Discrete spacetime and its applications

Eitan Bachmat

Dedicated to Percy deift on the occassion of his sixtieth birthday

A discrete spacetime is a random partially ordered set. It is obtained by sampling n points from a compact domain in a spacetime manifold w.r.t. the volume form. The partial order is given by restricting the causal structure on the domain to the sampled points. These objects were introduced by physicists in an attempt to reconcile gravity with quantum mechanics, [**39, 45, 18**]. They received very little direct attention in the mathematics literature.

Recently, discrete spacetimes have appeared in very diverse applications which are very far removed from their original motivation. In this self-contained survey we will present some results and numerical calculations regarding the asymptotic behavior of these objects. Some of the more basic results are stated as theorems since they are particularly useful in the applications. Numerical calculations are presented with the hope that they will incite some theoretical work, aimed at proving some related assertions. We then present a few of the applications by shamelessly following the pattern of the recent survey paper [**20**], which describes some related material. We will present a short list of seemingly unrelated problems and show how they fit into the discrete spacetime formalism. While this is a survey paper, it does contain some new material, namely, the statement of theorem 1.3, the (simple) solution to the dimension reconstruction problem, the numerical calculations of figure 2, the connect-the-dots application and the relation between discrete spacetime and the literature on maximal layers. A more comprehensive and extensive treatment of the subject along with full proofs of the basic results will be presented in a forthcoming publication.

Our own research relating to discrete spacetime owes much to the constant encouragement which was provided by Percy Deift and it is our sincere pleasure to dedicate this paper to him on the occasion of his 60th birthday.

1. The asymptotic behavior of discrete spacetime

A spacetime (Lorentzian manifold) M is an $m = d + 1$ dimensional real, connected, C^∞ manifold with a globally defined C^∞ symmetric tensor field g, of type $(0, 2)$, which is non-degenerate and Lorentzian. By Lorentzian we mean that for any $a \in M$ there is a basis of $T_a(M)$, the tangent space to M at a relative to

---

which g_a is given by the diagonal matrix with diagonal entries $(-1, 1, 1, \ldots, 1)$. We consider a set of local coordinates $x_0, \ldots, x_d$ on the manifold. In terms of the local coordinates it is convenient to consider the expression

$$(1) \qquad ds^2 = \sum_i \sum_j g_{i,j} \, dx_i \, dx_j$$

The metric defines a canonical volume form which, using the coordinates, may be written as

$$(2) \qquad \text{Vol}_a = \sqrt{|\det(g_a)|} \, d^m x$$

The metric g_a also classifies the tangent vectors $v \in T_a M$ into *timelike* vectors, if $g_a(v,v) > 0$, *null* or *lightlike*, if $g_a(v,v) = 0$ and *spacelike* if $g_a(v,v) < 0$. The set of timelike vectors consists of a double cone. We say that a spacetime is time orientable if there is a non-vanishing vector field V consisting of timelike vectors. We will assume that all our spacetime manifolds are time orientable. An orientation provides a continuous choice of one out of the two cones (the one containing the vector V_a) and we declare this cone to be the future-pointing cone. A parametrized curve in the manifold is said to be *future-pointing timelike* if all its tangent vectors belong to the future-pointing cone.

The simplest spacetime manifold is flat Minkowski space, which is $\mathbf{R}^m$ equipped with the constant metric

$$(3) \qquad ds^2 = dx_0^2 - dx_1^2 - dx_2^2 - \ldots - dx_d^2$$

Other natural examples to consider are the constant curvature hyperbolic (anti de-Sitter) and spherical (de-Sitter) spacetime manifolds. In two dimensions, hyperbolic spacetime with constant curvature -1 is given by the metric

$$(4) \qquad ds^2 = \frac{1}{cos^2(y)}(dx^2 - dy^2)$$

where $-\pi/2 < y < \pi/2$. Spherical spacetime is realized on the same manifold by simply inverting the sign in (4).

Consider a compact domain $D \subset M$. For the applications we have in mind it is sufficient to consider very "nice" domains, without global pathologies, or intricate smoothness issues, so we will assume for simplicity that the interior of D is an m dimensional, contractible, submanifold of M and that the boundary of D is C^∞, except possibly at a lower dimensional C^∞ submanifold. The causal relation (past-future) is defined on pairs $a, b \in D$ by $a \prec b$ iff there is a future-pointing timelike curve in D from a to b. We will also assume that the relation $\prec$ is a partial order relation, namely, there are no points $a \neq b \in D$ with future-pointing timelike curves (in D) from a to b and from b to a.

Given a spacetime M and a pair of points $a \prec b$, a natural domain to consider is the interval $[a, b]$ consisting, of all points z which satisfy $a \prec z \prec b$. The interval $[a, b]$ for a pair of points in the Minkowski plane is shown in figure 1, as the tilted square $|x + y| \leq \delta$, whose vertices are a, b, e, f. The same domain doubles as the interval of $[a, b]$ in spherical spacetime if $\delta < \pi/2$ and of $[e, f]$ in hyperbolic spacetime (restricting to $|y| < \pi/2$). The interval $[e, f]$ is a compact domain in hyperbolic spacetime if $\delta < \pi/2$.

A *discrete spacetime* associated with a compact domain D is a random partially ordered set on n points obtained by sampling the n points from D using the volume form and inducing the relation $\prec$ on the chosen points.
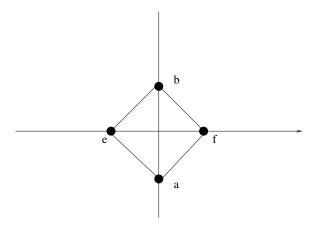
FIGURE 1. An interval $[a, b]$ in the Minkowski plane.

We are interested in the asymptotic properties of discrete spacetimes associated with a domain D as n becomes large. In particular, we may think of a discrete spacetime as a finite approximation to D. It is then natural to ask whether we can "reconstruct" D from larger and larger discrete spacetimes associated with it.

At first this seems difficult since we have thrown away the original metric and retained only the causal relation and the volume form. However, results of Hawking, King and McCarthy, [**27**], and of Malament, [**37**] show that, in the continuous setting, knowledge of causality and the volume form is essentially equivalent to knowledge of the metric.

We can consider more closely the type of quantities that we may want to reconstruct. Given a timelike curve C we may define its length $L(C)$, better known as *proper time*, to be $\int_C ds$. In relativity theory, this parameterization independent quantity measures the time which passes on a clock which is attached to a point-like object whose trajectory is given by the curve C. Proper time is also the action for the trajectory of a point-like object between two endpoints $a, b \in M$. This means that a free falling point-like object moving between events a and b will follow a critical path for this action, this is Einstein's equation of motion. The critical paths, which are maximizing in this case, are the timelike geodesics in analogy with the situation in Riemannian geometry (where they are minimizing).

Assume that $a, b \in D$, with $a \prec b$. We may define their "distance" in D to be $L_D(a, b) = \sup_C L(C)$, where the supremum is taken over all timelike curves $c \subset D$, whose initial point is a and whose endpoint is b. It is known that $L_D(a, b)$ is finite when D is compact and is achieved by a timelike curve, [**40**]. We shall call such a curve a *maximal curve*. A maximal curve may have non-geodesic segments on the boundary of D, but in the interior of D it must satisfy the geodesic equation. It must also be differentiable at all points where the boundary is differentiable.

We would like to have discrete analogues of the concepts of length and maximal curves. Given a partial order $\prec$, a *chain* of size k is a set of elements $x_1, \ldots, x_k$ such that $x_i \prec x_{i+1}$ for all $1 \le i < k$. A chain in a discrete spacetime P associated with D is the obvious discrete analogue of a timelike curve. Under this correspondence the analogue of curve length is simply the size of the chain, and the analogue of a maximal curve between a and b will be a maximal sized chain $x_1, \ldots, x_k$ in P such

that $a \prec x_1$ and $x_k \prec b$. We let $L_P(a, b)$ denote the size of such a maximal sized chain.

We will say that an event occurs with high probability (w.h.p.) if it occurs with probability approaching 1 as n tends towards infinity. The following folk theorem provides the asymptotic high probability link between the discrete distance $L_P(a, b)$ and the continuous distance $L_D(a, b)$.

THEOREM 1.1. *Let D be a compact spacetime domain with the metric normalized so that $\mathrm{Vol}(D) = 1$. Let $a, b \in D$ be such that $a \prec b$. Let P denote a discrete spacetime associated with D, then, for any $\varepsilon > 0$, w.h.p.*

$$(5) \qquad |\frac{L_P(a, b) - c_d L_D(a, b) n^{1/(d+1)}}{n^{1/(d+1)}}| < \varepsilon$$

where c_d is a constant which depends only on the dimension. Moreover, given any Riemannian metric which is compatible with the differential structure of M, and any $\varepsilon > 0$, the maximal chain in P between a and b will be contained w.h.p. in an ε-neighborhood of a maximal curve in D between a and b.

The result has a somewhat tortured history. It is implicit in [39]. It was proved in Minkowski space in [16], stated explicitly in [17] and proven in dimension $d + 1 = 2$ in [21]. The proof of [21], which is in a completely different context (no mention of spacetime geometry), generalizes to arbitrary dimension. Meanwhile, in the physics literature on discrete gravity it seems to retain the status of a conjecture, [49, 28]. A numerical study can be found in [28].

The constant c_d which appears in the theorem is known only for $d = 1$, where according to results of Kerov and Vershik and of Logan and Shepp, [51, 33] we have $c_1 = \sqrt{2}$. For $d > 1$ there are only bounds for c_d, [16], which are obtained via first moment methods.

The proof of theorem 1.1 is fairly elementary. It is first proved for an interval in Minkowski space. In this case it follows from a simple scaling argument and the transitivity of the Lorentz transformation group (the isometries) on timelike lines. The move from the local result to the global one is achieved using finite coverings of D by small patches where the metric is nearly constant, and by using the upper semi-continuity of the length functional, [40].

Given the law of large numbers $L_P(a, b) \approx c_D L_D(a, b) n^{1/(d+1)}$, it is natural to ask for a central limit theorem. Therefore, we define the random variable error term

$$(6) \qquad \Delta_{D,n}(a, b) = L_P(a, b) - c_D L_D(a, b) n^{1/(d+1)}$$

A general concentration result of Talagrand, [46], shows that w.h.p. $\mathrm{Var}(\Delta_{D,n}) = O(n^{1/(d+1)})$. For $d > 1$, essentially nothing is proven, beyond this fact. One expects that when the maximal curve is unique and entirely contained in the interior of D, the order of magnitude of $\Delta_{D,n}$ will only depend on the dimension of D. One may further expect that as n goes to infinity, $\Delta_{D,n}$ will have the limiting form

$$(7) \qquad n^{\beta_d/(d+1)} \chi_D$$

where β_d is a dimension dependent constant and χ_D is some distribution. When $d = 1$, and $D = [a, b]$ is an interval in the Minkowski plane, the famous Baik-Deift-Johansson theorem, [11] states that $\Delta_{D,n}$ does have a limiting form as above with $\beta = 1/3$ and $\chi_D = F_2$ where F_2 is the Tracy-Widom distribution of fluctuations of
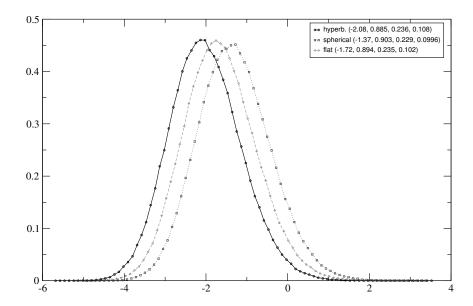
FIGURE 2. Δ_n for intervals in flat, spherical and hyperbolic space.

the largest eigenvalue in the GUE ensemble. Results of Baik and Rains [**12, 13**] show that χ_D for the triangular domain a, e, b in figure 1 is F_4, the Tracy-Widom distribution of GSE, and the distribution corresponding to the triangle, a, e, f is F_1, the Tracy-Widom distribution for GOE, see [**47**] for more on these distributions. These results are much deeper than the other results presented in this survey.

Figure 2 shows the distribution of $\Delta_{D,n}/n^{1/6}$ in discrete spacetimes with $500,000$ points associated with a hyperbolic, flat and spherical interval, which are represented by the domain $|x| + |y| \leq \pi/8$. Each distribution was composed of roughly $350,000$ samples. For the domain in question one can calculate $c_d L_D$ analytically to obtain $c_d L_D \approx 1.973\sqrt{n}$ for the hyperbolic metric, $c_d L_D \approx 2.026\sqrt{n}$ for the spherical and $c_d L_D = 2$ for the flat case. The figure also displays the average, variance, skew and kurtosis, for the distributions. The respective values for F_2 (flat space asymptotics) are approximately $-1.771, 0.902, 0.224, 0.093$, [**47**]. The numerics suggest that in hyperbolic and spherical spacetime the distribution χ_D has the form $F_2(\alpha x + \gamma)$, where γ has the same sign as the curvature and $\alpha = 1$ if the interval domain is calibrated so that $c_d L_D = 2$. A relation between the shift γ and curvature is expected since the curvature controls the length of curves which are close to the geodesic.

Moving to $d > 1$ nothing is proven even for the flat interval case. The behavior of $\Delta_{D,n}$ has been studied in the context of statistical physics because one of the simplest surface growth models, known as *Polynuclear growth* (PNG) can be described in terms of a discrete spacetime [**42, 41**]. In the PNG model one considers a crystal in contact with a super-saturated vapor. Once in a while a nucleus is formed on the existing surface and one assumes that the nucleus spreads evenly in all directions at a constant speed, forming a round new layer. When two such

layers collide they coalesce. In terms of the PNG model, theorem 1.1 describes the macroscopic shape of the surface while the properties of $\Delta_{D,n}$ describe finer aspects such as roughness of the surface. PNG is a model within a larger class of Kardar-Parisi-Zhang growth processes which are expected to exhibit some universal features, including the values of the exponents β_d. Some heuristic theoretical considerations in physics, [**32**] , suggest that β_2 and β_3 have the form $x/(2-x)$ for some x of the form $2/(k+2)$, where k is an odd integer. Numerical measurements predict that $\beta_2 = 0.24$, $\beta_3 = 0.165$, [**41**]. The closest values predicted by the heuristic argument would be $\beta_2 = 1/4$ ($k = 3$) and $\beta_3 = 1/6$ ($k = 5$), however, it is claimed in [**41**] that the numerical value $\beta_2 = 0.24$ seems significant enough to exclude $\beta_2 = 1/4$. A very interesting prediction of the physical heuristic arguments is that $\mathrm{Var}(\Delta_{D,n}) < n^\varepsilon$ for all $\varepsilon > 0$ and n large enough, when $d \geq 4$ (essentially $\beta_d = 0$ for $d \geq 4$.). The existence of such a critical dimension is subject to great debate in the physics literature, [**19, 32, 35**], and is currently not supported by numerical evidence. Numerics lead to an estimated value of β_4 in the range $0.1 - 0.15$. We note that many of the numerics are not computed on discrete spacetime, but rather, on closely related partially ordered sets in which the light cone is replaced by a "simplicial light cone". The main example is the unit m dimensional cube with the partial order relation $(x_0, \ldots, x_d) \prec (x'_0, \ldots, x'_d)$ iff $x_i \leq x'_i$ for all $0 \leq i \leq d$. It is expected that, after proper normalization, the fluctuations of the longest chain for this partial order are the same as those of the flat interval spacetime (universality). Numerical calculations are easier and faster for the simplicial partial order.

Another result which is strongly related to theorem 1.1 concerns the behavior of a random chain of length n, as n becomes large. We will state a semi discrete version.

THEOREM 1.2. *Let $a, b \in D$ Let $P = \{x_1, \ldots x_n\}$ be a discrete spacetime associated with D which is conditioned to be a chain between a and b , i.e., $a \prec x_1 \prec x_2 \prec \ldots \prec x_n \prec b$, then for any $\varepsilon > 0$, w.h.p., P is contained in an ε neighborhood of a maximal curve between a and b.*

Locally, where maximal curves and timelike geodesics coincide we can think of theorem 1.2 as providing a nice probabilistic interpretation for Einstein's law of motion, namely, a completely random trajectory of a particle between two endpoints is likely w.h.p. to closely follow a geodesic. So, perhaps God does play dice but as stated by a wise old man, "everything is foreseen, yet freedom of choice is given", Rabbi Akiva, Avot 3:19.

Theorem 1.2 was proven in [**21**], again in a very different context, for two dimensional domains, the proof extends to arbitrary dimension. A probabilistic interpretation along these lines to Einstein's equation of motion was sought in [**23**]. For similar and related results in the context of relativistic Markov processes and Brownian motion see [**24, 25, 26**].

So far we have considered maximal chains in a discrete spacetime. We can consider the analogues problems for anti-chains. An *anti-chain* is a set of points no pair of which is comparable w.r.t. $\prec$. In a continuous spacetime this is known as an *achronal set.* In the same way that chains are related to the concept of timelike curves, anti-chains are closely related to the concept of a spacelike hypersurface. A differentiable hypersurface is said to be spacelike if all its tangent vectors are spacelike. We have noted that geodesics locally maximize proper time. Proper time may be identified with the induced volume of a timelike curve. Similarly we

may consider *maximal hypersurfaces* which are spacelike hypersurfaces which locally maximize the induced volume among spacelike hypersurfaces. The corresponding Euler-Lagrange equation which is satisfied by a maximal hypersurface states that the scalar curvature vanishes along the hypersurface. In analogy with theorems 1.1 we can state the following result.

THEOREM 1.3. *Given P, a discrete spacetime associated with a domain D of unit volume, let $a(P)$ denote the size of a maximal anti-chain in P. Let $a(D)$ denote the maximal volume of an achronal spacelike hypersurface in D, then there exists a constant c'_d such that for all $\varepsilon > 0$, w.h.p.,*

$$(8) \qquad |\frac{a(P) - c'_d a(D) n^{d/(d+1)}}{n^{d/(d+1)}}| < \varepsilon$$

Moreover, if $A \subset P$ is an anti-chain of size $a(P)$, then, w.h.p., there is a spacelike hypersurface N of volume $a(D)$ such that A is in an ε neighborhood of N.

Similarly, we have the analog of theorem 1.2, which states that a large random anti-chain in D will cluster around an achronal spacelike hypersurface of maximal volume.

Geodesics and maximal hypersurfaces are examples of critical points of the volume functional. The Nambu-Goto action in string theory is also given by the volume functional, this time applied to surfaces which have a timelike tangent and a spacelike tangent at each point. Such surfaces describe the (classical) trajectory of a string. As before, one is interested in the critical points of this action. It would be interesting to provide some probabilistic interpretation to the critical points via causality preserving maps from a weighted family of discrete spacetimes.

We can use theorems 1.1 or 1.3 to provide a very simple procedure for reconstructing the dimension of D from a large associated discrete spacetime. Several solutions for this problem were proposed, [**38, 49, 43**], however, all of them rely on specific properties of intervals in Minkowski space and do not work for general domains in general spacetimes.

Let P be a given discrete spacetime associated with D. Randomly subdivide P into two disjoint sets P_1 and P_2 of size $2n/3$ and $n/3$ respectively. We compute the sizes k_1, k_2 of longest chains in P_1, P_2 and the sizes l_1, l_2 of the maximal anti-chains. By theorem 1.1 the ratio k_1/k_2 will approach $2^{1/(d+1)}$ w.h.p., therefore we can use $1/\log_2(k_1/k_2)$ as our estimate of the dimension of D. Similarly, by theorem 1.3 the ratio l_1/l_2 will approach $2^{d/(d+1)}$ and so $1/(1 - 1/\log_2(l_1/l_2))$ can serve as a dimension estimator. The basic idea of dividing P into two sets of uneven size is rather flexible and can be used in conjunction with many other properties of the partially ordered set in a similar way. The accuracy of the method depends on the properties of error terms like $\Delta_{D,n}$ or $a(P) - c'_d a(D) n^{d/(d+1)}$, to the best of our knowledge the latter has not been explored even numerically.

Having considered chains and anti-chains we may also consider a well known "bottom to top" peeling process which involves both. Let P be a finite partially ordered set. We denote by A_1 the set of minimal elements of P, by A_2 the set of minimal elements of $P - A_1$ and more generally, by A_i the set of minimal elements of $P - \cup A_j, j < i$. The number of such sets is well known to be equal to the size of the longest chain, hence we can analyze its asymptotic behavior using theorem 1.1. The peeling process will figure prominently in the applications which will be described below.

We can symmetrically define the "top to bottom" peeling process with A_i' defined recursively as the set of maximal elements. We call the set A_1', consisting of the maximal elements in P, the *maximal layer*. The maximal layer plays an important role in computational geometry, especially in the two dimensional case. Consequently, the asymptotics of the maximal layer have been studied in depth in the case $d = 1$, [22, 10]. We can consider D_{max}, the set of maximal points of the domain D. Generically, D_{max} will be a spacelike curve. We can consider its length $L(D_{max})$ w.r.t. the metric $-g$ for which it is timelike. We can restate the main result of [22] as saying that if $vol(D) = 1$, w.h.p. the size of the maximal layer will be approximately $\sqrt{\frac{\pi}{2}}L(D_{max})\sqrt{n}$, if $L(D_{max}) > 0$. This result generalizes to any dimension if we replace $L(D_{max})$ by $\text{Vol}(D_{max})$ and $\sqrt{\pi/2}$ by a constant, which is currently unknown. It is shown in [10] that the error term has a normal distribution with order of magnitude $n^{1/4}$. Some examples of finite volume, non-compact domains, for which the above asymptotic estimate does not hold are also presented. From the point of view of spacetime geometry, the problem is that in the counter examples the length function w.r.t. $-g$ is not upper semi-continuous at D_{max}, a problem which does not occur when D is compact.

The case where $\text{Vol}(L_{max}) = 0$, as occurs in intervals is also of interest. It has been shown in [44] that the number of maximal points in an interval is asymptotically

$$\frac{2}{(d+1)sin(\frac{2\pi}{d+1})\Gamma(\frac{2d}{d+1})}n^{\frac{d-1}{d+1}}$$

when $d > 1$. The constant term is related to the asymptotics of certain hypergeometric functions. It is interesting to contrast the result with the case of the simplicial light cone partial order on the cube, [9, 14, 15]. The expected size of the maximal layer in the simplicial case is $\frac{1}{d!}\ln(n)^d$. The large difference in results may be attributed to the nature of the light cones which is rounded in the first case, while having simplicial faces in the second case along which maximal points cluster. It is interseting to point out that the computation in [44] is motivated by an attempt to explain the entropy bounds which arise in modern cosmology.

For additional reconstruction results the reader may consult [36].

2. A list of problems

We present a few problems from different fields of applications.

1) **Scheduling of I/O (input/output) requests in disk drives:** Modern disk drives such as the one in a PC or laptop can handle multiple requests to read and write data. They can rearrange the order of requests so as to handle them more efficiently. In practice this is a very important feature and in some cases can double the performance of the disk drive. The problem of finding the optimal order of the requests is known as the *disk scheduling* problem. We may also ask for the amount of time that will be required to handle all the requests. This is known as *service time*. Since disk drives rotate at a constant speed we may ask equivalently for the number of disk rotations which are required to handle all the requests. We refer the reader to [2, 48, 29] for more information on disk scheduling.

As a toy model for this problem we may consider the simplistic case in which we assume that the read/write head of the disk accelerates and decelerates instantaneously. Under these assumptions Andrews, Bender and Zhang (ABZ) gave an algorithm which provides an optimal ordering of requests, [**2**].

What is the expected service time of the ABZ algorithm for n requests?

2) **Connecting-the-dots via Lifschitz functions:** The following problem was introduced in [**3**] in the context of estimating parameters for a multi-scale pattern recognition algorithm.

Given a set of n points sampled via some density in a bounded domain in $d+1$ dimensional Euclidean space.

What is the maximal number of points which can lie on the graph of a Lifschitz function (on the last d coordinates) with Lifschitz constant 1?

3) **Airplane boarding:** We consider passengers boarding an airplane. The input parameters for the process are

- u – the average amount of aisle length occupied by a passenger;
- w – the distance between successive rows;
- b – the number of passengers per row;
- D – the amount of time (delay) it takes passengers to clear the aisle, once they have arrived at their designated row;
- An airline boarding policy. For example, the policy may first allow passengers from the back third of the airplane to board, followed by the front third and finally the middle third. Another choice which is very popular is to board passengers from the back to the front, say, the back third, followed by the middle third and finally the front third.

Given u, w, b, D what is a good boarding policy for an airline to pursue?

3. Problem solutions

We show how the problems presented above are linked to discrete spacetime.

1) **Scheduling of I/O (input/output) requests in disk drives:** Disk locations can be represented in polar coordinates (r, θ), say with $1 \le r \le 2$ and $0 \le \theta \le 2\pi$ and $(r, 0)$ identified with $(r, 2\pi)$. We are also given a distribution $p(r, \theta)$ which reflects the popularity of different data in the disk. We can then assume that the data requests are sampled w.r.t. p. We assume for simplicity that the maximal radial speed of the disk is 1. We may then consider the Lorentzian metric

$$(9) \qquad ds^2 = 2p(r, \theta)(dr^2 - d\theta^2)$$

on the disk. Spacelike curves w.r.t. this metric correspond to possible motions of the disk head in a single rotation. Since the volume form for this metric is proportional to $p(r, \theta)$ we can think of n data requests as forming a discrete spacetime. We may apply the peeling process to this discrete spacetime. The ABZ algorithm has a very natural interpretation in terms of the peeling process. The i'th layer A_i is precisely the set of requests that will be handled in the i'th rotation of the disk. The service time will then be the number of layers in the peeling process which according to the results of the previous section is asymptotic to $L(C)\sqrt{n}$ where $L(C)$ is the length

of the maximal curve in the model. Since the boundary of the model is spacelike, the maximal curve will be a geodesic in this case. In disk drives, files are laid out in consecutive tracks, each track corresponding to a certain fixed radius r. This type of data arrangement naturally leads to a popularity density p which depends only on r, namely $p = p(r)$. In this case symmetry considerations show that the maximal curve is any curve which fixes the angle θ and calculating the length we get that the service time concentrates around $\sqrt{2} \int_0^1 \sqrt{p(r)} dr \sqrt{n}$. Since we have an infinite number of maximal curves the magnitude of the error term $\Delta_{D,n}$ can be somewhat larger than expected. When the distribution p is the uniform distribution, it can be shown using results from [30, 34] that $\Delta_{D,n}$ has order of magnitude $n^{1/6} \log(n)^{2/3}$. We expect that the variance will actually be smaller in this case, in analogy with extreme value distributions. For more information, see [6, 7].

We need to say a few words about the (non) realism of this model. The assumption that the disk does not accelerate or decelerate is far from true. The model which we have used predicts that a disk can handle an order of magnitude of $\sqrt{n}$ requests per rotation, a real disk will not handle more than 2-3 requests per rotation regardless of the number of requests, due to some fixed time delays. Nonetheless, the type of algorithms that are used in this case (ABZ) are indicative of successful algorithms in more complex and realistic models, therefore one does gain an understanding of the full problem by considering these simple cases.

2) **Connecting-the-dots via Lifschitz functions:** The relation between this problem and spacetime geometry becomes clear once we note that by definition the graph of a Lifschitz function is a spacelike hypersurface in Minkowski space. In order to think of the points as points in a discrete spacetime we have to match the volume form of the metric to that of the point distribution so we introduce the conformally flat metric

$$(10) \qquad ds^2 = p(x,t)^{2/(d+1)}(dx_0^2 - \sum_{i \geq 1} dx_i^2).$$

By theorem 1.3 the largest number of points will be picked in an ε neighborhood of a maximal hypersurface in the domain and the number of points will be proportional to the volume of the hypersurface.

3) **Airplane boarding:** We represent passengers by their position in the boarding queue x and their row number y, both normalized to produce a point (x,y) in the unit square. The airline boarding policy (and the way passengers react to it) leads to a distribution $p(x,y)$ on the joint row/queue position. For example, assume that the policy is to board the back third of the airplane, followed by the front third and finally the middle third. The corresponding distribution $p(x,y)$ is given by $p(x,y) = 3$ on the sub-squares $[0, 1/3] \times [2/3, 1]$, $[1/3, 2/3] \times [0, 1/3]$, $[2/3, 1] \times [1/3, 2/3]$, and $p(x,y) = 0$ elsewhere. The back-to-front policy with 3 equal group sizes will lead to a distribution with $p(x,y) = 3$ on the subsquares $[0, 1/3 \times [0, 1/3]$, $[1/3, 2/3] \times [1/3, 2/3]$ and $[2/3, 1] \times [2/3, 1]$.

We define $k = \frac{ub}{w}$. k measures the congestion of passengers in the airplane. It is the reciprocal of the portion of all passengers which can comfortably stand along the entire length of the airplane aisle. In real airplanes, k equals roughly 4. We also define $\alpha(x,y) = \int_y^1 p(x,z)dz$. α measures the density of passengers which want

to pass by the passenger (x, y) to get to their seats. Given p and k we define the Lorentzian metric

$$(11) \qquad ds^2 = 4D^2 p(x, y)(dx(dx + k\alpha(x, y)\, dy))$$

Since $p(x, y)$ is proportional to the volume form, we can think of the passengers as a discrete spacetime sampled from the unit square equipped with this metric. It can be shown, [4] that the causal relation $\prec$ associated with the metric (11) asymptotically coincides with the notion of passengers blocking other passengers from getting to their seats. Consequently, the boarding process roughly coincides with the bottom-to-top peeling process applied to this discrete spacetime. The first layer A_1, corresponds to passengers which can sit down without being blocked by other passengers. More generally, the set A_i asymptotically corresponds to the passengers which can sit unobstructed once the passengers in $\cup A_j$, $j < i$, have sat down. The boarding time is then given by $L(C)\sqrt{n}$, where C is the longest (maximal) curve in the unit square with the metric given in (11). We must admit that when the number of passengers is 200 or so the maximal curve does not follow the geodesic well. Furthermore, it turns out that large portions of the maximal curve typically lie on the boundary of the support of $p(x, y)$. This causes significantly greater variations in boarding time and makes the error term $\Delta_{D,n}$ much larger, [4]. Nonetheless, rather miraculously the ranking of boarding policies which is produced by the estimate $\ell(C)\sqrt{n}$ nearly coincides with the rankings given by very detailed computer simulations of the boarding process, [5, 50]. The methods above led to a detailed understanding of back-to-front policies as practiced by many airlines. Such policies are parameterized by the infinite dimensional simplex, consisting of sequences $r_0 = 0 < r_1 < \ldots < r_{m-1} < 1 = r_m$, for any positive integer m. Given such a sequence, n passengers, with d passengers per row, the corresponding policy boards passengers in rows $r_{m-1}n/d$ to $r_m n/d = n/d$, followed by passengers in rows $r_{m-2}n/d$ to $r_{m-1}n/d$ and so on. Using these methods it was shown in [8] that for $k > 1$ the expected boarding time of any back-to-front policy is at most

$$(12) \qquad \frac{\sqrt{k} + \frac{1 - \ln(2)}{\sqrt{k}}}{\sqrt{k} - 1}$$

faster than random boarding (no airline policy). For the realistic value of $k = 4$ it shows that we cannot obtain a speed up of more than 20% in boarding time using back-to-front and in fact most back-to-front policies yield performance which is worse than random. The estimate given in (12), which is based on constructing long (but not maximal) curves whose length satisfies a recursive inequality, is not tight. More refined calculations show that the optimal back-to-front policy is only about 8% better than random boarding, [31]. The policy first boards about 20 % of the passengers from the back of the airplane and then the rest of the passengers.

References

[1] D. Aldous and P. Diaconis, Longest increasing subsequences: From patience sorting to the Baik-Deift-Johansson theorem, *Bull. of the AMS*, vol. 36, 413-432, 1999.

[2] M. Andrews, M.A. Bender, and L. Zhang, New algorithms for the disk scheduling problem, *Algorithmica*, vol. 32, 277-301, 2002.

[3] E. Arias-Castro, D.L. Donoho, X. Huo, and C.A. Tovey, Connect-the-dots: How many random points can a regular curve pass through?, *Advances in applied probability* vol. 37, 571-603, 2005. Correction, vol. 38, 579, 2006.

[4] E. Bachmat, D. Berend, L. Sapir, S. Skiena, and N. Stolyarov, Analysis of airplane boarding via space-time geometry and random matrix theory, *Journal of physics A: Mathematical and general*, Vol. 39, L453-L459, 2006.

[5] E. Bachmat, D. Berend, L. Sapir, S. Skiena and N. Stolyarov, Analysis of airplane boarding times, 2006. Available at www.cs.bgu.ac.il/∼ebachmat/#pubs.

[6] E. Bachmat, Average case analysis for batched disk scheduling and increasing subsequences, *Proc. of the Symp. on Theo. of Comp. (STOC)*, Montreal, 2002.

[7] E. Bachmat, Average case analysis of disk scheduling, increasing subsequences and space-time geometry, *Algorithmica*, vol. 49, 212-231, 2007.

[8] E. Bachmat and M. Elkin, Bounds on the performance of back-to-front boarding policies. Available at www.cs.bgu.ac.il/∼ebachmat/#pubs.

[9] Z. Bai, L. Devroye, H. Hwang and H. Tsai, Maxima in hypercubes, *Random structures and algorithms*, vol. 27, 290-309, 2005.

[10] Z. Bai, H. Hwang, W. Liang and T. Tsai, limit theorems for the number of maxima in random samples from planar regions, *Electronic journal of probability*, vol. 6, paper 3, 2001.

[11] J. Baik, P.A. Deift, and K. Johansson, On the distribution of the length of the length of the longest increasing subsequence of random permutations, *Journal of the AMS*, vol. 12, 1119-1178, 1999.

[12] J. Baik and E. Rains, Algebraic aspects of increasing subsequences, *Duke journal of mathematics*, vol.109, 1-61, 2001.

[13] J. Baik and E. Rains, The asymptotics of monotone subsequences of involutions, *Duke journal of mathematics*, vol.109, 205-281, 2001.

[14] Y. Baryshnikov, Supporting-points processes and some applications, *Probability theory and related fields*, vol. 117, 163-182, 2000.

[15] B. Berezovskii and S. Travkin, Supervision of queues of requests in computer systems, *Automation and remote control*, vol. 36, 1719-1725, 1975.

[16] B. Bollobas and G. Brightwell, Box-spaces and random partial orders, *Transactions of the AMS*, vol. 324, 59-72, 1991.

[17] G. Brightwell and R. Gregory, Structure of random discrete space-time, *Physical review letters*, vol. 66, 260-263, 1991.

[18] L. Bombelli, J. Lee, D. Meyer and R.D. Sorkin, Spacetime as a causal set, *Physical review letters*, vol. 59, 521-524, 1987.

[19] F. Colaiori and M.A. Moore, Upper critical dimension, dynamic exponent, and scaling functions in the mode-coupling theory for the Kardar-Parisi-Zhang equation, *Physical review letters*, vol. 86, 3946-3949, 2001.

[20] P. Deift, Universality for mathematical and physical systems, plenary talk delivered at the ICM 2006 in Madrid, to appear in the ICM 2006 proceedings, math-ph/0603038.

[21] J.D. Deuschel and O. Zeitouni, Limiting curves for iid records, Annals of probability, vol. 23, 852-878, 1995.

[22] L. Devroye, Records, the maximal layer, and uniform distributions in monotone sets, *Computers, mathematics and applications*, vol. 25, 19-31, 1993.

[23] F. Dowker, J. Henson and R.D. Sorkin, Quantum gravity phenomenology, Lorentz invariance and discreteness, *Mod. Phys. Lett.*, vol. A19, 1829-1840, 2004.

[24] R. Dudley, Lorentz-Invariant Markov processes in relativistic phase space, *Arkive Mat. Astro. Fys*, vol. 6, 241-268, 1965.

[25] R. Dudley, A note on Lorentz-Invariant processes. *Arkive Mat. Astro. Fys.*, vol. 6, 575-581, 1965.

[26] R. Dudely, Assymptotics of some relativistic Markov processes, *Proceedings of the national academy of science USA*, vol. 70, 3551-3555, 1973.

[27] S. Hawking, A. King and P. McCarthy, A new topology for curved space-time which incorporates the causal, differential and conformal structures, *Jounal of mathematical physics*, vol. 17, 174, 1976.

[28] R. Ilie, G. Thompson and D. Reid, A numerical study of the correspondence between paths in a causal set and geodesics in the continuum, gr-qc/0512073.

[29] D. Jacobson and J. Wilkes, Disk scheduling algorithms based on rotational position, *HP labs technical report*, HPL-CSP-91-7rev1, 1991.

[30] K. Johansson, Transversal fluctuations for increasing subsequences on the plane, *Probability theory and related fields*, vol. 116, 445-456, 2000.

[31] V. Khachaturov and E. Bachmat, Optimal back-to-front boarding policies, preprint, 2007.

[32] M. Lassig, On growth, disorder, and field theory, *Journal of physics: Condensed matter*, vol. 10, 9905-9950, 1998.

[33] B.F. Logan and L.A. Shepp, A variational problem for random young tableaux, *Advances in mathematics*, vol. 26, 206-222, 1977.

[34] M. Lowe and F. Merkl, Moderate deviations for longest increasing subsequences: The upper tail, *Communications in pure and applied mathematics*, vol. 54, 1488-1520, 2001.

[35] E. Marinari, A. Pagnani, G. Parisi and Z. Racz, Width distributions and the upper critical dimension of Kardar-Parisi-Zhang interfaces, *Physical review E*, vol. 65, 026136 (2002).

[36] S. Major, D. Rideout and S. Surya, On recovering continuum topology from a causal set, gr-qc/0604124.

[37] D. Malament, The class of continuous timelike curves determines the topology of space-time, *Journal of mathematical physics*, vol. 18, 1399, 1977.

[38] D. Meyer, The dimension of causal sets, Ph.D thesis, M.I.T., 1988.

[39] J. Myrheim, Statistical geometry, CERN preprint TH-2538, 1978.

[40] R. Penrose, *Techniques of differential topology in relativity*, Regional conference series in applied mathematics vol.7, SIAM, 1972.

[41] M. Prahofer, Stochastic surface growth, *Ph.D thesis*, TU Munich, 2003. Available at http://www-m5.ma.tum.de/pers/praehofer/papers/

[42] M. Prahofer and H. Spohn, Universal distributions for growth processes in 1+1 dimensions and random matrices, *Physical review letters*, vol. 84, 4882-4885, 2000.

[43] D. Reid, The manifold dimension of a causal set: tests in conformally flat spacetimes, gr-qc/0207103.

[44] D. Rideout and S. Zohren, Evidence for the entropy bound from fundumentally discrete gravity, *Classical and quantum gravity*, 6195-6213, 2006.

[45] G. 't Hooft, Quantum gravity: A fundamental problem and some radical ideas, *in Recent developments in gravitation*, proceedings of the 1978 Cargese summer institute, edited by M. Levy and S. Deser, Plenum, 1979.

[46] M. Talagrand, Concentration of measure and isoperimetric inequalities in product spaces, *Publ. Math. I.H.E.S*, vol. 81, 73-203, 1995.

[47] C. tracy and H. Widom, Distribution functions for largest eigenvalues and their applications, *Proceedings of the ICM 2002*, vol. I, 587-596.

[48] M. Seltzer, Chen P. and J. Ousterhout, Disk scheduling revisited, *Proceedings of the Usenix technical conference*, 313-324, winter 1990.

[49] R.D. Sorkin, Causal sets: Discrete gravity, Notes for the Valdivia summer school, gr-qc/0309009.

[50] H. Van Landeghem and A. Beuselinck, Reducing passenger boarding time in airplanes: A simulation approach. *European J. of Operations Research* vol. 142, 294-308, 2002.

[51] A.M. Vershik and S.V. Kerov., Asymptotics of the plancherel measure of the symmetric group and the limiting form of young tables. *Soviet Math. Dokl.*, vol.18, 527-531, 1977.

DEPARTMENT OF COMPUTER SCIENCE, BEN-GURION U., BEER-SHEVA, ISRAEL, 84105, EMAIL: EBACHMAT@CS.BGU.AC.IL