Found Comput Math
DOI 10.1007/s10208-013-9152-1

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

The Complexity of Some Topological Inference Problems

Shmuel Weinberger

Received: 19 November 2012 / Revised: 12 March 2013 / Accepted: 12 March 2013

Abstract We give some lower bounds on the description, sample, and computational complexities of the problems of computing dimension, homology, and topological type of a manifold, and detecting singularities for a polyhedron.

Keywords Sample complexity · Inference · Entropy · Homeomorphism · Gromov–Hausdorff space · Concentration of measure

Mathematics Subject Classification (2010) Primary 60-07 · 62G99 · Secondary 51K99 · 03D40

The problem of making sense of large and high dimensional data sets is an extremely difficult and important one. In recent years, there has been a growing amount of work, with a number of notable successes (see, e.g., the surveys [6, 14, 19] and the book [38] for an overview, and the recent paper [29] for an outstanding example) in applying geometric or topological methods to this problem. This is, prima facie, reasonable in that the crudeness of topological equivalence forgives many sins—small perturbations of a space are apt to be homeomorphic to the original space—so the robust ideas from topology have the potential to help with a qualitative understanding of noisy data.

Communicated by Gunnar Carlsson.

S. Weinberger (✉)
Department of Mathematics, University of Chicago, 5734 S. University Ave., Chicago, IL 60637, USA
e-mail: shmuel@math.uchicago.edu

Published online: 16 April 2013

The basic idea behind this approach is to imagine that the data that one is considering should be viewed not as merely a sample of an existing population, so that it is a small fraction of some other, much larger, but finite, set, but rather to imagine that it is a sample from a "Platonic ideal" that is assumed to be a "nice" topological space.[1] The goal is to infer information about this topological space, and use that information to better understand the data. The ultimate hope is that one can define invariants of data sets by this method, even if they do not come from any larger topological space.

The basic problem of "topological inference," that is, the determination of a topological space from samples taken from it, can be considered in both "clean" and "noisy" settings. In the clean setting, the sample points (thought of as lying in a Euclidean space) are assumed to be points of the original space, and in the noisy setting, they are some sort of random perturbation of points of the original space. Even in the clean setting, the problem is not at all well posed unless one makes an a priori hypothesis regarding the space. After all, otherwise, the sample could be the whole space, or one could always fit a (windy) curve through all of the points.

In much of the literature (see, e.g., [1, 3, 9, 11, 12, 16, 30, 31]), the assumption made regards the scale at which the nontrivial topology occurs. Assuming that our Platonic model is a smoothly embedded compact submanifold of Euclidean space, we shall assume that there is a number τ (the reach, or feature size, or condition number) so that the normal exponential map is a diffeomorphism for vectors of length $<\tau$ (i.e., in the τ-tubular neighborhood of M, each point has a unique nearest neighbor in M). For a practical data set, such an assumption should be viewed as a provisional subject to test and retest. Obviously, without such a condition it would be impossible to distinguish between a wire and the thin rubber insulation that surrounds it.

Our main focus here is on the sample complexity (or almost equivalently, the information-based complexity; see, e.g., [34, 35]) of some basic problems of topological inference, such as dimension, diffeomorphism (or homotopy or homeomorphism) type, and, in the polyhedral context, detection of singularities. That is, we shall estimate the number of sample points necessary to solve these topological problems in various contexts. In dimension 3 and higher, our results on homeomorphism give lower bounds on the "description complexity" of the solution, so they give lower bounds on the number of measurements necessary for solving the problem (as opposed to the number of samples).

In a couple of cases, we will point out other resource-bounded complexities of the problems (see the undecidability results on singularity detection). The previous works on these problems tended to focus on algorithms to solve the problems and the analysis of those specific algorithms. The work here is complementary and elementary, giving crude lower bounds that are independent of algorithm. However, our observations tend to show that some crude aspects of the known algorithms cannot be improved upon *in general* (although there is much room for detailed improvement, such as lowering the base of an exponential). We hope that these results shed light on the type of problems and data sets for which topological methods are most likely to be effective.

We state our results in Sect. 1, prove them in Sect. 2, and discuss their implications in Sect. 3.

[1] This feels more reasonable for data coming from the physical sciences than for biological or social data.

1 Statement of Results

1.1 Problem 1: Dimension

Let us consider this in its starkest form. We have two possibilities for our manifold: it is either S^n or an equator, $S^{n-1} \subset S^n$. We shall assume that whatever model we are discussing, the points are chosen at random and uniformly from the platonic space.

In that case, if the equator is specified, one can solve the problem with one data point. If this point lies on that equator, then one outputs that the dimension is $n-1$, if not, one outputs n. With probability 0, this will give a "wrong" answer. If the equator is not specified, then one needs $n+1$ points, and if they all lie on an equator, one outputs $n-1$ and otherwise n.

However, if we now convolve with some small amount of Gaussian noise (with some fixed variance that is independent of n), the situation completely changes. The number of samples necessary to decide between the two hypotheses with a certainty of 0.51 grows exponentially with n. The same is true without noise, if our assumption is that either the manifold is the unit sphere S^n or a hypersurface of it which is C^2 close the equator.

On the other hand, even if we are in a fixed Euclidean space, and have an a priori bound on τ and the diameter D of the set (in its Euclidean metric sense), then [7, 8, 30, 31] give a calculation of dimension that grows exponentially with n. (See [13] for a survey of persistent homology methods useful for such purposes. Also see [31] for a discussion of situations in which one can show that these problems can be bounded in terms of the dimension of the submanifold as opposed to the ambient dimension.)

Remark 1.1 These lower bounds do not apply if one has adaptive search methods for the points of M. Essentially this is a bound for unsupervised dimension learning.

1.2 Problem 2: Topological Types

Now we consider a d-dimensional $\tau = 1$ connected compact submanifold of R^k whose diameter is bounded by D. If $d = 0$ or 1, then there is a unique manifold with these conditions. For higher dimensions, we have the following.

Proposition 1.2 *The number of possible topological types grows as D^k for $d = 2$. For $d \geq 3$, it grows as the exponential of this number.*

Proposition 1.3 *Without the assumption of connectivity, the number of topological types grows as D^k for $d = 0, 1$, as $\exp(D^{k/2})$ for $d = 2$, and as $\exp(D^k)$ for $d > 2$.*

Remark 1.4 These propositions are all valid if topological type is interpreted as meaning homeomorphism, diffeomorphism, or homotopy type.

Given that the number of possible answers grows so quickly, it takes a great many bits to even *express* this answer. One needs at least $O(D^k)$ samples to be able to determine the answer for $d > 2$ in any setting, deterministic or random.

Proposition 1.5 *For all d, the sample complexity is $O(D^k)$ (in the disconnected case, or for $d > 2$ in the connected case).*

1.3 Problem 3: Singularities

For many purposes, the assumption that the platonic space is a manifold is unreasonable. Often there are special points where the behavior is different, either because they are extrema of some defining inequality implicit in the system or because they correspond to some symmetry. (For instance, in the space of shapes, studied in [25], singularities can correspond to coincidence of points or more generally to k-points within the ensemble spanning a subspace of dimension $< k - 1$.)

Our interest is not in the number of samples it takes to determine a triangulation of our target space, but rather the computational complexity of solving some natural algorithmic questions about them. So let us assume that we are given a polyhedron, i.e., a union of simplices (e.g., encoded by a list of vertices and their sets that span the simplices).

Proposition 1.6

(a) *It is algorithmically undecidable to tell whether a point in a polyhedron is a singular point or not.*
(b) *It is undecidable whether or not the singularity set of a polyhedron is connected or whether a particular pure stratum is.*
(c) *One cannot algorithmically compute the dimension of the stratum on which a point lies.*

All of the results of this proposition are only true in dimensions >5. For four-dimensional polyhedra, these problems are actually solvable. In dimension 5 (a) is open, but (b) and (c) hold.

Finally, we comment as follows.

Proposition 1.7 *The problem of homeomorphism for two-dimensional polyhedra is decidable, but for homotopy equivalence, it is not. The number of homotopy types of two-dimensional complexes (of bounded geometry)*[2] *in R^k with diameter D grows as* $\exp(D^k)$.

2 Proofs

2.1 Dimension

The propositions regarding sample complexity are relatively straightforward consequences of the volume estimates that arise in the "concentration of measure" phenomenon.

Suppose that we have two probability measures on X, μ and ν, that we are trying to distinguish. A point is chosen from one of the two measures, and we would like to

[2] This is the analog of the condition on τ in the manifold case.

have a method, i.e., a function $f : X \to \{\mu, \nu\}$, that tells us whether we guess that x was chosen from μ or from ν. For example, if $\mu = \nu$, then any method will be wrong $1/2$ of the time.

Clearly, the best one can do is to choose which of $\frac{d\mu}{d\lambda}$ or $\frac{d\nu}{d\lambda}$ is larger (i.e., the Radon–Nikodym derivative) where λ is any measure (e.g., $\mu + \nu$) for which both μ and ν are absolutely continuous. The number of errors obtained by this method is

$$\mathrm{Error}(\mu, \nu) = \frac{1}{2}\int_{\{x \mid \frac{d\mu}{d\lambda}(x) < \frac{d\nu}{d\lambda(x)}\}} d\mu + \frac{1}{2}\int_{\{x \mid \frac{d\mu}{d\lambda}(x) \geq \frac{d\nu}{d\lambda(x)}\}} d\nu.$$

(Note that this is symmetric.) Now, we are not interested in the case of $X = S^n$, but rather—if we are considering methods associated to choosing k-samples from these distributions—the product measures associated to our original possibilities on $(S^n)^k$. However, the estimates in, e.g., [2] show that unless k grows exponentially in n, this Error will be $1/2 - o(1)$. For note that, in this regime, points produced by the S^n distribution will almost certainly be attributed mistakenly to the S^{n-1} distribution (because in a constant size tubular neighborhood of the equator, the S^{n-1} distribution has infinitesimally slightly larger measure).[3]

The deterministic version is even simpler. Without an exponential number of points, almost all of them will be quite near the equator, but not near each other, and it is easy to find a hypersurface with large τ that goes through them.

2.2 Topological Types

The results on the description complexity of homeomorphism come from a different source. The upper bound of the type given (with a terrible estimate on the base of the exponential) follows immediately from known algorithms on topological inference. With the given number of points, one can completely reconstruct the manifolds (see [9, 30] for the case of hypersurfaces, [3, 11, 28] for the general case and also [17]).

The reverse requires construction of the manifolds. (In dimension 2, the results follow immediately from the classification of surfaces, together with, in the disconnected case, the well-known asymptotics of Hardy–Ramanujan [22] for the partition function.)

For higher dimensions, we will use some theory of 3-manifolds (and then cross with tori to yield the situation for dimension >3).[4] Doing it in this way constructs many aspherical manifolds (i.e., manifolds whose higher homotopy groups are all trivial), but they are distinguished by their fundamental groups.[5] We will drape our manifolds around the lattice points of a large box $[-D, D]^k$.[6] Label each point

[3] In other words, if the ν measure of points where $\frac{d\lambda}{d\nu}$ is $1 - \epsilon$, then with s samples one has at least a $(1 - \epsilon)^s/2$ chance of being incorrect.

[4] In dimension >4 many more subtle features can be built into these manifolds, using the methods of [36]. These are not necessary for the current purposes; the construction we give can be viewed as an adaptation of [5].

[5] These manifolds actually have nonpositively curved Riemannian metrics using the work of [26].

[6] We are unconcerned about the $2\sqrt{k}$ factor we've introduced to the diameter (or the extra additive constant that arises for draping the manifold around these points).

by $+/-$. There are the requisite number of configurations; we just have to construct manifolds according to each of these configurations.

Notice that the points of the box have different combinatorial structures, e.g., different numbers of neighbors. Label (arbitrarily) the k edges coming out of a vertex. We shall need for each k, two different k-component links in S^3 whose complement has a hyperbolic structure (with non-homeomorphic complements: take the complement of a nonsplittable link, whose individual components are knots of varying genus, and such that each knot has a hyperbolic complement). Thurston's theorem readily provides infinitely many of these (Theorem 0.2 of [33]). That these complements are non-homeomorphic follows from the observation that the genera of the knots (as an unordered k-tuple) is a homeomorphism invariant of a link complement, since it is topologically defined as the minimal genus of the unique homology class that exists after filling in the other components in the unique way that makes the result into a knot complement (e.g., by [20]). Use one of these for each $+$ vertex, and the other for each $-$.

Now one can embed these link complements around the corresponding vertices and glue boundary tori around corresponding edges—with a uniform bound on τ. The manifold is aspherical, since it is obtained by gluing 3-manifolds together along their incompressible boundary components (see, e.g., [23], Chap. 13). The diffeomorphisms (or equivalently homotopy equivalences, or isomorphisms among fundamental groups, as we have produced closed Haken manifolds; again see [23], Chap. 13) among these manifolds are all induced by labeled graph isomorphisms among the initial combinatorial data, because of the canonical torus (Jaco–Shalen–Johannsen) decomposition of 3-manifolds (the main result of [24]). Each configuration is thus isomorphic to at most $2^k k!$ others, i.e., a multiplicative constant with respect to the dependence of D giving the desired result.

The observation concerning sample complexity is straightforward: with fewer points, an adversary can always find a region big enough to make a change in the topology.

2.3 Singularities

The results on singularities depend on various undecidability results about manifolds.

(a) The suspension of a homology sphere is a manifold iff the homology sphere is simply connected (or, equivalently, is the sphere, due to the Poincaré conjecture). This is undecidable for homology spheres of dimension >4 (according to a theorem of Novikov, proved in [27]). In other words, if one had an algorithm that could detect singularities, we would input the cone on a nonsimply connected homology sphere and see if the cone point were outputted as a singular stratum. Regarding (b) and (c): Let M^k be a manifold other than the sphere, and Σ^{k+1} a homology sphere that cannot be distinguished from the sphere. Now consider $S^1 \times M^k \# \Sigma \# \Sigma$. There is a map: $S^1 \times M^k \# \Sigma \# \Sigma \to S^1$ (that crushes each of the copies of Σ to points, and then projects to the first coordinate) so that in the inverse image of each point aside from two of them is a copy of M. The inverse images of these two points are (small intervals $\times M^k$)$\#\Sigma$. If Σ is a sphere, then the singularity set is a circle; otherwise, there are two most singular points, and the one-dimensional pure stratum consists of two open intervals. Thus, these most singular points either lie on a 0-stratum or a 1-stratum depending on whether or not Σ is a sphere.

The homeomorphism problem for 2-complexes is a straightforward consequence of the result for surfaces. (One first finds a homeomorphism between the 1-skeleta, i.e., the unions of graphs and disjoint points, and then runs into easily computed obstructions to extending these to the rest of the polyhedron.) The impossibility of determining homotopy type follows readily from the unsolvability of the word problem, as any finitely presented group is the fundamental group of a finite 2-complex. (More precisely, a 2-complex with vanishing first homology is homotopy equivalent to a wedge of 2-spheres (of number necessarily the second Betti number of the complex) iff it is simply connected.)

The quantitative bound of homotopy types can be obtained by considering the spines of the 3-manifolds constructed above (i.e., any closed 3-manifold M has a two-dimensional subcomplex which is a deformation retract of M—any point of the subcomplex); it can be constructed with bounded geometry if M is given a triangulation with bounded geometry, i.e., one where there are bounds on the number of the valence of all vertices, the lengths of edges, and the angles between adjacent edges, since this can be constructed as a subcomplex of any given triangulation of the manifold. (See [10] for the construction of nice triangulations of manifolds based on differential-geometric bounds on their geometry.)

3 Remarks

The results given above might have a negative air to them: some very natural problems have very large (e.g., superexponential) complexities when measured in various ways. Our view is that they should help guide researchers who wish to apply geometric topological tools. Some possible implications or interpretations are as follows:

(1) High dimensions are distinctly harder than low, even from the point of view of qualitative description of answers. Thus, one is more apt to discover "low dimensional features" of general objects of arbitrary dimension than to discover their higher dimensional features. An important example of this is the development of novel means of clustering or developing of graphical caricatures of large data sets (such as Mapper [32] with its remarkable application in [29]).

(2) The above point is true both in the supervised and unsupervised setting, in that the description complexity of the topology is overwhelming (with a phase transition in dimension 3 for nonsingular situations).

(3) It thus becomes important to develop methodology to compute topological and geometric invariants, rather than analyze the full spaces. Even non-abelian low dimensional invariants, such as the fundamental group, are susceptible to the growth of descriptive complexity. However, solvable quotients are more likely to be computable and can still contain applications to, e.g., entropy bounds.

(4) This point is analogous to the insight that underlies the theory of testability of graph properties (see, e.g., [18]). See [15] for first steps in the direction of extending these ideas to higher dimensions.

In this way rather than asking, say, how many components there are, one should ask how many "components" there are with at least 5 % of the volume, also understanding that the components might be connected by very thin sets.

(5) This requirement, forced upon us by complexity issues, leads to the study of mm-spaces, that is metric spaces that have a measure as well as their metric (see [21]). This is completely natural in a data analytic setting where the metric is determined by the structural properties of the type of data considered, but where the measure reflects the incidence of individual types in the population.

(6) If one examines the sketch proof, one discovers that a large part of the explosion of topological types is due to the fact that the control on diameter allows a rather large growth in volume, which is what allows a lot of variability for the number of topological types (at least when the situation is "non-abelian", as is the case when the dimension is at least 3). A precursor to this (and it has analogues in our setting) is the result of [5] on the number of closed hyperbolic manifolds (in dimension >3) with volume $<V$ that grows as $\exp(V \log V)$,[7] while according to [37] the analogous statement for hyperbolic manifolds with diameter $<D$ is $\exp(\exp(D))$.

Thus, a first test that should be applied before trying to apply geometric methods is whether the volume is much smaller than suggested by the given diameter. (A second test, in view of item 3, is to see whether the Betti numbers normalized by the volume are small or large: the latter indicating some kind of "foam" rather than a geometric complexity that one can hope to exploit.)

(7) Singularities do introduce new, serious computational difficulties. However, this is not true of all types of singularities; for instance, one can find the boundary of a manifold algorithmically. Consequently, there is room for much useful work in the singular setting, but one needs to be aware of not overreaching.

Acknowledgements This research was partially supported by NSF and AFOSR grants.

I would like to thank Robert Adler, Yuliy Baryshnikov, Gunnar Carlsson, Frederic Chazal, Herbert Edelsbrunner, John Harer, Konstantin Mischaikow, Sayan Mukherjee, Partha Niyogi, Alex Nabutovsky, Vin de Silva, Steve Smale, Jonathan Taylor, Katharine Turner, and Robert Young for useful conversations about related material.

References

1. N. Amenta, M. Bern, Surface reconstruction by Voronoi filtering, *Discrete Comput. Geom.* **22**, 481–504 (1999).
2. K. Ball, *An Elementary Introduction to Modern Convex Geometry, in Flavors of Geometry*. MSRI Lecture Notes (Cambridge University Press, Cambridge, 1997).
3. J.-D. Boissonnat, L.J. Guibas, S.Y. Oudot, Manifold reconstruction in arbitrary dimensions using witness complexes, *Discrete Comput. Geom.* **42**(1), 37–70 (2009).
4. J. Bourgain, On Lipschitz embedding of finite metric spaces in Hilbert space, *Isr. J. Math.* **52**, 46–52 (1985).

[7] As noted above, it is a consequence of the known positive results on the topological inference problem (e.g., [3, 11, 17, 28]) that the embedded data analysis problem has only an $\exp(V)$ growth, not the superexponential growth of the abstract hyperbolic manifold result. Ultimately, this is related to the fact that there is necessarily logarithmic distortion when one tries to embed general graphs in Euclidean space (according to [4]).

5. M. Burger, T. Gelander, A. Lubotzky, S. Mozes, Counting hyperbolic manifolds, *Geom. Funct. Anal.* **12**, 1161–1173 (2002).
6. G. Carlsson, Topology and data, *Bull. Am. Math. Soc.* **46**, 255–308 (2009).
7. F. Chazal, D. Cohen-Steiner, A. Lieutier, A sampling theory for compact sets in Euclidean space, *Discrete Comput. Geom.* **41**(3), 461–479 (2009).
8. F. Chazal, D. Cohen-Steiner, Q. Mérigot, Geometric inference for probability measures, *Journal on Foundations of Computational Mathematics* **11**(6), 733–751 (2011).
9. F. Chazal, A. Lieutier, Smooth manifold reconstruction from noisy and non uniform approximation with guarantees, *Comput. Geom.* **40**, 156–170 (2008).
10. J. Cheeger, W. Muller, G. Schrader, On the curvature of piecewise flat spaces, *Commun. Math. Phys.* **3**, 405–454 (1984).
11. S. Cheng, T. Dey, E. Ramos, Manifold reconstruction from point samples, in *Proc. 16th Sympos. Discrete Algorithms* (2005), pp. 1018–1027.
12. H. Edelsbrunner, D. Cohen-Steiner, J. Harer, Stability of persistence diagrams, *Discrete Comput. Geom.* **37**, 103–120 (2007).
13. H. Edelsbrunner, J. Harer, Persistent homology, a survey, in *Surveys in Discrete and Computational Geometry: Twenty Years After, AMS–IMS–SIAM Joint Summer Research Conference*. Contemp. Math., vol. 453 (Am. Math. Soc., Providence, 2008), pp. 257–282.
14. H. Edelsbrunner, D. Morozov, Persistent homology: theory and practice. IST preprint, http://pub.ist.ac.at/~edels/Papers/2012-P-11-PHTheoryPractice.pdf.
15. G. Elek, Betti numbers are testable. arXiv:0907.5302 (2009).
16. H. Federer, *Geometric Measure Theory* (Springer, New York, 1969).
17. C. Fefferman, B. Klartag, Fitting a C^m-smooth function to data II, *Rev. Mat. Iberoam.* **25**(1), 49–273 (2009).
18. E. Fischer, The art of uninformed decisions: a primer to property testing, *Bull. Eur. Assoc. Theor. Comput. Sci.* **75**, 97–126 (2001).
19. R. Ghrist, Barcodes, the persistent homology of data, *Bull. Am. Math. Soc.* **45**, 61–75 (2008).
20. C. Gordon, J. Luecke, Knots are determined by their complements, *J. Am. Math. Soc.* **2**(2), 371–415 (1989).
21. M. Gromov, *Metric Structures for Riemannian and Non-Riemannian Spaces* (Birkhäuser, Boston, 2001).
22. G.H. Hardy, S. Ramanujan, Asymptotic formulae in combinatory analysis, *Proc. Lond. Math. Soc.* **17**, 75–115 (1918).
23. J. Hempel, *3-Manifolds*. Ann. of Math. Studies, vol. 86 (Princeton University Press/University of Tokyo Press, Princeton/Tokyo, 1976).
24. W. Jaco, P. Shalen, Seifert fibered spaces in 3-manifolds, *Mem. Am. Math. Soc.* **21**, 220 (1979).
25. G. Kendall, D. Barden, T. Carne, H. Le, *Shapes and Shape Theory* (Wiley, New York, 1999).
26. B. Leeb, 3-manifolds with(out) metrics of nonpositive curvature, *Invent. Math.* **122**(2), 277–289 (1995).
27. A. Nabutovsky, Geometry of the space of triangulations of a compact manifold, *Commun. Math. Phys.* **181**, 303–330 (1996).
28. H. Naryanan, S. Mitter, Sample complexity of testing the manifold hypothesis, in *Neural Information Processing Systems (NIPS)* (2010).
29. M. Nicolau, A. Levine, G. Carlsson, Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival, *Proc. Natl. Acod. Sci. USA* **108**(17), 7265–7270 (2011).
30. P. Niyogi, S. Smale, S. Weinberger, Finding the homology of submanifolds with high confidence from random samples, *Discrete Comput. Geom.* **39**, 419–441 (2008).
31. P. Niyogi, S. Smale, S. Weinberger, A topological view of unsupervised learning, *SIAM J. Comput.* **40**, 646–663 (2011).
32. G. Singh, F. Memoli, G. Carlsson, Topological methods for the analysis of high dimensional data sets and 3D object recognition, in *Eurosymposium on Point Based Graphics*, (2007).
33. W. Thurston, Hyperbolic structures on 3-manifolds I: deformations of acylindrical manifolds, *Ann. Math.* **124**, 203–246 (1986).
34. J.F. Traub, A.G. Werschulz, *Complexity and Information* (Oxford University Press, Oxford, 1998).
35. J.F. Traub, H. Woźniakowski, G.W. Wasilkowski, *Information-Based Complexity* (Academic Press, New York, 1988).
36. S. Weinberger, *Computers, Rigidity, and Moduli* (Princeton University Press, Princeton, 2006).
37. R. Young, Counting hyperbolic manifolds with bounded diameter, *Geom. Dedic.* **116**, 61–65 (2005).
38. A. Zomorodian, *Topology for Computing* (Cambridge University Press, Cambridge, 2005).