/3 - (\theta/2 - \pi/6) = \pi/2 - \theta/2$. Length w can be calculated in terms of length q with the law of sines:
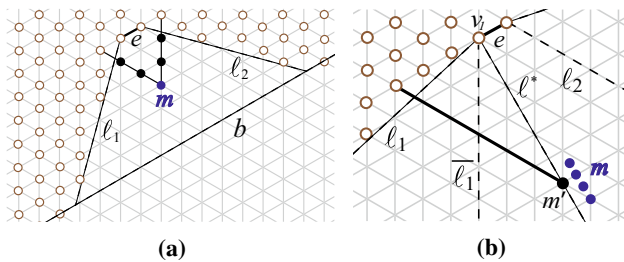
Fig. 8 From the proof of Lemma 12: **a** an example of a shortest path between land locations on opposite sides of the gap passing through midpoint m. **b** The four possible locations for midpoint m for which a shortest path passing through or below m contains m', and a shortest path from m' to a land location (solid line)

$$w = \frac{\sin\left(\frac{\theta}{2} - \frac{\pi}{6}\right)}{\sin\left(\frac{\pi}{2} - \frac{\theta}{2}\right)} \, q = \frac{\sin\frac{\theta}{2}\cos\frac{\pi}{6} - \cos\frac{\theta}{2}\sin\frac{\pi}{6}}{\cos\frac{\theta}{2}} \, q$$

$$= \frac{\frac{\sqrt{3}}{2}\sin\frac{\theta}{2} - \frac{1}{2}\cos\frac{\theta}{2}}{\cos\frac{\theta}{2}} \, q = \frac{q\sqrt{3}}{2}\tan\frac{\theta}{2} - \frac{q}{2}.$$

Because q is an integer, it follows that

$$k = q + \lceil w \rceil = \left\lceil q + \frac{q\sqrt{3}}{2}\tan\frac{\theta}{2} - \frac{q}{2} \right\rceil$$

$$= \left\lceil \left(\frac{1}{2} + \frac{\sqrt{3}}{2}\tan\frac{\theta}{2}\right)q \right\rceil,$$

which is the desired result. □

For simplicity, we do the bulk of our analysis using q instead of k. The previous lemma shows that proving an expression is exponentially small in q implies it is also exponentially small in k.

Lemma 11 *For any V-shaped land mass of gap depth q and angle $\theta \geq \pi/3$, any configuration σ has perimeter at least*

$$p(\sigma) \geq \left(2\sqrt{3}\tan\frac{\theta}{2}\right)q + 6.$$

Proof We first bound the distance between the two objects on either side of the gap. Using the length w from the proof of Lemma 10, the distance between the two objects in any configuration is $q + 2\lceil w \rceil + 3 \geq q + 2w + 3$ (see Fig. 7). The perimeter of any particle configuration is at least twice this distance, so for any σ,

$$p(\sigma) \geq 2q + 4w + 6 = 2q + 4\left(\frac{q\sqrt{3}}{2}\tan\frac{\theta}{2} - \frac{q}{2}\right) + 6$$

$$= \left(2\sqrt{3}\tan\frac{\theta}{2}\right)q + 6,$$

which is the desired bound. □

Lemma 12 *For any V-shaped land mass of gap depth q and angle $\theta > \pi/3$, any configuration $\sigma \in S_2$ (passing below or through midpoint m of the gap) has gap perimeter $g(\sigma) \geq \frac{q}{2}$.*

Proof If $\sigma \in S_2$, i.e., if its inner perimeter passes through or below m, then it must contain a path that starts and ends at land locations and also passes through or below m. We consider all such paths and give a lower bound on the number of gap locations they must contain. The shortest such paths start and end on opposite sides of the gap, so we focus on paths of this type.

If m is a vertex of G_Δ, one shortest path between land locations passing through m leaves m along the two lattice lines not parallel to e and follows them until reaching the land mass, as in Fig. 8a. If m is on a lattice edge, a shortest path passing below m is constructed in the same way, beginning from each of the edge's endpoints. Otherwise, if m is neither a lattice point nor on a lattice edge, the same procedure is followed for the first lattice point or lattice edge below m. In all cases, let m' be the point of intersection between this path and ℓ^*, the line perpendicular to e through v_1. Fig. 8b shows all the possible locations of m producing a particular m'. Inspection shows that in all of these cases, m' is contained in the $2\lfloor\frac{q+1}{4}\rfloor$-th lattice line below e.

Let $\overline{\ell_1}$ be the line from v_1 to b forming an angle of $2\pi/3$ with e; see Fig. 8b. Because $\theta > \pi/3$, all vertices of G_Δ contained in $\overline{\ell_1}$, except v_1, are gap locations. Any shortest path from m' to a land location must share a vertex of G_Δ with line $\overline{\ell_1}$. Because m' is in the $2\lfloor\frac{q+1}{4}\rfloor$-th lattice line below e, any path from m' to $\overline{\ell_1}$ is of length at least $\lfloor\frac{q+1}{4}\rfloor$ and contains at least $\lfloor\frac{q+1}{4}\rfloor + 1$ gap locations, including both of its endpoints. By symmetry, this means any path between land locations passing below m, and thus any inner perimeter of a particle configuration passing below m, contains at least

$$2\left(\left\lfloor\frac{q+1}{4}\right\rfloor + 1\right) \geq 2\left(\frac{q-2}{4} + 1\right) \geq \frac{q}{2}$$

gap locations, as claimed. □

Theorem 3 *Let $\lambda > 2 + \sqrt{2} =: \nu$ and $\gamma > (\lambda\nu)^4$. Then there exists a constant $\theta_2 > \pi/3$ such that for all V-shaped land masses with angle $\theta \in (\pi/3, \theta_2)$, the probability that the inner perimeter goes through or below midpoint m is exponentially small in k, the height of the gap, provided k is sufficiently large.*

Proof Recall S_2 is the set of all configurations whose inner perimeter goes through or below m. We show that $\pi(S_2)$ is exponentially small in k, the height of the gap. By definition,

$$\pi(S_2) = \frac{\sum_{\sigma \in S_2} \lambda^{-p(\sigma)} \gamma^{-g(\sigma)}}{Z}.$$

By Lemma 6, the number of particles and objects in σ_0 for a land mass of height k is $4k + 7$. Since σ_0 is a path of width 2 and every particle occupies a land location, $p(\sigma_0) = 4k + 7$ and $g(\sigma_0) = 0$. Thus,

$$Z = \sum_{\sigma \in \Omega} \lambda^{-p(\sigma)} \gamma^{-g(\sigma)} \geq \lambda^{-p(\sigma_0)} \gamma^{-g(\sigma_0)} = \lambda^{-4k-7}.$$

It is simpler to work with gap depth q instead of gap height k. By Lemma 10, k satisfies $k \leq \left(\frac{1}{2} + \frac{\sqrt{3}}{2} \tan \frac{\theta}{2}\right) q + 1$, so

$$Z \geq \lambda^{-4k-7} \geq \lambda^{-4\left(\frac{1}{2} + \frac{\sqrt{3}}{2} \tan \frac{\theta}{2}\right)q - 4 - 7} = \lambda^{-\left(2 + 2\sqrt{3} \tan \frac{\theta}{2}\right)q - 11}.$$

Combining this with Lemma 12,

$$\pi(S_2) = \sum_{\sigma \in S_2} \frac{\lambda^{-p(\sigma)} \gamma^{-g(\sigma)}}{Z}$$

$$\leq \lambda^{\left(2 + 2\sqrt{3} \tan \frac{\theta}{2}\right)q + 11} \sum_{\sigma \in S_2} \lambda^{-p(\sigma)} \gamma^{-\frac{q}{2}}.$$

Let p_{min} (resp., p_{max}) be the minimum (resp., maximum) possible perimeter for a valid particle configuration in S_2. By Lemma 11, $p_{min} \geq 2\sqrt{3} \tan(\theta/2) q$. As shown in the proof of Theorem 2, $p_{max} = 8k + 12$; in terms of q, by Lemma 10,

$$p_{max} \leq 8\left(\frac{q}{2} + \frac{q\sqrt{3}}{2} \tan \frac{\theta}{2} + 1\right) + 12$$

$$= 4q + 4q\sqrt{3} \tan \frac{\theta}{2} + 20.$$

Using the result from Cannon et al. (2016) which upper bounds the number of particle configurations with perimeter p by the expression $f(p)v^p$, for some subexponential function f, we have that:

$$\pi(S_2) \leq \lambda^{\left(2 + 2\sqrt{3} \tan \frac{\theta}{2}\right)q + 11} \sum_{p = p_{min}}^{p_{max}} f(p) v^p \lambda^{-p} \gamma^{-\frac{q}{2}}$$

$$\leq \lambda^{\left(2 + 2\sqrt{3} \tan \frac{\theta}{2}\right)q + 11} \left(\sum_{p = p_{min}}^{p_{max}} f(p)\right) \left(\frac{v}{\lambda}\right)^{p_{min}} \gamma^{-\frac{q}{2}}$$

$$\leq \left(\lambda^{11} \sum_{p = p_{min}}^{p_{max}} f(p)\right) \cdot \left(\lambda^{\left(2 + 2\sqrt{3} \tan \frac{\theta}{2}\right)} \left(\frac{v}{\lambda}\right)^{2\sqrt{3} \tan \frac{\theta}{2}} \gamma^{-\frac{1}{2}}\right)^q$$

$$= \left(\lambda^{11} \sum_{p = p_{min}}^{p_{max}} f(p)\right) \left(\lambda^2 v^{2\sqrt{3} \tan \frac{\theta}{2}} \gamma^{-\frac{1}{2}}\right)^q.$$

The first parentheses is a function $f_1(q)$ that is subexponential in q, as it has a polynomial number of summands based on our calculations of p_{min} and p_{max} (which are expressions in terms of q), and each summand is subexponential. When the term in the second set of parentheses

above is less than one, the second factor (this term raised to the q power) is exponentially small in q, the gap depth, and thus for sufficiently large q this term dominates and the entire expression is exponentially small in q. This holds whenever θ satisfies:

$$\theta < 2 \tan^{-1}\left(\frac{1}{2\sqrt{3}} \log_v\left(\gamma^{1/2} \lambda^{-2}\right)\right)$$

$$= 2 \tan^{-1}\left(\frac{1}{\sqrt{3}} \log_v\left(\frac{\gamma^{1/4}}{\lambda}\right)\right) =: \theta_2.$$

Whenever $\gamma^{1/4}/\lambda > v$—i.e., whenever $\gamma > (\lambda v)^4$—the argument of $\tan^{-1}$ above is at least $1/\sqrt{3}$, and thus $\theta_2 > \pi/3$. It follows that whenever $\gamma > (\lambda v)^4$ and $\theta \in (\pi/3, \theta_2)$,

$$\pi(S_2) < f_1(q) \psi^q,$$

where $f_1(q)$ is subexponentially large in q and $\psi < 1$ so the second term is exponentially small in q. For sufficiently large q, the second term dominates, and we conclude the weight of set S_2 at stationarity is exponentially small in q. Because k and q differ only by additive and multiplicative constants, it is also exponentially small in k, the gap height, for sufficiently large k. $\square$

As was the case for small angles, here also we have that by Lemma 6, there are $n = 4k + 5$ particles. Thus, we have that the probability the inner perimeter goes through or below midpoint m when θ is sufficiently large is also exponentially small in n.

If we again use the example value of $\lambda = 4$ (as in the simulations depicted in Figs. 9 and 10), Theorem 3 requires $\gamma > (\lambda v)^4 \approx 34,786$. This value is large, but importantly is constant (i.e., it does not depend on n) and is only an artifact of our proof. For example, when $\lambda = 4$ and $\gamma = 10^5$, our methods show that the resulting bridge remains above midpoint m with high probability for any angle between $\pi/3 = 60°$ and $\theta_2 \approx 1.2234 \sim 70.10°$. On the other hand, an experiment with $\lambda = 4$, $\gamma = 2$, and $\theta = 90°$ is shown in Fig. 10c to remain well above the midpoint m, suggesting that this behavior is stable for much smaller values of γ and a much larger range of angles than we were able to prove.

As for Theorems 1 and 2, we are unable to give explicit bounds on the "sufficiently large k" required by the statement of Theorem 3 because the exact form of $f(p)$ in its proof is unknown, but we expect and observe that it holds even for the small k for which our proof does not apply.

5 Simulations

We can see the performance of our algorithm from simulation results on a variety of instances. Figure 9 shows snapshots over time for a bridge shortcutting a V-shaped gap with internal angle $\theta = \pi/3$ and biases $\lambda = 4, \gamma = 2$. Qualitatively, this bridge matches the shape and position of the army ant bridges in Reid et al. (2015). Figure 10 shows the results of an experiment that held λ, γ, and the number of iterations of $\mathcal{M}$ constant, varying only the internal angle of the V-shaped land mass. The particle system exhibits behavior consistent with the theoretical results in Sect. 4 and the army ant bridges in Reid et al. (2015), shortcutting closer to the bottom of the gap when θ is small and staying almost entirely on land when θ is large. Lastly, Fig. 11 shows the resulting bridge structure when the land mass is N-shaped, demonstrating that our algorithm can be generalized beyond the original inspiration of V-shaped land masses to shortcut multiple gaps in more complex structures.

These simulations demonstrate the successful application of our stochastic approach to shortcut bridging. Moreover, experimenting with variants suggests this approach may be useful for other related applications in the future.

6 Conclusions and future directions

In summary, we presented a Markov chain $\mathcal{M}$ that can be directly translated to a stochastic, distributed, local, asynchronous algorithm $\mathcal{A}$ that provably solves the shortcut bridging problem. Furthermore, in the special case of
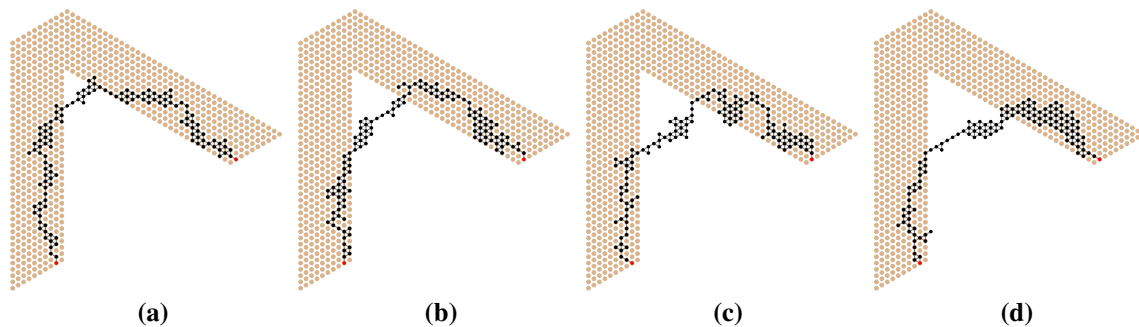

Fig. 9 A particle system using biases $\lambda = 4$ and $\gamma = 2$ to shortcut a V-shaped land mass with $\theta = \pi/3$ after **a** 2 million, **b** 4 million, **c** 6 million, and **d** 8 million iterations of Markov chain $\mathcal{M}$, beginning in configuration σ_0 shown in Fig. 2a
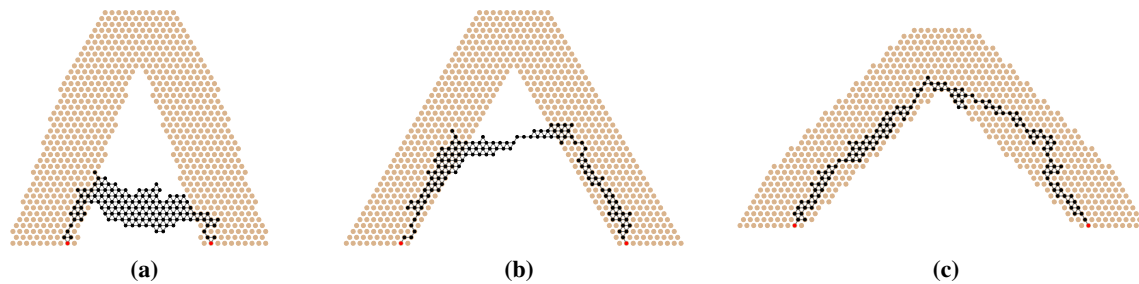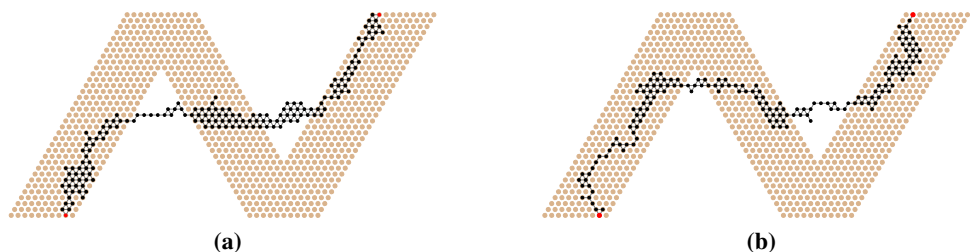

Fig. 10 A particle system using biases $\lambda = 4$ and $\gamma = 2$ to shortcut a V-shaped land mass with angle **a** $\pi/6$, **b** $\pi/3$, and **c** $\pi/2$ after 20 million iterations of Markov chain $\mathcal{M}$. For a given angle, the land mass L and initial configuration σ_0 were constructed as described in Sect. 4

Fig. 11 A particle system using biases $\lambda = 4$ and $\gamma = 2$ to shortcut an N-shaped land mass after **a** 10 million and **b** 20 million iterations of Markov chain $\mathcal{M}$, beginning in configuration σ_0 shown Fig. 2b

bridging over the gap in a V-shaped land mass, we rigorously analyzed the effect of the land mass's internal angle, showing that below one threshold angle the bridge will shortcut near the bottom of the gap, and above another threshold angle the bridge will remain close to land, with high probability.

Several directions of further investigation seem promising. The successful application of our stochastic approach to shortcut bridging suggests it may be useful for other types of problems as well; one related behavior of particular interest is "exploration bridging", where a particle system first explores its environment to discover sites of interest, and then converges to a bridge-like structure between them. We are also interested in formulating alternative local rules for shortcut bridging which yield bridges that appear more "structurally sound," though we suspect the information needed to do so may be difficult to encode in our particle systems due to the constant-size memory constraint of the amoebot model.

References

Andrés Arroyo M, Cannon S, Daymude JJ, Randall D, Richa AW (2017) A stochastic approach to shortcut bridging in programmable matter. In: DNA computing and molecular programming, DNA23, pp 122–138

Angluin D, Aspnes J, Diamadi Z, Fischer MJ, Peralta R (2006) Computation in networks of passively mobile finite-state sensors. Distrib Comput 18(4):235–253

Baxter RJ, Enting IG, Tsang SK (1980) Hard-square lattice gas. J Stat Phys 22:465–489

Blanca A, Chen Y, Galvin D, Randall D, Tetali P (2018) Phase coexistence for the hard-core model on $\mathbb{Z}^2$. Comb Prob Comput. https://doi.org/10.1017/S0963548318000238

Camazine S, Visscher KP, Finley J, Vetter SR (1999) House-hunting by honey bee swarms: collective decisions and individual behaviors. Insectes Sociaux 46(4):348–360

Cannon S, Daymude JJ, Randall D, Richa AW (2016) A Markov chain algorithm for compression in self-organizing particle systems. In: Proceedings of the 2016 ACM symposium on principles of distributed computing, PODC'16, pp 279–288

Chazelle B (2009) Natural algorithms. In: Proceedings of the 20th ACM–SIAM symposium on discrete algorithms, SODA'09, pp 422–431

Chen M, Xin D, Woods D (2015) Parallel computation using active self-assembly. Nat Comput 14(2):225–250

Cheung KC, Demaine ED, Bachrach JR, Griffith S (2011) Programmable assembly with universally foldable strings (moteins). IEEE Trans Robot 27(4):718–729

Chirikjian GS (1994) Kinematics of a metamorphic robotic system. In: Proceedings of the 1994 IEEE international conference on robotics and automation, ICRA'94, vol 1, pp 449–455

Cieliebak M, Flocchini P, Prencipe G, Santoro N (2012) Distributed computing by mobile robots: gathering. SIAM J Comput 41(4):829–879

Daymude JJ, Gmyr R, Richa AW, Scheideler C, Strothmann T (2017a) Improved leader election for self-organizing programmable matter. In: Algorithms for sensor systems, ALGO-SENSORS'17, pp 127–140

Daymude JJ, Richa AW, Scheideler C (2017b) The amoebot model. https://sops.engineering.asu.edu/sops/amoebot. Accessed 16 Sept 2018.

Derakhshandeh Z, Dolev S, Gmyr R, Richa AW, Scheideler C, Strothmann T (2014) Brief announcement: amoebot—a new model for programmable matter. In: Proceedings of the 26th ACM symposium on parallelism in algorithms and architectures, SPAA'14, pp 220–222

Derakhshandeh Z, Gmyr R, Richa AW, Scheideler C, Strothmann T (2017) Universal coating for programmable matter. Theor Comput Sci 671:56–68

Douglas S, Dietz H, Liedl T, Högberg B, Graf F, Shih W (2009) Self-assembly of DNA into nanoscale three-dimensional shapes. Nature 459:414–418

Duminil-Copin H, Smirnov S (2012) The connective constant of the honeycomb lattice equals $\sqrt{2 + \sqrt{2}}$. Ann Math 275(3):1653–1665

Feller W (1968) An introduction to probability theory and its applications, vol 1. Wiley, New York

Flocchini P, Prencipe G, Santoro N, Widmayer P (2008) Arbitrary pattern formation by asynchronous, anonymous, oblivious robots. Theor Comput Sci 407(1):412–447

Hastings WK (1970) Monte carlo sampling methods using Markov chains and their applications. Biometrika 57(1):97–109

Jeanson R, Rivault C, Deneubourg J, Blanco S, Fournier R, Jost C, Theraulaz G (2005) Self-organized aggregation in cockroaches. Animal Behav 69(1):169–180

Levin DA, Peres Y, Wilmer EL (2009) Markov chains and mixing times. American Mathematical Society, Providence

Lynch N (1996) Distributed algorithms. Morgan Kauffman, Burlington

Mohammed A, Šulc P, Zenk J, Schulman R (2017) Self-assembling DNA nanotubes to connect molecular landmarks. Nat Nanotechnol 12:312–316

Reid CR, Latty T (2016) Collective behaviour and swarm intelligence in slime moulds. FEMS Microbiol Rev 40(6):798–806

Reid CR, Lutz MJ, Powell S, Kao AB, Couzin ID, Garnier S (2015) Army ants dynamically adjust living bridges in response to a cost-benefit trade-off. Proc Natl Acad Sci 112(49):15113–15118

Restrepo R, Shin J, Tetali P, Vigoda E, Yang L (2013) Improving mixing conditions on the grid for counting and sampling independent sets. Prob Theory Relat Fields 156:75–99

Rubenstein M, Cornejo A, Nagpal R (2014) Programmable self-assembly in a thousand-robot swarm. Science 345(6198):795–799

Walter JE, Brooks ME, Little DF, Amato NM (2004a) Enveloping multi-pocket obstacles with hexagonal metamorphic robots. In: Proceedings of the 2004 IEEE international conference on robotics and automation, ICRA'04, vol 3, pp 2204–2209

Walter JE, Welch JL, Amato NM (2004b) Distributed reconfiguration of metamorphic robot chains. Distrib Comput 17(2):171–189

Wei B, Dai M, Yin P (2012) Complex shapes self-assembled from single-stranded DNA tiles. Nature 485:623–626

Woods D (2015) Intrinsic universality and the computational power of self-assembly. Philos Trans R Soc A 373(2046):20140214

Woods D, Chen HL, Goodfriend S, Dabby N, Winfree E, Yin P (2013) Active self-assembly of algorithmic shapes and patterns in polylogarithmic time. In: Proceedings of the 4th innovations in theoretical computer science conference, ITCS'13, pp 353–354

Yim M, Shen WM, Salemi B, Rus D, Moll M, Lipson H, Klavins E, Chirikjian GS (2007) Modular self-reconfigurable robot systems. IEEE Robot Autom Mag 14(1):43–52