Spherical Cubes:
Optimal Foams from Computational Hardness Amplification

Guy Kindler[1], Ryan O'Donnell[2], Anup Rao[3] & Avi Wigderson[4]

Abstract. *Foam* problems are concerned with how one can partition space into "bubbles" which minimize surface area. We investigate the case where one unit-volume bubble is required to tile *d*-dimensional space in a periodic fashion according to the standard, cubical lattice. While a *cube* requires surface area 2*d*, we construct such a bubble having surface area very close to that of a sphere; i.e., proportional to $\sqrt{d}$ (the minimum possible even without the periodicity constraint). Our method for constructing this "spherical cube" has a surprising inspiration: foundational questions in the theory of computation—specifically the issue of "hardness amplification." We additionally show an algorithmic application of our new foam: a method for "coordinated discretization" of high-dimensional data points which has near-optimal resistance to noise. Finally, we provide the most efficient known cubical foam in 3 dimensions.

[1] *School of Computer Science and Engineering, Hebrew University of Jerusalem, Givat Ram Campus, 91904 Jerusalem, Israel.*

[2] *Computer Science Department, School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh PA 15213.*

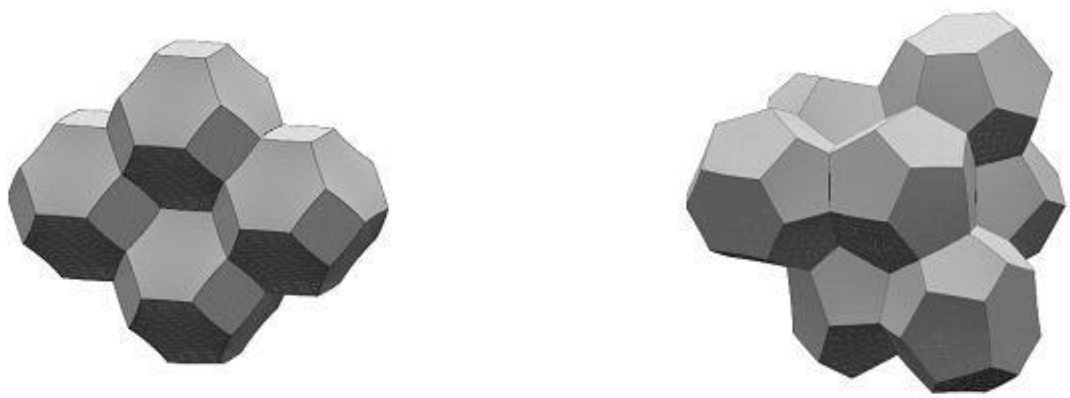
[3] *Department of Computer Science and Engineering, University of Washington, Box 352350, Seattle WA 98159.*

[4] *School of Mathematics, Institute for Advanced Study, Einstein Dr., Princeton NJ 08540.*

How can space be tiled by shapes of a given volume so that the average surface area is minimized? This fundamental question has been a focus of study for scientists in many disciplines, from physicists studying soap bubbles (*1*), to chemists studying crystal structures (*2*), biologists studying cell aggregation (*3*), mathematicians studying sphere-packings (*4*), materials scientists studying polymers (*5*), and even artists and architects (*6*). In this work we present a new approach to the construction of tiling shapes, based on methods from computer science. This approach leads to an asymptotically optimal solution of the Cubical Foam Problem, defined below.

Foams. Questions about minimal surface area tilings of space have a very long history. In the 19th century Thomson (Lord Kelvin) introduced the *Kelvin Foam Problem* (*7*), which asks how 3-dimensional space can be partitioned into bubbles of volume 1 such that the average surface area of the bubbles in the foam is minimized. This question (which turns out to be extremely difficult) is motivated not only by its mathematical appeal, but also by interest in the physics of foams in nature, since surface tension makes bubbles seek to minimize their surface area.

One of the very few known ways of designing foams with small surface area is to first construct a *lattice* of periodically arranged points, and then to take the *Voronoi cells* around each lattice point. The Voronoi cell of a lattice point *x* is defined to be the bubble which includes all points which are closer to *x* than to any other lattice point. The solution Kelvin proposed in 1887 for his problem was based on the Voronoi foam associated to the *body-centered cubic* lattice. The bubbles in this foam have surface area $\tfrac{3}{4}\sqrt[3]{120\sqrt{3}+148} \approx 5.315$. Kelvin further suggested letting this foam "relax" so that it conforms with Plateau's Rules for soap bubbles (*1*); modern computer simulations (*8*) show that this decreases the surface area to about 5.306 (*9*). Kelvin's foam was widely believed to be optimal until 1994, when Weaire and Phelan (*10*) exhibited a foam with an improved average surface area of about 5.288. The Weaire–Phelan foam is also formed by relaxing the Voronoi foam for a certain periodic arrangement of points in space (a subset of a lattice). Weaire and Phelan based their arrangement of points on the crystal structure of a certain silicon-sodium clathrate; whether or not their foam optimally solves Kelvin's problem is still an open question.



A **B**

Fig. 1. (A) Four bubbles in the Kelvin Foam, formed by relaxing the Voronoi cells of the body-centered cubic lattice. **(B)** Seven bubbles in the Weaire–Phelan Foam, formed by relaxing the Voronoi cells of the *A15 Packing*.

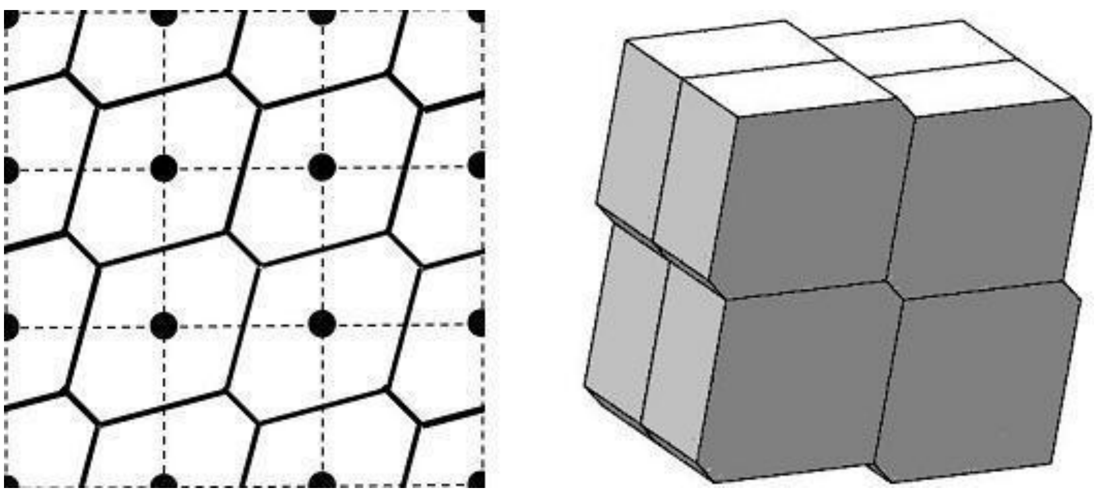
As with the closely related question of sphere-packing (i.e., how efficiently equal-sized spheres can be packed in space), it is natural to study the Kelvin Foam Problem in dimensions other than 3. In 2 dimensions, it was long believed that the best solution is to tile space with regular hexagons—the "Voronoi foam" of the triangular lattice. Although Pappus of Alexandria claimed this solution was optimal in the 4th century, a mathematical proof was found only in 1999, by Hales (*11*). In higher dimensions, a lower bound on average surface area follows from the Isoperimetric Inequality: the surface area of any bubble of volume 1 must be at least as large as that of a ball of volume 1. As the number of dimensions d grows, this lower bound asymptotically approaches $\sqrt{2\pi e}\sqrt{d}$. An upper bound which matches this lower bound up to a factor of 2 can be obtained by taking the Voronoi foam of a d-dimensional lattice in which the covering-radius to packing-radius ratio tends to 2—such a lattice can be obtained by a probabilistic construction (*12*). Hence the minimum surface area in the d-dimensional Kelvin Foam Problem grows in proportion to the *square-root* of the dimension.

In our work we consider tilings that are periodic with respect to a specific and very natural lattice, the *integer lattice* (also known as the *cubic lattice*). This lattice consists of the points in d-dimensional space whose coordinates are all integers. We address the following question:

Cubical Foam Problem: *What is the least surface area of a bubble that partitions d-dimensional space periodically according to the integer lattice?*

The Voronoi foam for the integer lattice consists of cubes of side-length 1. In d dimensions, these cubes have surface area $2d$. This grows *linearly* with the dimension, much higher than the known lower bound of $\sqrt{d}$. Are there more “spherical” cubes, which still tile by the integer lattice but have surface area closer to that of a ball? Prior to our work, there was no general conjecture as to whether integer-lattice foams require surface area proportional to the dimension, to the square root of the dimension, or to something in between.

The Cubical Foam Problem seems to have been first formally raised by Choe in 1989 (*13*). Choe showed that in 2 dimensions, the unit square (whose "surface area"—i.e., perimeter—is 4) is not the optimal solution. Rather, the optimal solution is the "isosceles" hexagon shown in Fig. 2A, with 120° angles, side lengths $\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{6}}$ and $\sqrt{\frac{2}{3}}$, and perimeter about 3.864. Choe gave the 3-dimensional version as an open problem. No solution has even been conjectured, and prior to our work the best known solution was simply to add depth to the Choe hexagon, transforming it into the prism shown in Fig. 2B with surface area 5.864 (*14*).



A B

Fig. 2. (A) The Choe Hexagon, Choe's optimal solution to the Cubical Foam Problem in 2 dimensions. **(B)** The hexagons extruded into 3-dimensional prisms. The resulting 3-dimensional cubical foam is not optimal; a solution with smaller average surface area is presented in this work.

The high-dimensional version of the Cubical Foam Problem was raised by Feige, Kindler, and O'Donnell (*14*) in 2007, who noted a surprising connection to a certain problem in theoretical computer science about *computational hardness amplification*. This connection is explained later in this article. A subsequent result of Raz (*15*) on the limits of such amplification, using an idea from a related paper of Holenstein (*16*), provided us with the tools to solve the high-dimensional Cubical Foam Problem.

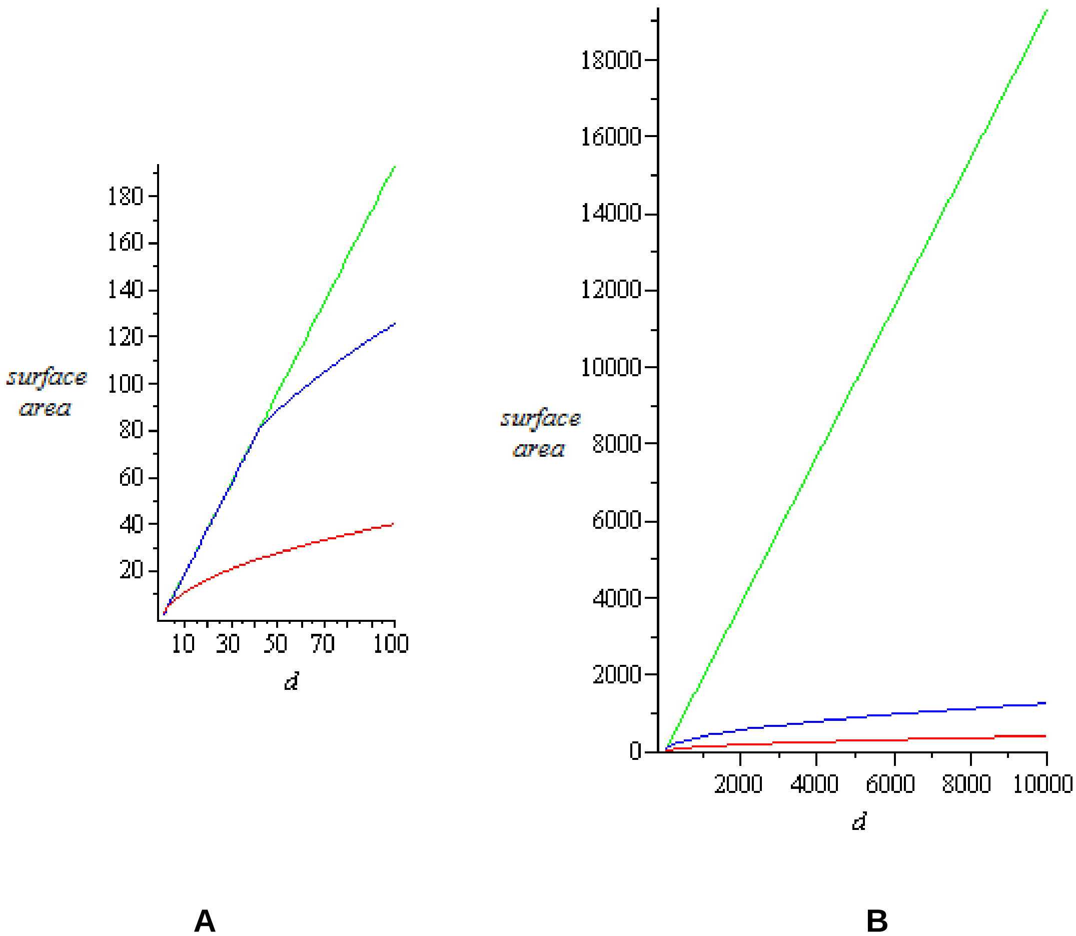


Fig. 3. Cubical Foam solutions, surface area vs. dimension *d*. The green plot shows the previous best construction, Choe prisms. The red plot shows the theoretical lower bound given by the sphere's surface area. The blue plot shows the results achieved in this paper. **(A)** *d* = 1…100. **(B)** *d* = 1…10,000.

Our results.

- We give a probabilistic construction proving the existence of a bubble that partitions *d*-dimensional space according to the cubic lattice and whose surface area is at most $4\pi\sqrt{d}$. Thus our bubble is nearly "spherical", in the sense that its surface area is larger than that of a sphere by only a constant multiplicative factor (about 3.04). This is in contrast to best previous constructions with "cube-like" surface area proportional to *d*. We conclude that the optimal solution to the Cubical Foam Problem has surface area proportional to the square-root of the dimension, just as in the more general Kelvin Foam Problem. Thus in high dimensions, integer-lattice tilings are essentially as efficient as *any* tilings.

- We show that our construction also yields a highly noise-resistant procedure for the "coordinated discretization" of data. Specifically, the construction can be viewed as a randomized procedure which assigns each vector of real numbers $(x_1,\ldots,x_d)$ to a vector of integers $(r_1,\ldots,r_d)$, with the following two guarantees. The "closeness" guarantee is that each r_i is always simply x_i rounded up or down to the nearest integer. The "noise-resistance" guarantee is that if two real vectors $(x_1,\ldots,x_d)$ and $(y_1,\ldots,y_d)$ are at Euclidean distance ε in $\mathbf{R}^d$, then our procedure assigns them to the *same* integer vector except with probability proportional to ε (more precisely, probability at most $2\pi\cdot\varepsilon$). Somewhat remarkably, the noise-resistance guarantee is independent of *d*, the dimension of the vectors. Previously known coordinated discretization procedures had either a worse "closeness" guarantee, with $|x_i - r_i|$ possibly as large as $\sqrt{d}$, or a worse "noise-resistance" guarantee (*17, 18*), with "miscoordination" probability proportional to $\sqrt{d}\cdot\varepsilon$ rather than to our ε.

- Using traditional methods tangential to the main focus of the paper, we give an explicit construction for the Cubical Foam Problem in 3 dimensions with surface area about 5.602. This beats the "Choe prism's" surface area of 5.864 by about 4.5%.

The connection to computational hardness amplification. Our method for constructing a "sphere-like" cubical foam has its roots in an unexpected source: the subject of *computational hardness amplification*, a basic topic of study in the theory of computation. Consider a computational task *T*, such as solving a system of equations, finding the best move in a chess position, optimizing a schedule under constraints, etc. The difficulty, or hardness, of *T* is measured in terms of the computational resources required to obtain a solution of a given quality. "Hardness amplification" asks about ways of transforming *T* into an even harder task; for example, by asking that *T* be solved on *d* inputs simultaneously. Is this "*d* times harder", or can a clever reuse of resources allow for a computation that achieves more than one would naively expect?

This question arises in many areas of computational theory, including cryptography, pseudorandomness, and optimization, and has proven to be extremely subtle. Our foam construction is motivated by the hardness amplification problem in the context of *constraint satisfaction problems (CSPs),* a major topic in computer science, operations research, statistical physics, and information theory (*19*). Here the focus is on the *accuracy* that efficient algorithms can have when solving CSPs. Assuming the standard hypothesis $P \neq NP$, it follows that efficient algorithms cannot guarantee a 100%-accurate solution to a satisfiable CSP. A seminal hardness amplification result, the *PCP Theorem* (*20*, *21*) from 1992, improved this to show "$(1-\varepsilon_0)$-hardness". By this we mean that given a satisfiable CSP, no efficient algorithm can guarantee finding a solution satisfying even a $(1-\varepsilon_0)$-fraction of the constraints. Here ε_0 is a (small) positive constant. Raz's celebrated *Parallel*

Repetition Theorem (*22*) from 1995 dramatically strengthened this: he showed δ-hardness for any $\delta > 0$. Indeed, Raz showed this is true even for "bipartite" CSPs, in which each constraint involves only two variables.

Raz's work was a hardness amplification result: given a "$(1-\varepsilon)$-hard" bipartite CSP, Raz showed that solving *d* instances of it "in parallel" constitutes a roughly $(1-\varepsilon^{32})^d$-hard CSP. This was later (*16*) improved to $(1-\varepsilon^{3})^d$ (and to $(1-\varepsilon^{2})^d$ for CSPs arising from the PCP Theorem (*23*)), but the question of *Strong Parallel Repetition*—namely, whether the resulting CSP is in fact $(1-\varepsilon)^d$-hard—remained open for many years. Note that $(1-\varepsilon)^d \approx 1 - d\cdot\varepsilon$ assuming $\varepsilon << 1/d$; thus the Strong Parallel Repetition question asks to confirm the intuition that having to simultaneously satisfy *d* constraints should amplify the inaccuracy of efficient algorithms from ε to $d\cdot\varepsilon$.

Herein lies the connection to foams. In 2007, Feige, Kindler, and O'Donnell (*14*) investigated parallel repetition of the *Odd Cycle CSPs*, showing hardness amplification from $1-\varepsilon$ to roughly $1-\sqrt{d}\cdot\varepsilon$. They also observed a connection to foam problems, proving that if there exists a cubical foam with surface area *A* then amplification beyond $1-A\cdot\varepsilon$ is impossible. Hence proving Strong Parallel Repetition would require showing that *d*-dimensional cubical foams require surface area proportional to *d*. However, in 2008 Raz (*15*) showed that Strong Parallel Repetition fails for Odd Cycle CSPs, and that amplification to $1-\sqrt{d}\cdot\varepsilon$ is optimal. In retrospect, one can view Raz's work as constructing a kind of "discrete cubical foam" with "surface area" proportional to $\sqrt{d}$. One key tool in Raz's proof was the *Consistent Sampling Lemma*, first used in the context of parallel repetition by Holenstein (*16*).

Our solution to the Cubical Foam Problem involves generalizing Raz's discrete methods to real space, and "opening up" the proof of the Consistent Sampling Lemma. We also use the "Buffon's Needle" method to estimate surface area, and we optimize our results using Fourier analysis. Full mathematical proofs are supplied in the supporting online text.

Our cubical foam and discretization procedure. Before describing our "sphere-like" cubical foam, we will give some motivation for its construction. As stated earlier, our construction can also be interpreted as a very noise-resistant "randomized discretization procedure" for rounding off vectors of real numbers to vectors of integers. Let us make some definitions.

Definitions: *A* discretization *is a mapping which assigns each point* $x = (x_1, \ldots, x_d)$ *in* $\mathbf{R}^d$ *to an integer point* $r = (r_1, \ldots, r_d)$ *in* $\mathbf{Z}^d$ *such that each* r_i *is either* x_i *rounded up or* x_i *rounded down. A discretization is* periodic *if for all* x *in* $\mathbf{R}^d$, x *is assigned to* r *if and only if* $x + s$ *is assigned to* $r + s$ *for each* $s \in \mathbf{Z}^d$. *Given a periodic discretization, we define its* principal bubble *to be the set of all points in* $\mathbf{R}^d$ *assigned to* $(0, \ldots, 0)$.

The principal bubble of a periodic discretization tiles $\mathbf{R}^d$ according to the integer lattice. Thus any periodic discretization immediately yields a cubical foam. We will in fact give a *randomized* procedure whose output is a periodic discretization (hence also a cubical foam). As described earlier, we say that such a procedure is "noise-resistant" if every two nearby points $x, y \in \mathbf{R}^d$ are unlikely to be assigned to different integer points. Intuitively, we expect the bubbles produced by a noise-resistant procedure to have small surface area, because x and y are assigned to different integer points only if the line segment joining them crosses the surface of a bubble.

We will later see that finding a periodic discretization in which nearby pairs x and y are usually assigned to the same integer point is very similar to a bipartite CSP whose "variables" are the points of $\mathbf{R}^d$. Indeed, Raz's analysis (*15*) of the Odd Cycle CSPs suggests the investigation of a related problem (which we state somewhat imprecisely for the sake of brevity): Assign each point x in the unit cube $[0,1)^d$ a random "shift" $z \in [0,1)^d$ in such a way that nearby points are very likely to be assigned the same shift.

For now, let f be a probability density function on $[0,1)^d$ for shifts; let f_x be its translation by $-x$, so f_x is a probability density on $-x+[0,1)^d$; and, let $\tilde{f}_x$ be the periodic extension of this function, so $\tilde{f}_x(z) = f(z + x \bmod 1)$. Holenstein's Consistent Sampling Lemma gives a method for assigning shifts to all points such that the probability x and y are assigned different shifts is essentially $\int_{[0,1)^d} \left| \tilde{f}_x - \tilde{f}_y \right|$. Thus we are led to seek a function f for which this is small whenever x and y are nearby points.

In Raz's CSP analysis it was sufficient to use the Consistent Sampling Lemma as a "black box". However for our construction of a periodic discretization and foam in $\mathbf{R}^d$, we need to analyze the actual proof of the lemma. In our setting, Holenstein's proof would draw a sequence of pairs (Z_i,H_i), with Z_i random in $[0,1)^d$ and H_i random in $(0,\|f\|_\infty)$, where $\|f\|_\infty$ denotes the maximum value of f. Then x would be assigned the shift Z_i for the minimal i with $\tilde{f}_x(Z_i) = \tilde{f}_{Z_i}(x) > H_i$. This method forms the basis of our construction. Given a density function f and a number $h > 0$, consider the shape D consisting of $\{x : f(x) > h\} \subseteq (0,1)^d$ together with its boundary. We call D (or a translate of D) a "droplet". We will want droplets to have smooth boundaries which do not touch the boundary of $[0,1]^d$. For this reason we will require that f's periodic extension $\tilde{f}$ be analytic, equal to 0 on the boundary of $[0,1]^d$; we will call such a density function f *proper*. Given a proper density function f, we can now describe our randomized algorithm for producing a periodic discretization and associated cubical foam:

Algorithm 1: Periodic discretization (and foam) construction, given *f*

1. Let all points in $\mathbf{R}^d$ be "unassigned".
2. For stage i = 1, 2, 3, ... until all points are assigned:
3. Choose a uniformly random shift $Z_i \in [0,1)^d$.
4. Choose a uniformly random parameter $H_i \in (0, \|f\|_\infty)$.
5. Let droplet D_i be $\{x : f_{Z_i}(x) > H_i\}$, together with its boundary.
6. Assign all *currently unassigned* points in D_i to $(0,\ldots,0)$, and extend this assignment periodically.
7. "Color" all newly assigned points from Step 6 with "color *i*".

We remark that the "coloring" done in Step 7 is not essential to the algorithm per se, but is helpful for understanding and reasoning about it.

It is not hard to prove that this algorithm indeed always ends after a finite number of stages (see supporting online text for all mathematical proofs). It is also clear that regardless of the algorithm's random choices, it always produces a periodic discretization. Thus the points assigned to $(0,\ldots,0)$ by the algorithm always constitute a principal bubble which partitions space according to the integer lattice.

We illustrate a sample run of the algorithm in Fig. 4, with d = 2 and $f(x_1, x_2) = 4\sin^2(\pi x_1)\sin^2(\pi x_2)$. Therein the integer lattice is outlined in gray, with the origin $(0,0)$ depicted as a gray dot. The first three panels illustrate stages 1, 2, and 3 of the algorithm. In each stage, the black dot represents $-Z_i$ and the black dashed line outlines the droplet D_i. Colors 1, 2, and 3 are green, yellow, and red, respectively; we have used dark colors to show the points assigned to $(0,0)$, and light colors to show their periodic translations, assigned to $(-1,1)$, $(0,1)$, $(1,1)$, $(-1,0)$, $(1,0)$, etc.

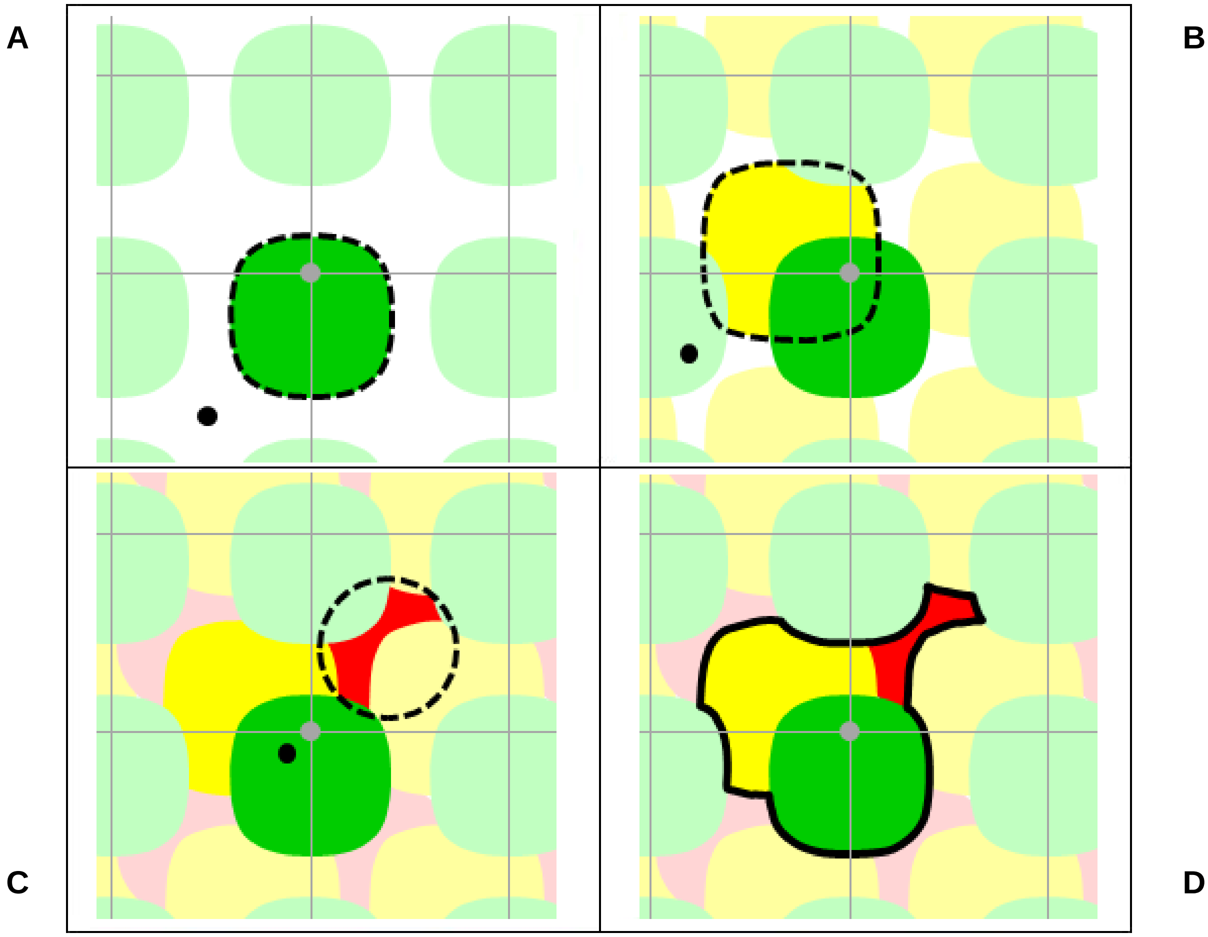


Fig. 4. Example construction. **(A)** The assignment after stage 1: all points in the dark green droplet are assigned to (0,0); the light green translations are assigned periodically. **(B)** The assignment after stage 2; the unassigned (uncolored) points within the outlined droplet are colored dark yellow and are assigned to (0,0). **(C)** The assignment after stage 3, using red. The algorithm terminates after this stage—all points in $\mathbf{R}^2$ have been assigned. **(D)** Here we outline the final bubble which partitions $\mathbf{R}^2$ periodically according to the integer lattice.

Analysis of our construction. We establish the following theorems about Algorithm 1, with all proofs given in the supporting online text.

First we compute the probability of "miscoordinating" a pair x and y in terms of f:

Theorem 1: *Fix a proper density function f. Let* $\overline{xy}$ *be a short line segment in* $\mathbf{R}^d$*; say* $y = x + \varepsilon \cdot u$*, where u is a unit vector and* $\varepsilon > 0$ *is sufficiently small. For a given execution of Algorithm 1, let N denote the times the segment* $\overline{xy}$ *crosses the boundary between differently colored regions. Then*

$$\mathrm{E}[N] \sim \varepsilon \cdot \int_{[0,1)^d} \left|\langle \nabla f, u \rangle\right| . \tag{1}$$

(Here E denotes expectation and the notation ~ means equality up to an error of order ε^2.) Next, using the relationship between noise-resistance and surface area we show:

Theorem 2: *Given an execution of Algorithm 1, let A denote the surface area of the boundary between color regions within* $[0,1)^d$*. Then*

$$\mathrm{E}[A] = \int_{[0,1)^d} \left\|\nabla f\right\| . \tag{2}$$

Finally, we find an f so as to minimize the noise-resistance and surface area:

Theorem 3: *There exists a proper density function f with* $\int \left\|\nabla f\right\| \leq 2\pi\sqrt{d}$*, namely*

$$f(x) = \prod_{i=1}^{d} \left(2\sin^2(\pi x_i)\right). \tag{3}$$

Moreover, for each unit vector u, f satisfies $\int \left|\langle \nabla f, u \rangle\right| \leq 2\pi$.

The bounds we obtain for our cubical foam solution and for the noise-resistance of our coordinated discretization procedure follow easily from these theorems. The bubble *B* output by Algorithm 1 has surface area at most 2*A*, where *A* is the quantity in Theorem 2; hence with the *f* from Theorem 3, the expected value of *B*'s surface area is at most $4\pi\sqrt{d}$. Hence there must *exist* a bubble that tiles $\mathbf{R}^d$ according to the integer lattice with surface area at most $4\pi\sqrt{d}$. As for the noise-resistance of Algorithm 1 as a coordinated discretization procedure: if points $x, y \in \mathbf{R}^d$ at distance ε are assigned different integer points by the algorithm then *N*, the number of times segment $\overline{xy}$ crosses the boundary between color regions, must be at least 1. But the probability of this, i.e., Pr[$N \geq 1$], is at most E[*N*]. Combining Theorems 1 and 3, the probability is at most $2\pi\varepsilon$, as claimed (up to an error of order ε^2, but in fact this can be eliminated).

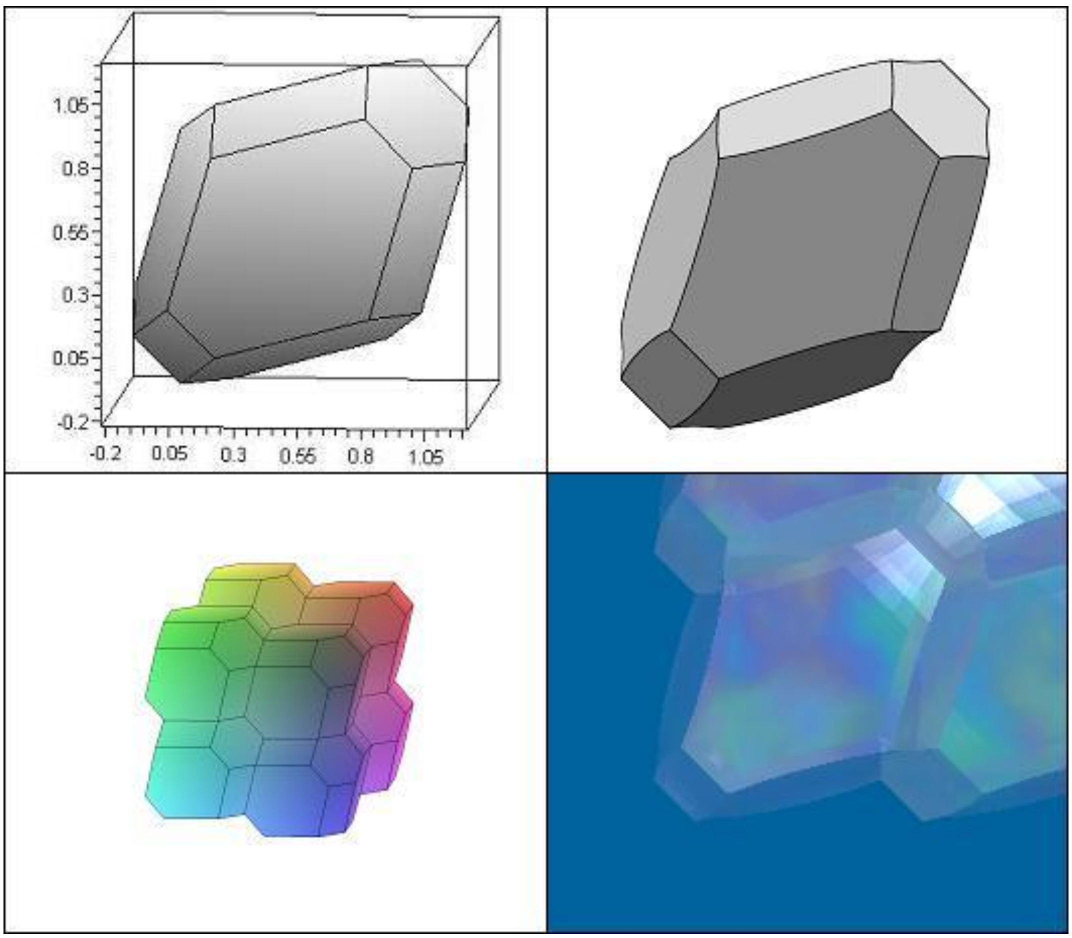
A 3-dimensional cubical foam. Although we have asymptotically solved the Cubical Foam Problem up to a small constant factor, in the physically natural case of $d = 3$ our construction does not improve on the "Choe prism" (or even on the cube). Here we present an improved 3-dimensional cubical foam, constructed via an ad hoc method.

The 2-dimensional minimizer given by Choe in Fig. 2A can (when translated by (½, ½)) be described as follows: Start with a "base" facet centered at the origin; specifically, an edge from $(-s,s)$ to $(s,-s)$ for some parameter *s*. This already gives all vertices, by periodic extension. The hexagonal bubble is the convex hull of the two base points, their translates within $[0,1)^2$, and their translates by (1,1). One chooses *s* to minimize the resulting surface area (perimeter).

We similarly construct a tiling shape *B* in 3 dimensions. We form a "base" facet centered at the origin which is a regular hexagon, with vertices $\pm(0,-t,t)$, $\pm(-t,0,t)$, and

$\pm(-t,t,0)$, for some $t \in (0,1/3)$. Again, this gives all vertices, by periodic extension. We take B to be the convex hull of the 6 base points, along with their 6 translates within $[0,1)^3$ and their 6 translates within $(0,1]^3$. The polytope B has 14 facets: 2 opposing base regular hexagons, 6 larger "isosceles" hexagons, and 6 rectangles. An illustration is in Fig. 5A

A **B**



C **D**

Fig. 5. Our new 3-dimensional cubical foam. **(A)** The unrelaxed tile. **(B)** The tile after it has relaxed according to Plateau's Rules. **(C)** The unrelaxed tile forming a foam according to the integer lattice. **(D)** Illustrating the relaxed foam as soap bubbles.

One may calculate (see the supporting online text) that B has surface area

$$6\left(\sqrt{3}t^2 + \sqrt{2}t\sqrt{1-4t+6t^2} + (1-t)\sqrt{1-2t+3t^2}\right), \tag{11}$$

which is minimized when $t \approx 0.1880$, having minimal value about 5.6121. This already beats the Choe prism.

We can further improve this solution by letting B relax (within the torus $\mathbf{R}^3/\mathbf{Z}^3$) as a soap bubble. Using Brakke's Surface Evolver (*8*), we obtain the relaxed bubble B' shown in Fig. 5B. We remark that it has slightly wavy faces and curved edges, and that the vertices have moved according to $t \approx 0.1814$. The surface area of B' is slightly less than 5.602, according to Surface Evolver.

Discussion. We have given a probabilistic construction of a cubical foam with near-spherical surface area. The construction uses ideas that are new to the study of foams, and is inspired by work on the limits of "hardness amplification" in certain computational optimization problems. Our construction gives the first suggestion that in high dimensions, optimal foams might not be derived from Voronoi cells and may be quite unlike polyhedra.

We have also given an algorithmic application of our foam's construction: a very "noise-resistant" procedure for rounding off vectors of d real numbers to integers. This discretization algorithm may not be practical for very large d, as Algorithm 1 is likely to run for a number of stages which is exponential in d. An important open problem is to find a coordinated discretization procedure with similar noise-resistance but taking time which grows only polynomially in d.

Finally, the construction of our cubical foam used randomness in an essential way; randomness is also used in other efficient high-dimensional constructions of foams (such as high-dimensional Kelvin foams). Although randomness is clearly *required* for noise-resistant coordinated discretization, it is an intriguing question as to whether it is necessary for the construction of foams, or whether "explicit" or "derandomized" constructions exist. Subsequent to the preliminary announcement of our work (*24*), Alon and Klartag (*25*) showed an alternative derivation of our cubical foam via Cheeger's isoperimetric inequality; their analysis also shows that there exists a fixed parameter *h* that can be used as H_i throughout Algorithm 1. In other words, a good foam can be derived from the random translations of a single droplet of the form $D = closure(\{x \in \mathbf{R}^d : \prod_{i=1}^{d} \left(2\sin^2(\pi x_i)\right) > h\})$.

However it still remains unknown how to construct an *explicit* "spherical cube".

References and Notes.

1. J. Plateau. *Statique Expérimentale et Théoretique des Liquides Soumis aux Seules Forces Moléculaires*. (Gauthier-villars, Paris, 1873.)

2. F. C. Frank, J. S. Kasper. Complex alloy structures regarded as sphere packings. I. Definitions and basic principles. *Acta Cryst.* **11**(3), 184–190 (1958).

3. W. N. Arnold, J. S. Lacy. Permeability of the cell envelope and osmotic behavior in Saccharomyces cerevisiae. *J. Bacteriol.* **131**(2), 564–571 (1977).

4. T. C. Hales. A proof of the Kepler conjecture. *Ann. Math.* **162**(3), 1065–1186 (2005).

5. L. E. Nielsen, R. F. Landel. *Mechanical Properties of Polymers and Composites*, 2nd edn (CRC Press, Boca Raton, 1993).

6. P. Ball. Science in culture: Beijing bubbles. *Nature* **448**(7151), 256–256 (2007).

7. W. Thomson. On the division of space with minimum partitional area. *Philos. Mag.* **24**, 503–514 (1887).

8. K. A. Brakke. The surface evolver. *Experimental Math.***1**(2), 141–165 (1992).

9. R. Kusner, J. M. Sullivan. Comparing the Weaire–Phelan equal-volume foam to Kelvin's foam. *Forma* **11**(3) 233–242 (1996).

10. D. Weaire, R. Phelan. A counter-example to Kelvin's conjecture on minimal surfaces. *Phil. Mag.* Lett. 69(2), 107–110 (1994).

11. T. C. Hales. The honeycomb conjecture. Disc. Comp. Geom. 25(1), 1–22 (2001).

12. G. Butler. Simultaneous packing and covering in Euclidean space. *Proc. London Math Soc.* **3**(4), 721–735 (1972).

13. J. Choe. On the existence and regularity of fundamental domains with least boundary area. *J. Differential Geom.* **29**(3), 623–663 (1989).

14. U. Feige, G. Kindler, R. O'Donnell. Understanding parallel repetition requires understanding foams. *Proc. 22nd Ann. IEEE Conf. on Computational Complexity* 179–192 (2007).

15. R. Raz. A counterexample to strong parallel repetition. *Proc. 49th Ann. IEEE Symp. on Foundations of Computing* 369–373 (2008).

16. T. Holenstein. Parallel repetition: simplifications and the no-signaling case. *Theory of Computing* **5**(1), 141–172 (2009).

17. M. Charikar, C. Chekuri, A. Goel, S. Guha, S. Plotkin. Approximating a finite metric by a small number of tree metrics. *Proc. 39th Ann. IEEE Symp. on Foundations of Computing* 379–388 (1998).

18. J. R. Lee, A. Naor. Extending Lipschitz functions via random metric partitions. *Invent. Math.* **160**(1), 59–95 (2005).

19. D. Achlioptas, A. Naor, Y. Peres. Rigorous location of phase transitions in hard optimization problems. *Nature* **435**, 759–764 (2005).

20. S. Arora, S. Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM* **45**(1), 70–122 (1998).

21. S. Arora, C. Lund, R. Motwani, M. Sudan, M. Szegedy. Proof verification and the hardness of approximation problems. *J. ACM* **45**(3), 501–555 (1998).

22. R. Raz. A parallel repetition theorem. *SIAM J. Comput.* **27**(3), 763–803 (1998).

23. A. Rao. Parallel repetition in projection games and a concentration bound. *Proc. 40th Ann. ACM Symp. on Theory of Computing* 1–10 (2008).

24. G. Kinder, R. O'Donnell, A. Rao, A. Wigderson. Spherical cubes and rounding in high dimensions. *Proc. 49th Ann. IEEE Symp. on Foundations of Computing* 189–198 (2008).

25. N. Alon, B. Klartag. Economical toric spines via Cheeger's inequality. *J. Top. Anal.* **1**(2), 101–111 (2009).

26. A. Requicha. Mathematical models of rigid solid objects. *Tech. memo. 28, Production Automation Project, U. Rochester* (1977).

27. E. Bierstone, P. Milman. Semianalytic and subanalytic sets. *Publ. Math. IHES* **67**(1), 5–42 (1988).

28. A. Requicha, R. Tilove. Mathematical foundations of Constructive Solid Geometry: general topology of closed regular sets. *Tech. memo. 27, Production Automation Project, U. Rochester* (1978).

29. L. Santaló. *Integral Geometry and Geometric Probability*, 2nd edn. (Cambridge University Press, Cambridge, 2002).

Acknowledgments. G.K. is supported in part by the Koshland fellowship and by the Binational Science Foundation (BSF). R.O. is supported in part by NSF grants CCF-0747250 and CCF-0915893, BSF grant 2008477, and Sloan, Okawa, and von Neumann fellowships. A.R. is supported in part by NSF grants CCR-0324906 and CCF-1016565. A.W. is supported by NSF grants CCF-0832797 and DMS-0835373. This material is based upon work supported by the National Science Foundation under the above-named grants. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation (NSF). We thank N. Alon, K. Ball, H. Cohn, P. D'Ancona, M. Goresky, J. Håstad, L. Hoffman, J. R. Lee, A. Klivans, D. Micciancio, V. Milman, O. Regev, and J. Sullivan for their insights. We thank R. Nelson for his assistance with ray-tracing code. Image production information, full proofs, and supplemental images appear in supporting online material. The authors contributed equally to this work.

SOM Text

Herein we give complete proof details for the mathematical claims made in the article.

Regularity properties of the droplets and color regions

In this section we establish some regularity properties of the droplets and color regions formed by Algorithm 1.

Prior to introducing Algorithm 1 in the main document we defined the notion of a "proper" probability density function f on $[0,1)^d$; this was one whose periodic extension $\widetilde{f}$ is real analytic and 0 on the boundary of $[0,1]^d$. The first condition is chosen to ensure that the final bubble B has piecewise-smooth boundary. The second condition is to ensure that each droplet

$$D_i = \mathrm{cl}(\{x \in -Z_i + [0,1)^d : f_{Z_i}(x) > H_i\}),$$

is in fact contained within $-Z_i + (0,1)^d$. (Recall that $Z_i \in [0,1)^d$, f_{Z_i} denotes f shifted by $-Z_i$, $0 < H_i < \|f\|_\infty$, and $\mathrm{cl}(G)$ is the topological closure of the set G.) We require this fact that $D_i \subset -Z_i + (0,1)^d$ so that Algorithm 1 produces a well-defined periodic assignment; otherwise, D_i might contain distinct points which are integer shifts of one another.

Let us make these observations precise. Recall that a set K is *closed regular* if $K = \mathrm{cl}(\mathrm{int}(K))$, where $\mathrm{int}(K)$ is the topological interior of K. A set is *semianalytic* if it is everywhere locally equal to a finite Boolean-algebraic combination of sets of the form $\{x : g(x) > 0\}$ where g is real analytic. For more background on such sets, see (*26,27*).

Lemma 0.1. *Each D_i is closed regular and semianalytic. Furthermore, $D_i \subset -Z_i + (0,1)^d$.*

Proof. There is no loss of generality in assuming $Z_i = 0$. We then have $D_i = \mathrm{cl}(G_i)$ where $G_i = \{x \in [0,1)^d : f(x) > H_i\}$. Since $\widetilde{f}$ is 0 on the boundary of $[0,1]^d$, and $H_i > 0$, we must have $G_i \subseteq (0,1)^d$. Thus $G_i = \{x : g(x) > H_i\}$, where g denotes the restriction of f (and also $\widetilde{f}$) to $(0,1)^d$. Since g is analytic on $(0,1)^d$, the set G_i is semianalytic. It is also open by virtue of g's continuity. Thus $D_i = \mathrm{cl}(G_i) = \mathrm{cl}(\mathrm{int}(G_i)) \subseteq [0,1]^d$ is closed regular. D_i is also semianalytic, using the fact that the closure of a semianalytic set is semianalytic. Finally, by continuity of $\widetilde{f}$ we have $\widetilde{f}(x) \geq H_i > 0$ for all $x \in D_i$. It follows that D_i cannot meet the boundary of $[0,1]^d$; hence $D_i \subset (0,1)^d$ as needed. □

Let us now make some more observations about Algorithm 1. In the main article, we described it as also generating new "color regions" in each stage. Let us write

$$\widetilde{D_i} = D_i + \mathbb{Z}^d, \qquad C_i = D_i \setminus (\widetilde{D_1} \cup \cdots \cup \widetilde{D_{i-1}}), \qquad \widetilde{C_i} = C_i + \mathbb{Z}^d.$$

Thus C_i is the set of points assigned to $(0,\ldots,0)$ in stage i of the algorithm, and $\widetilde{C_i}$ is the ith color region. One easily shows by induction that

$$\widetilde{C_1} \sqcup \cdots \sqcup \widetilde{C_i} = \widetilde{D_1} \cup \cdots \cup \widetilde{D_i} \quad \forall i.$$

The algorithm halts as soon as the above set is all of $\mathbb{R}^d$. Later in this supporting text we show that with probability 1 the algorithm halts after some finite number of stages n. In this case the colors are numbered 1 through n, and the final tiling bubble is

$$B = C_1 \cup \cdots \cup C_n. \tag{1}$$

For technical reasons we will prefer to "regularize" the color regions. We define the ith regularized color region to be

$$C_i' = \mathrm{cl}(\mathrm{int}(C_i)),$$

a closed regular set. We also define $\widetilde{C_i'} = C_i' + \mathbb{Z}_d$. We require the following technical lemmas:

Lemma 0.2. *$C_i \subseteq C_i' \subseteq \mathrm{cl}(C_i) \subseteq D_i$ for each i.*

Proof. The latter two inclusions are easy: $\mathrm{int}(C_i) \subseteq C_i \subseteq D_i$, hence

$$C_i' = \mathrm{cl}(\mathrm{int}(C_i)) \subseteq \mathrm{cl}(C_i) \subseteq \mathrm{cl}(D_i) = D_i.$$

As for the inclusion $C_i \subseteq C_i'$, let us write $C_i = D_i \setminus F$, where $F = \widetilde{D_1} \cup \cdots \cup \widetilde{D_{i-1}}$ is closed and D_i is closed regular. Writing $D_i = D$ for simplicity, we have

$$\begin{aligned}
D \setminus F &= (\mathrm{int}(D) \cup \partial D) \setminus F && \text{(since } D \text{ is closed)}\\
&= (\mathrm{int}(D) \setminus F) \cup (\partial D \setminus F)\\
&= (\mathrm{int}(D) \setminus F) \cup (\partial \mathrm{int}(D) \setminus F) && \text{(since } D \text{ is closed regular)}\\
&= (\mathrm{int}(D) \setminus F) \cup (\partial \mathrm{int}(D) \cap \mathrm{int}(F^c)) && (\mathrm{int}(F^c) = F^c \text{ since } F \text{ is closed)}\\
&\subseteq (\mathrm{int}(D) \setminus F) \cup (\partial(\mathrm{int}(D) \cap F^c)) && \text{(Property 2.7.13 in (28))}\\
&= \mathrm{cl}(\mathrm{int}(D) \setminus F)\\
&= \mathrm{cl}(\mathrm{int}(D \setminus F)) && \text{(using Property 2.6.12 in (28) and } F \text{ closed),}
\end{aligned}$$

as claimed. □

Lemma 0.3. *$\partial C_i'$ is contained in ∂D_i and is disjoint from $\mathrm{int}(\widetilde{D_1} \cup \cdots \cup \widetilde{D_{i-1}})$ for each i.*

Proof. Using the notation F from the previous lemma's proof, we have

$$\partial C_i' = \partial \mathrm{int}(C_i) = \partial(\mathrm{int}(D_i \cap F^c)) = \partial(\mathrm{int}(D_i) \cap \mathrm{int}(F^c)) = \partial(\mathrm{int}(D_i) \cap F^c) \tag{2}$$

where we used that F is closed. This set is contained in $\partial \mathrm{int}(D_i) \subseteq \partial D_i$; hence it remains to show that (2) is disjoint from $\mathrm{int}(F)$. We have

$$(2) = \partial((\mathrm{int}(D_i) \cap F^c)^c) = \partial(\mathrm{int}(D_i)^c \cup F),$$

and since the boundary of a set is disjoint from its interior, we conclude that the above set is disjoint from $\mathrm{int}(\mathrm{int}(D_i)^c \cup F) \supseteq \mathrm{int}(F)$, as needed. □

Lemma 0.4. *The sets $\widetilde{D_i}$ and $\widetilde{C_i'}$ are closed regular.*

Proof. This follows easily from the fact that the sets D_i and C_i' are closed regular and are *isolated* from their integer translates by the grid of open cubes $-Z_i + (0,1)^d + \mathbb{Z}^d$. □

Lemma 0.5. *The sets $\widetilde{D_i}$, C_i, C_i', $\widetilde{C}_i$, $\widetilde{C_i'}$, and B are semianalytic.*

Proof. By Lemma 0.1, the integer translates of D_i are isolated from one another. Hence each point of $\mathbb{R}^d$ has a neighborhood on which $\widetilde{D_i}$ is a finite union of translates of D_i. Since these translates are semianalytic (Lemma 0.1), it follows that $\widetilde{D_i}$ is semianalytic. Since finite Boolean-algebraic combination of semianalytic sets are semianalytic, it follows that each C_i is semianalytic, as is B. C_i' is semianalytic since semianalyticity is preserved under interior and closure. Finally, $\widetilde{C}_i$ and $\widetilde{C_i'}$ are semianalytic for the same reason $\widetilde{D_i}$ is. □

A consequence of B being semianalytic and compact is that its boundary ∂B is a piecewise smooth surface (i.e., the disjoint union of finitely many smooth $(d-1)$-dimensional surfaces along with sets of Hausdorff dimension at most $d-2$).

An algorithm on the torus

We now describe how Algorithm 1 can be thought of as also taking place in the torus, $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. Let $p : \mathbb{R}^d \to [0,1)^d$ be the natural map $x \mapsto (x_1 \bmod 1, \dots, x_d \bmod 1)$. As Algorithm 1 is producing droplets D_i, color regions C_i, and regularized color regions C_i', we may also consider the images of these sets in the torus, $p(D_i)$, $p(C_i)$, $p(C_i')$. By Lemma 0.1 p is a bijection on D_i, hence also on C_i and C_i' by Lemma 0.2. We have $\widetilde{D_i} = p^{-1}(p(D_i))$ and similarly for $\widetilde{C}_i$, $\widetilde{C_i'}$. It follows from Lemma 0.4 that $p(D_i)$ and $p(C_i')$ are closed regular in the torus topology, and from Lemma 0.5 that they are semianalytic therein.

Indeed, we can think of the algorithm taking place on the torus $\mathbb{T}^d$ as follows: At each stage, Z_i and H_i are chosen randomly; then $p(D_i)$ is defined to be the translation by $-Z_i$ (within $\mathbb{T}^d$) of $\mathrm{cl}(\{x \in \mathbb{T}^d : f(x) > H_i\}) \subseteq \mathbb{T}^d$, where f is thought of as an analytic function on the torus. The color region $p(C_i)$ is defined by $p(C_i) = p(D_i) \setminus (p(C_1) \cup \cdots \cup p(C_{i-1}))$; the regularized color region $p(C_i')$ is $\mathrm{cl}(\mathrm{int}(p(C_i)))$ (in the torus topology). The algorithm halts once all points of $\mathbb{T}^d$ are colored. The downside of viewing the algorithm exclusively on the torus is that it is less clear what the final tiling bubble $B \subset \mathbb{R}^d$ is. Thus, it is best to think of the algorithm as taking place in parallel on $\mathbb{R}^d$ and on $\mathbb{T}^d$.

Deducing our main results from Theorems 1, 2, 3

In the main article, after the statements of Theorem 1, 2, and 3 we give a paragraph explaining why these imply our main results. In this section we spell out the deductions formally.

We will formally view Theorems 1 and 2 in the main document as being about the algorithm on the torus $\mathbb{T}^d$, and we will interpret "the boundary between color regions" as meaning

$$S = \bigcup_{i=1}^{n} \partial p(C_i').$$

Our first task is to justify the statement that if the $(d-1)$-dimensional area of S is A, then the surface area of B is at most $2A$. By Lemma 0.2,

$$B = \bigcup_{i=1}^{n} C_i \subseteq \bigcup_{i=1}^{n} C_i' \quad \Rightarrow \quad \partial B \subseteq \bigcup_{i=1}^{n} \partial C_i'.$$

Writing $|\cdot|$ for $(d-1)$-dimensional area, we conclude that the surface area of B is at most

$$\sum_{i=1}^{n} |\partial C_i'| = \sum_{i=1}^{n} |\partial p(C_i')|,$$

the equality holding because C_i' is contained in an open cube of side-length 1. We will show that this quantity in fact equals $2|S| = 2A$. In doing so, we may clearly assume that each set $p(C_i')$ is nonempty. Then since the sets $p(C_1'), \ldots, p(C_n')$ are semianalytic and closed regular subsets of the compact torus $\mathbb{T}^d$, they may be triangulated into homogeneously d-dimensional topological sub-polyhedra of $\mathbb{T}^d$. Further, since they meet only at their boundaries (as $C_i' \subseteq \mathrm{cl}(C_i)$ by Lemma 0.2, and the C_i's are disjoint) and since their union is all of $\mathbb{T}^d$, these triangulations may be commonly refined to a homogeneously d-dimensional triangulation of $\mathbb{T}^d$. In this triangulation, each set $\partial p(C_i')$ will be the union of some topological $(d-1)$-dimensional simplices with mutually disjoint interiors. The surface S is the union of all these $(d-1)$-simplices. Each $(d-1)$-simplex K is contained in precisely two topological d-dimensional simplices, which in turn are contained in distinct sets $p(C_j')$ and $p(C_k')$. Further, since the $p(C_i')$'s are homogeneously triangulated, K cannot be contained in any additional $p(C_\ell')$. Thus indeed $\sum_{i=1}^{n} |\partial p(C_i')| = 2|S|$.

Our second task is to justify the analysis of Algorithm 1 as a randomized discretization procedure. First, assume that the distance between points $x, y \in \mathbb{R}^d$ is "sufficiently short" as required for Theorem 1; we will later relax this assumption. Let us justify the claim that if x and y are assigned to different integer points then the line segment joining them—or rather, its image in the torus $p(\overline{xy})$—must intersect the boundary between color regions S. Suppose that $p(\overline{xy})$ does not meet S; since $p(\overline{xy})$ meets *some* $p(C_i')$, it must be in the interior of this $p(C_i')$ and cannot intersect any other $p(C_j')$. Thus all of $p(\overline{xy})$ has color i, and so the same of true of $\overline{xy} \subset \mathbb{R}^d$. Thus the only way x and y could be assigned to different integer points is if x and y were in different integer translates of $C_i' \subset \mathbb{R}^d$. But these translates are isolated from one another by distinct open cubes of side-length 1, meaning that the segment $\overline{xy}$ would have to meet some $\widetilde{C_j'}$ for $j \neq i$.

Finally, we eliminate the assumption that the distance ϵ between x and y be sufficiently short. From Theorems 1 and 3 in the main document we have that there are universal constants $w, W > 0$ such that if $\epsilon \leq w$ then x and y are assigned to different integer points with probability at most $2\pi\epsilon + W\epsilon^2$. If $\epsilon > w$, consider evenly spaced points $x = x_0, x_1, x_2, \ldots, x_{\epsilon/\delta} = y$ at distance δ along $\overline{xy}$, where $\delta \leq w$ satisfies $\epsilon/\delta \in \mathbb{N}$. For each segment $\overline{x_i x_{i+1}}$ we know that x_i and x_{i+1} are assigned to the same integer point except with probability at most $2\pi\delta + W\delta^2$. By a union bound, the probability that $x = x_0$ and $y = x_{\epsilon/\delta}$ are assigned to different integer points is at most

$$(\epsilon/\delta)(2\pi\delta + W\delta^2) = 2\pi\epsilon + W\epsilon\delta.$$

Since this holds for arbitrarily small δ, we conclude that in fact the probability is at most $2\pi\epsilon$, as claimed in the main document.

Analysis of the number of stages Algorithm 1 takes

In this section we view Algorithm 1 as taking place on the torus $\mathbb{T}^d$. Following the description of Algorithm 1 in the main document we state that the algorithm always (i.e., with probability 1) ends after a finite number of stages. This is a consequence of the Borel–Cantelli Lemma together with the following result:

Theorem. *For every proper density function f, there exists $\epsilon > 0$ and $M \in \mathbb{N}$ such that the probability of Algorithm 1 taking more than $M + m$ stages is at most $(1-\epsilon)^m$, for each $m \in \mathbb{N}$.*

Proof. Since f is nonnegative and $\int_{\mathbb{T}^d} f = 1$, there must exist some positive $h_0 > 0$ such that $G_0 := \{x \in \mathbb{T}^d : f(x) > h_0\}$ has positive measure. Since f is continuous, G_0 is open, and so G_0 contains some open cube L with positive measure; say L has side-length δ. Partition $\mathbb{T}^d$ into subcubes of positive side-length $\delta' \in (0, \delta/2]$. Let K denote any such subcube. We claim that in each stage of Algorithm 1, the chosen droplet $D = \mathrm{cl}(\{x : f_Z(x) > H\})$ has probability at least $\epsilon := (\delta/2)^d \cdot (h_0/\|f\|_\infty) > 0$ of covering K. To see this, first note that $H \le h_0$ with probability at least $h_0/\|f\|_\infty$. Assuming this, D will be some translate within $\mathbb{T}^d$ of a set which contains G_0 and hence L. Further, this translate is uniformly random, since Z is uniform on $\mathbb{T}^d$ and independent of H. But a uniformly random translate of L has probability at least $(\delta/2)^d$ of containing K, since L has side-length δ and K has side-length at most $\delta/2$.

Hence for each subcube K, the probability that Algorithm 1 has not chosen a droplet that covers it within the first N stages is at most $(1-\epsilon)^N$. On the other hand, the algorithm must terminate by the time it has chosen droplets which cover every subcube, since all points of $\mathbb{T}^d$ must be colored by this time. Taking a union bound over all $(1/\delta')^d$ subcubes, we see that the probability of Algorithm 1 not terminating after the first $M + m$ stages is at most $(1/\delta')^d(1-\epsilon)^{M+m}$, which is at most $(1-\epsilon)^m$ if we take $M \ge d\ln(1/\delta')/\epsilon$. □

We now also verify the claim, made in the Discussion section of the article, that with the property density function given in Theorem 3, Algorithm 1 is likely to take a number of stages which is exponential in d.

Theorem. *There are universal constants $1 < c < C$ such that Algorithm 1, when run with $f(x) = \prod_{i=1}^d (2\sin^2(\pi x_i))$, takes between c^d and C^d stages except with probability at most c^{-d}.*

Proof. For the upper bound we can follow the proof of the previous theorem. If we set $h_0 = 1$, then $G_0 = \{x \in \mathbb{T}^d : f(x) > h_0\}$ contains every point in $L := (\frac{1}{4}, \frac{3}{4})^d$, since $2\sin^2(\pi x_i) > 1$ whenever $\frac{1}{4} < x_i < \frac{3}{4}$. Thus we can take $\delta = 1/2$ and $\delta' = 1/4$. Further, as $\|f\|_\infty = 2^d$, the quantity ϵ equals $(1/4)^d \cdot (1/2^d) = 8^{-d}$. We may therefore select $M = (\ln 4)d8^d$ and $m = d8^d$ and conclude that Algorithm 1 terminates after $M + m$ stages except with probability at most e^{-d}, as needed.

As for the lower bound, we have

$$\mathbf{Pr}[H_i \le (1.9)^d \text{ for some } 1 \le i \le (1.05)^d] \le (1.05)^d \cdot \frac{(1.9)^d}{2^d} < (.999)^d.$$

We will complete the proof by showing that if $H_i > (1.9)^d$ for each $1 \le i \le (1.05)^d$ then the algorithm does not terminate within the first $(1.05)^d$ stages. To see this, note that for $H_i > (1.9)^d$, the associated droplet $D_i = \mathrm{cl}(\{x \in \mathbb{T}^d : f_{Z_i}(x) > H_i\})$ is contained within a translate of the set

$$V := \{x \in \mathbb{T}^d : \textstyle\prod_{i=1}^d (2\sin^2(\pi x_i)) \ge (1.9)^d\}.$$

We claim that every point $x \in V$ has at least $.96d$ of its coordinates in the range $[\frac{1}{6}, \frac{5}{6}]$. For otherwise, we would have

$$f(x) \le 2^{.96d} \cdot (2\sin^2(\pi/6))^{.04d} = 2^{.92d} < (1.9)^d.$$

It follows that A is contained within the following union of boxes U:

$$U = \bigcup_{\substack{I \subset \{1,\ldots,d\} \\ |I| = \lceil .96d \rceil}} [\tfrac{1}{6}, \tfrac{5}{6}]^I \times [0,1]^{\{1,\ldots,d\} \setminus I}.$$

But the volume of U is at most

$$\binom{d}{.96d} (\tfrac{4}{6})^{.96d} \leq 2^{H_2(.96)d} (\tfrac{2}{3})^d \leq (.79)^d,$$

and so this is an upper bound on the volume of each droplet D_i. Thus the algorithm will not have colored all points of $\mathbb{T}^d$ within the first $(1.05)^d$ stages as $(1.05)^d \cdot (.79)^d < 1$. □

Proof of Theorem 1

We view Theorem 1 as referring to Algorithm 1 on the torus $\mathbb{T}^d$; in this section we will write D_i, C_i, C_i' instead of $p(D_i)$, $p(C_i)$, $p(C_i')$ for simplicity, and we think of f as an analytic probability density function on the torus. We identify $x, y \in \mathbb{R}^d$ with their images in the torus and write ℓ for the image of the segment joining them, $p(\overline{xy})$. We assume the distance ϵ is at most 1 so that ℓ has no self-intersections. We write n for the number of stages the algorithm takes (a random variable); we know that n is finite with probability 1. Recall that the boundary between color regions, S, is taken to mean $\bigcup_{i=1}^n \partial C_i'$ (this is a random set). If N is the random variable denoting the number of intersections between ℓ and S, our task is to show

$$\left| \mathbf{E}[N] - \epsilon \cdot \int_{\mathbb{T}^d} |\langle \nabla f, u \rangle| \right| \leq W\epsilon^2$$

assuming $\epsilon \leq w$, where $w, W > 0$ are universal constants depending only on f.

The desired result is a consequence of the following two lemmas:

Lemma 0.6. *Consider D_1, the first droplet chosen by Algorithm 1 (in the torus). Let I_1 denote the event that $D_1 \cap \ell \neq \emptyset$ and let M_1 be the random variable $\#(\partial D_1 \cap \ell)$. Finally, let $\kappa = \mathbf{E}[M_1 \mid I_1]$. Then*

$$\left| \kappa - \epsilon \cdot \int_{\mathbb{T}^d} |\langle \nabla f, u \rangle| \right| \leq W'\epsilon^2$$

provided $\epsilon \leq w'$, where $w', W' > 0$ are universal constants depending only on f.

Lemma 0.7. *Let κ be as in the previous lemma let ϵ be sufficiently small so that $\kappa < 1$. Then $\kappa \leq \mathbf{E}[N] \leq \kappa/(1-\kappa)$.*

Proof. (Lemma 0.6.) First, note that κ is well-defined: it is easy to show $\mathbf{Pr}[I_1] > 0$ using the proof technique of the previous section. Let us now understand how I_1 and M_1 are determined by the random variables $Z_1, H_1 \in \mathbb{T}^d$. For a given $z \in \mathbb{T}^d$ let us write $g_z : [0, \epsilon] \to \mathbb{R}^{\geq 0}$ for the restriction of the function $f_z(x) = f(x - z)$ to the segment ℓ and write $G(z) = \|g_z\|_\infty$. Recall that $D_1 = \mathrm{cl}(\{x \in \mathbb{T}^d : f_{Z_1}(x) > H_1\})$. Thus

$$\{(z,h) : h < G(z)\} \subseteq I_1 \subseteq \{(z,h) : h \leq G(z)\}$$

and note that the event $\{(z,h) : h = G(z)\}$ has probability 0. As for M_1, write $P = \{s \in [0,\epsilon] : g_{Z_1}(s) = H_1\}$. We have

$$P \supseteq \partial D_1 \cap \ell \supseteq P \setminus \{s \in [0,\epsilon] : g'_{Z_1}(s) = 0\}.$$

The latter inclusion is because any point x where $f_{Z_1}(x) = H_1$ yet $x \notin \partial D_1$ must be a local extremum for f_{Z_1}. Each g'_{Z_1} is real analytic on $[0,\epsilon]$ and hence either has only finitely many zeros on $[0,\epsilon]$ or is identically zero there; in the latter case g_{Z_1} is constant on $[0,\epsilon]$. Regardless, we may conclude that $M_1 = \#P$ with probability 1. Combining these facts about I_1 and M_1 we obtain

$$\kappa = \mathbf{E}[M_1 \mid I_1] = \frac{\int_{\mathbb{T}^d} \int_0^{G(z)} \#\{s \in [0,\epsilon] : g_z(s) = h\}\, dh\, dz}{\int_{\mathbb{T}^d} G(z)\, dz}. \tag{3}$$

The remainder of the proof involves simple estimation. Regarding the denominator of (3), by Taylor's theorem we have $G(z) \approx g_z(0) = f(z)$ up to an additive error of at most $W_1\epsilon$, where W_1 is the maximum length of f's gradient. Since $\int_{\mathbb{T}^d} f = 1$, we conclude that the denominator in (3) equals 1 up to $\pm W_1\epsilon$. Thus we can complete the proof by showing that the numerator of (3) equals $\epsilon \int |\langle \nabla f, u\rangle|$ up to an additive error of at most $W_2\epsilon$, where W_2 depends only on f.

Regarding the numerator of (3), the inner integral equals the vertical distance traveled by a particle moving along the curve g_z; hence it also equals $\int_0^\epsilon |g'_z|$. (This equality is a simpler statement statement than the Crofton formula and would hold even if g_z were merely $\mathcal{C}^1$.) Using Taylor's theorem again, $|g'_z(h)| \approx |g'_z(0)|$ up to an additive error of at most $W_2\epsilon$, where W_2 is the maximum magnitude of f's Hessian's eigenvalues. It follows that, up to an additive $W_2\epsilon^2$, the inner integral equals $\epsilon|g'_z(0)| = \epsilon|\langle \nabla f_z(x), u\rangle|$. It thus remains to observe that

$$\int_{\mathbb{T}^d} |\langle \nabla f_z(x), u\rangle|\, dz = \int_{\mathbb{T}^d} |\langle \nabla f, u\rangle|$$

as required, since the integral on the left does not depend on x (as expected). □

Proof. (Lemma 0.7.) Generalizing the notation from the previous lemma, let I_j denote the event that $D_j \cap \ell \neq \emptyset$ and let M_j be the random variable $\#(\partial D_j \cap \ell)$. Further, let E_j be the event that $\ell \subset \operatorname{int}(D_j)$, and note that

$$E_j = I_j \wedge (M_j = 0).$$

We wish to define J to be the least index such that the event E_J occurs. Now it may be that $\cup_j E_j$ has not occurred by the time the algorithm halts. In this case, it is convenient to think of the algorithm as continuing until some E_j occurs. Note that nothing changes by doing this, since the new color regions C_j and regularized color regions C'_j will all be empty. Since $\mathbf{Pr}[I_j] > 0$ and $\mathbf{E}[M_j \mid I_j] = \kappa < 1$, each event E_j has positive probability. Therefore J, the least index such that E_J occurs, will be finite with probability 1.

Let $j_1 < j_2 < \cdots < j_K$ be the stages in which I_j occurs, up until $j_K = J$. We claim that

$$M_{j_1} \leq N \leq \sum_{k=1}^{K} M_{j_k}. \tag{4}$$

Let us first justify $N \geq M_{j_1}$. For this we need to show that $\partial D_{j_1} \cap \ell \subseteq \partial C'_{j_1} \cap \ell$. Since $C'_{j_1} \subseteq D_{j_1}$, it suffices to show that for each point $u \in \partial D_{j_1} \cap \ell$ and each neighborhood $U \ni u$, there is a point of C'_{j_1} in U. Write $G = (D_1 \cup \cdots \cup D_{j_1-1})^c$, an open set. Since j_1 is the first stage for which I_j occurs, we have $\ell \subset G$. Thus there is a neighborhood $U_0 \ni u$ entirely contained in G. Since $u \in \partial D_{j_1}$ there is a point $v \in D_{j_1}$ in the neighborhood $U_0 \cap U$. Since $v \in G$ it follows that $v \in C_{j_1}$, and hence $v \in C'_{j_1}$ by Lemma 0.2.

Next we justify $N \leq \sum_{k=1}^{K} M_{j_k}$. Since $M_j \neq 0$ implies I_j occurs, we have

$$\sum_{k=1}^{K} M_{j_k} = \sum_{j=1}^{J} M_j \geq \sum_{j=1}^{J} \#(\partial C'_j \cap \ell),$$

where we used $\partial C'_j \subseteq \partial D_j$ (Lemma 0.3). It remains to show that $\ell \cap \partial C'_j = \emptyset$ for all $j > J$. By definition of J, we have that $\ell \subset \mathrm{int}(D_J)$. But Lemma 0.3 says that $C'_j \supseteq \partial C'_j$ is disjoint from $\mathrm{int}(D_J)$ for all $j > J$; hence indeed $\partial C'_j$ is disjoint from ℓ for all $j > J$. This completes the justification of (4).

Note that the distribution of each random variable M_{j_k} is the same as that of $M_1 \mid I_1$; hence $\mathbf{E}[M_{j_k}] = \kappa$ for all $k \leq K$. Taking expectations in (4) we conclude

$$\kappa \leq \mathbf{E}[N] \leq \mathbf{E}[K]\kappa,$$

where we used Wald's Theorem. Now K is distributed as the least index for which a sequence of i.i.d. random variables, $M_{j_1}, \ldots, M_{j_K}$, is 0. Since M_{j_k} is integer-valued, the probability it is 0 is at least $1 - \mathbf{E}[M_{j_k}] = 1 - \kappa$. Hence $\mathbf{E}[K] \leq 1/(1-\kappa)$, the mean of a geometric random variable with parameter $1 - \kappa$. This completes the proof. □

Proof of Theorem 2

To deduce Theorem 2 from Theorem 1, we need a method for computing surface area. We use the following "Buffon's Needle Theorem":

Theorem. *Let S be a piecewise smooth surface in the torus $\mathbb{T}^d$. "Drop a needle" ℓ of length $0 < \epsilon < 1$ into the torus; i.e., let $x \in \mathbb{T}^d$ be uniformly random, let u be a uniformly random unit vector, let $y = x + \epsilon u$, and let $\ell = \overline{xy}$. If N denotes the number of intersections of ℓ with S, then*

$$\mathop{\mathbf{E}}_{\ell}[N] = c_d \cdot \epsilon \cdot \mathrm{area}(S),$$

where c_d is the dimension-dependent constant $\mathbf{E}[|u_1|]$.

This theorem is stated as (8.11) in Santaló's textbook (*29*) in the $d = 2$ case; the extension to higher dimensions is discussed on page 274.

Given an execution of Algorithm 1, we can apply Buffon's Needle Theorem to the resulting S (which is indeed piecewise-smooth since the sets $p(C'_i)$ are semianalytic and compact) to compute its area A:

$$A = \mathop{\mathbf{E}}_{\ell}[N]/(c_d\epsilon).$$

Thus including the randomness of Algorithm 1 we deduce

$$\mathop{\mathbf{E}}_{S}[A] = \mathop{\mathbf{E}}_{S}[\mathop{\mathbf{E}}_{\ell}[N]/(c_d\epsilon)] = \mathop{\mathbf{E}}_{\ell}[\mathop{\mathbf{E}}_{S}[N]]/(c_d\epsilon).$$

But for each fixed ℓ, Theorem 1 tells us that $\mathbf{E}_S[N] = \epsilon \int |\langle \nabla f, u\rangle|$ up to an additive $W\epsilon^2$, where W is a constant depending only on f. Hence

$$\mathop{\mathbf{E}}_S[A] = \mathop{\mathbf{E}}_\ell[\textstyle\int |\langle \nabla f, u\rangle|]/c_d \pm W\epsilon,$$

and so in fact there is exact equality by taking $\epsilon \to 0$. The proof of Theorem 2 is completed by noting that $\mathbf{E}_\ell[|\langle \nabla f(z), u\rangle|] = c_d \cdot \|\nabla f(z)\|$ for any z, and therefore

$$\mathop{\mathbf{E}}_\ell[\textstyle\int |\langle \nabla f, u\rangle|]/c_d = \left(\displaystyle\int_{\mathbb{T}^d} \mathop{\mathbf{E}}_\ell[|\langle \nabla f(z), u\rangle|]\, dz\right)/c_d = \int_{\mathbb{T}^d} \|\nabla f\|$$

as required.

Proof of Theorem 3

Recall that we seek a property density function $f : [0,1)^d \to \mathbb{R}^{\geq 0}$ such that $\int \|\nabla f\|$ is small and such that $\int |\langle \nabla f, u\rangle|$ is small for every fixed unit vector u. Let us express f as $f = g^2$ for some $g : [0,1]^d \to \mathbb{R}$. For f to be a property density function, the following properties of g are required:

1. $\int g^2 = 1$.
2. g is 0 on the boundary of $[0,1]^d$.
3. g^2 is analytic when extended periodically to all of $\mathbb{R}^d$.

Then we can bound $\int \|\nabla f\|$ as follows:

$$\int \|\nabla f\| = 2\int |g| \cdot \|\nabla g\| \leq 2\sqrt{\int g^2}\sqrt{\int \|\nabla g\|^2} = 2\sqrt{\int \|\nabla g\|^2}; \tag{5}$$

here we used Cauchy–Schwarz as well as property 1 of g. Hence the first part of Theorem 3 in the main article is implied by the following:

Theorem. *For piecewise $\mathcal{C}^1$ functions g satisfying properties 1 and 2 above, the minimum possible value of $\int \|\nabla g\|^2$ is $\pi^2 d$ and it occurs when*

$$g(x) = \prod_{i=1}^{d} \sqrt{2}\sin(\pi x_i). \tag{6}$$

This g also satisfies property 3 above.

Proof. Suppose $g : [0,1]^d \to \mathbb{R}$ is piecewise $\mathcal{C}^1$ and satisfies properties 1 and 2. We expand in terms of its (multidimensional) Fourier sine series:

$$g(x) = \sum_{\omega \in \mathbb{N}^d} \widehat{g}(\omega) \prod_{i=1}^{d} \sqrt{2}\sin(\pi \omega_i x_i), \tag{7}$$

where

$$\widehat{g}(x) = \int g(x) \prod_{i=1}^{d} \sqrt{2} \sin(\pi \omega_i x_i).$$

We remark that we have pointwise convergence everywhere in (7). This is because g is piecewise $\mathcal{C}^1$ and because $\sqrt{2}\sin(\pi x_i)$ is 0 on the boundary of $[0,1]$, hence its odd extension to $[-1,1]$ is continuous. More crucially, these conditions also justify term-by-term differentiation of g's sine series, giving us the expansion

$$D_j g(x) = \sum_{\omega in \mathbb{N}^d} \pi \omega_j \widehat{g}(x) \cdot (\sqrt{2}\cos(\pi \omega_j x_j)) \cdot \prod_{i \neq j} \sqrt{2} \sin(\pi x_i). \tag{8}$$

Recall that

$$\int_0^1 (\sqrt{2}\sin(\pi t))^2 \, dt = \int_0^1 (\sqrt{2}\cos(\pi t))^2 \, dt = 1$$

(this is Parseval's identity for sine and cosine series). We use this fact when computing $\int g^2$ and $\int \|\nabla g\|^2$ from (7) and (8):

$$1 = \int g^2 = \sum_{\omega \in \mathbb{N}^d} \widehat{g}(\omega)^2, \tag{9}$$

$$\int \|\nabla g\|^2 = \sum_{i=1}^{d} \sum_{\omega \in \mathbb{N}^d} \pi^2 \omega_i^2 \widehat{g}(\omega)^2 = \pi^2 \sum_{\omega \in \mathbb{N}^d} \widehat{g}(\omega)^2 \|\omega\|^2. \tag{10}$$

Now subject to the condition in (9), it's clear that (10) is minimized when the Fourier sine spectrum of g is concentrated on the frequency ω which minimizes $\|\omega\|^2$, namely $\omega = (1, 1, \ldots, 1)$, for which $\|\omega\|^2 = d$. This means that (6) indeed gives the minimizing g, and the minimum value for $\int \|\nabla g\|^2$ is $\pi^2 d$.

It remains to verify that this minimizing g indeed satisfies property 3; i.e., that g^2 is analytic when extended periodically to all of $\mathbb{R}^d$. Note that this is not true of g itself, since g's periodic extension is $\overline{g}(x) = \prod_{i=1}^{d} \sqrt{2} |\sin(\pi x_i)|$. But on $[0,1]^d$ we have

$$g(x)^2 = \prod_{i=1}^{d} 2\sin^2(\pi x_i) = \prod_{i=1}^{d} (1 - \cos(2\pi x_i)), \tag{11}$$

by a trigonometric identity. Since $\cos(2\pi x_i)$ is periodic on $[0,1]$, the above formula also gives the periodic extension of g^2 to $\mathbb{R}^d$. And the function in (11) is evidently an analytic function on $\mathbb{R}^d$. $\square$

The following completes the proof of Theorem 3 from the main article:

Theorem. *When $f = g^2$ for g as in* (6)*, it holds that $\int |\langle \nabla f, u \rangle| \leq 2\pi$ for all unit vectors $u \in \mathbb{R}^d$.*

Proof. Similar to the deduction of (5), we have

$$\int |\langle \nabla f, u \rangle| = 2 \int |g| \cdot |\langle \nabla g, u \rangle| \leq 2 \sqrt{\int g^2} \sqrt{\int \langle \nabla g, u \rangle^2} = 2 \sqrt{\int \langle \nabla g, u \rangle^2},$$

and hence it suffices to show that

$$\pi^2 \geq \int \langle \nabla g, u \rangle^2 = \sum_{j,k=1}^{d} u_i u_j \int D_j g \; D_k g. \tag{12}$$

In fact, we show that for our particular choice of g we have equality in (12). For the $j \neq k$ terms in (12) we have

$$\begin{aligned} D_j g(x) \cdot D_k g(x) &= (\sqrt{2}\pi \cos(\pi x_j)) \prod_{i \neq j} \sqrt{2} \sin(\pi x_i) \cdot (\sqrt{2}\pi \cos(\pi x_k)) \prod_{i \neq k} \sqrt{2} \sin(\pi x_i) \\ &= (2\pi^2 \sin(\pi x_j) \cos(\pi x_j))(2\pi^2 \sin(\pi x_k) \cos(\pi x_k)) \prod_{i \neq j,k} (\sqrt{2} \sin(\pi x_i))^2. \end{aligned}$$

As $\int 2\pi^2 \sin(\pi x_j) \cos(\pi x_j) = 0$ (and similarly for k), the $j \neq k$ terms drop out of (12). As for the $k = j$ terms in (12), with our g we get

$$\sum_{j=1}^{d} u_j^2 \int (D_j g)^2 = \sum_{j=1}^{d} u_j^2 \int (\sqrt{2}\pi \cos(\pi x_j))^2 \prod_{i \neq j} (\sqrt{2} \sin(\pi x_i))^2 \, dx = \sum_{j=1}^{d} u_j^2 \pi^2,$$

using Parseval's identity. But this equals π^2 as claimed, since u is a unit vector. □

The surface area of our 3-dimensional tiling bubble

In the main article we described a (non-relaxed) polyhedral bubble B, parameterized by some $t \in (0, 1)$, which tiles $\mathbb{R}^3$ periodically according to the integer lattice $\mathbb{Z}^3$. We claimed that surface area of B is given by formula (11) from the main article:

$$6 \left(\sqrt{3}t^2 + \sqrt{2}t\sqrt{1 - 4t + 6t^2} + (1 - t)\sqrt{1 - 2t + 3t^2} \right).$$

Here we verify this formula. (The claim that this quantity has minimal value approximately 5.6121 when $t \approx 0.1880$ can be checked numerically; we used Maple.)

As noted in the article, the polyhedral bubble B has 14 facets: 2 opposing base regular hexagons, 6 larger "isosceles" hexagons, and 6 rectangles. Let F be the base hexagon centered at the origin, let X be the isosceles hexagon "on the bottom", and let R be the rectangle "on the top"; these may be described as

$$F = \text{convex-hull}\Big((0, -t, t),\ (-t, 0, t),\ (-t, t, 0),\ (0, t, -t),\ (t, 0, -t),\ (t, -t, 0)\Big),$$

$$X = \text{convex-hull}\Big((0, t, -t),\ (t, 0, -t),\ (1 - t, t, 0),\ (1, 1 - t, t),\ (1 - t, 1, t),\ (t, 1 - t, 0)\Big),$$

$$R = \text{convex-hull}\Big((0, -t, t),\ (-t, 0, t),\ (0, t, 1 - t),\ (t, 0, 1 - t)\Big).$$

where we listed the vertices in order around the perimeter of each facet. We have

$$\text{surface-area}(B) = 2 \times \text{area}(F) + 6 \times \text{area}(X) + 6 \times \text{area}(R), \tag{13}$$

so it remains to compute the area of F, X, and R. The area of F is 6 times the area of the equilateral triangle with vertices $(0,-t,t)$, $(-t,0,t)$, and $(0,0,0)$. This triangle has area equal to half the magnitude of the cross product $(0,-t,t)\times(-t,0,t)$, namely, $(\sqrt{3}/2)t^2$. Hence

$$\text{area}(F) = 6 \times (\sqrt{3}/2)t^2 = (3\sqrt{3})t^2.$$

The area of the "isosceles" hexagon X can be computed as twice the area of the trapezoid with parallel sides joining $(0,t,-t)$ to $(t,0,-t)$ and $(t,1-t,0)$ to $(1-t,t,0)$. The first of these sides has length $s_1 = \sqrt{2}t$; the second has length $s_2 = \sqrt{2}(1-2t)$. The height of the trapezoid is the distance between the midpoints of these sides, $(t/2,t/2,-t)$ and $(1/2,1/2,0)$; hence the height is $h = \sqrt{2(1/2-t/2)^2+t^2} = \sqrt{1/2-t+3t^2/2}$. Thus

$$\text{area}(X) = 2 \times \frac{s_1+s_2}{2} h = (1-t)\sqrt{1-2t+3t^2}.$$

Finally, the area of the rectangle R is just the product of the lengths of its two side vectors, $(-t,t,0)$ and $(t,t,1-2t)$. Hence

$$\text{area}(R) = \sqrt{2t^2}\sqrt{2t^2+(1-2t)^2} = \sqrt{2}t\sqrt{1-4t+6t^2}.$$

Substituting the calculation of area(F), area(X), and area(R) into (13) gives our claimed formula for the surface area of B.

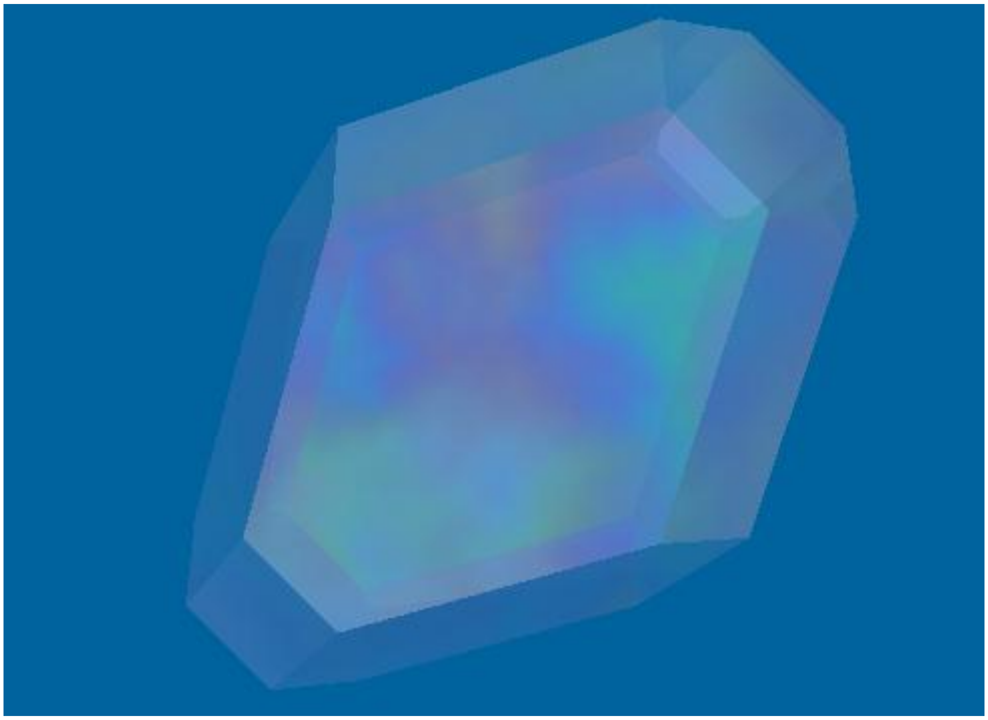


Fig. S1.
Our 3-dimensional bubble before relaxation. This shows our new 3-dimensional tiling bubble in its original polyhedral shape, before being relaxed according to Plateau's rules for soap bubbles.

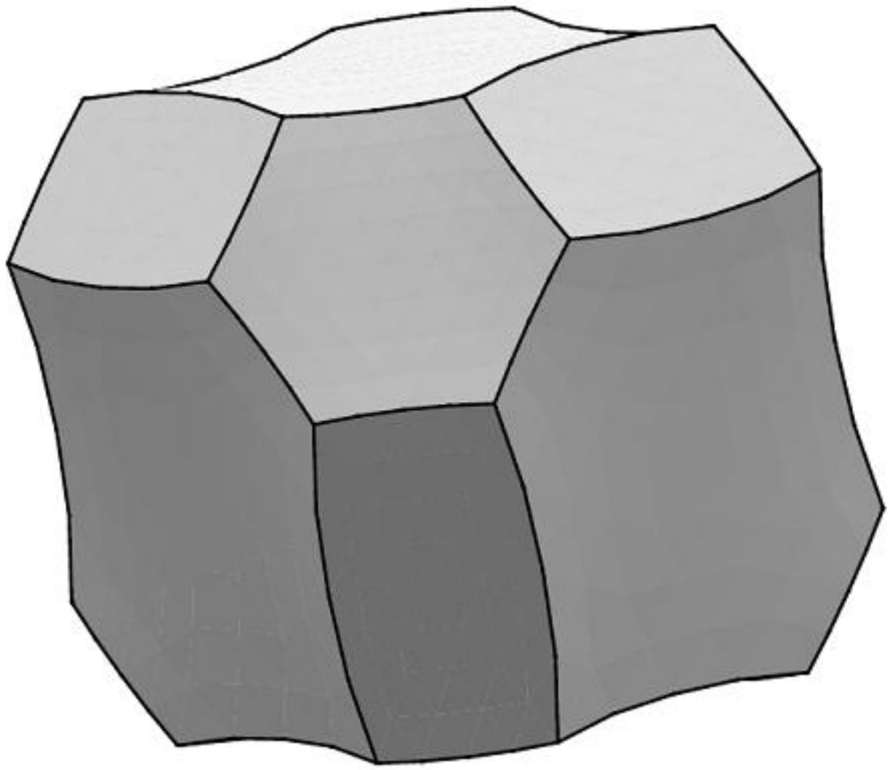


Fig. S2

The relaxed 3-dimensional bubble. This shows (another angle on) the relaxed version of our 3-dimensional bubble, which tiles $\mathbf{R}^3$ according to the integer lattice.

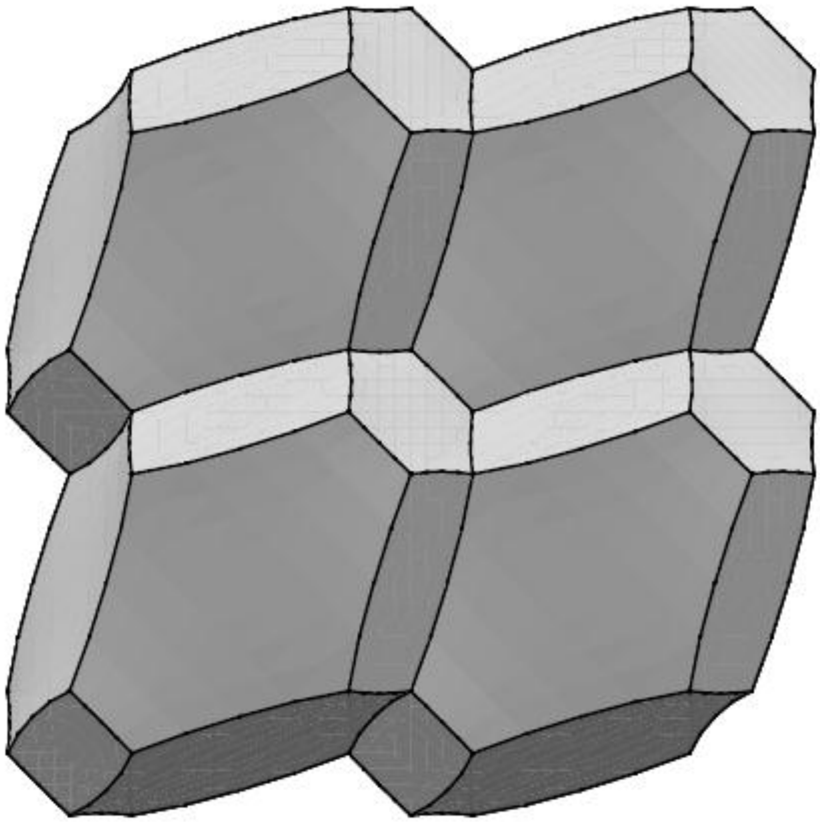


Fig. S3
Four copies of the bubble. This shows how four copies of the relaxed 3-dimensional bubble fit together in a cubical pattern.

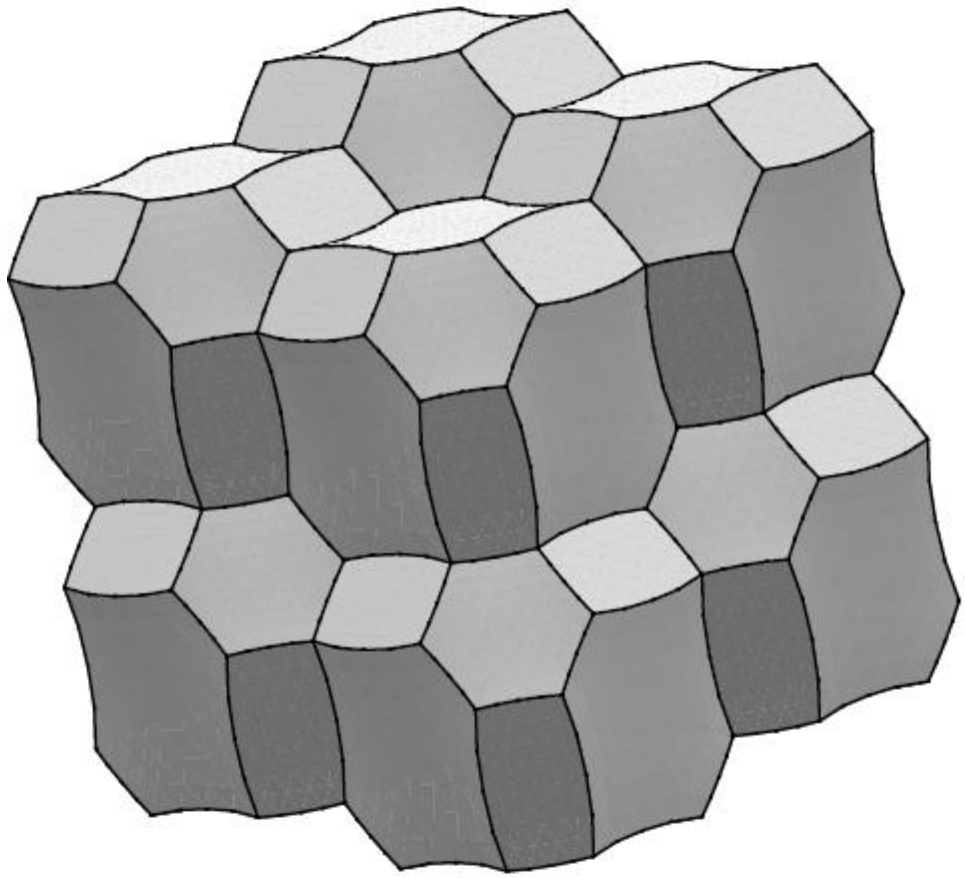


Fig. S4
Eight copies of the bubble. This shows a rotated view of how eight copies of the relaxed 3-dimensional bubble fit together in a cubical pattern.