

Chapter 6

Probabilistic Networks and Network Algorithms

Timothy Law Snyder

Department of Computer Science, Georgetown University, Washington, DC 20057, U.S.A.

J. Michael Steele

Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, U.S.A.

1. Introduction

The uses of probability in the theory of networks are extensive, and new applications emerge at an increasing rate. Still, when compared with the purely deterministic aspects of network theory, the part that calls upon probability theory is in its infancy. Certainly there are areas where the uses of probability have developed into a reasonably complete theory, but in many instances the results that have been obtained have to be regarded as fragmented and incomplete. This situation presents considerable opportunity for researchers, and the purpose of this chapter is to highlight aspects of the current state of the theory with an eye toward the developments and the tools that seem most likely to be of value in further investigations.

Probability enters into the theory of networks and network algorithms in several different ways. The most direct way is through probabilistic modeling of some aspect of the network. For example, in some freight management models the cost of transportation along the arcs of the network are modeled by random variables. In models such as these probability helps us grasp a little better a world that comes with its own physical randomness.

A second important way probability enters is through more stylized stochastic models where the aim is to provide deeper insight into our technical understanding of the methods of operations research. Here there is considerably less emphasis on building detailed models that hope to capture aspects of randomness that live in a specific application context; rather, the aim is to provide mathematically tractable models of reasonable generality that can be used to explore a variety of different computational or estimation methods. Among the types of issues that have been studied in such models are the efficacies of deterministic algorithms and of deterministic heuristic methods. Many of the 'average case' analyses of algorithms would fit into this second role for probability.

The third path by which probability enters into network theory is through randomized algorithms. This is the newest of the roles for probability, but it is a role that is of increasing importance. To make certain of the distinction that makes an algorithm ‘randomized,’ consider a version of depth-first search where one chooses the next vertex to be explored by selecting it at random from a set of candidates. Here one does not call on any modeling of the network, which may in fact be specified in a way that is completely deterministic. The use of probability here is purely technical in the sense that it is employed to serve our computations, not to model some external physical randomness, or even to capture the notion of an ‘average case.’

In the material that follows, one does well to keep these differing uses of probability in clear sight. Still, the distinctions may not always be pristine, mostly because two or more roles for probability can be present in the same problem. As an example, consider the computation or estimation of the reliability polynomial $R(p)$ of a network. Here one begins with a simple, physically motivated stochastic model. Given a specific graph intended to represent a communication network, one models the possibility of degraded communication in the network by allowing edges to ‘fail’ with probability p . The key problem is the determination of the probability $R(p)$ that for each pair of vertices a and b in the graph there exists a path from a to b that consists only of edges that have not failed. As the problem sits, it offers a simple but useful stochastic model, and one can go about the calculation or estimation of $R(p)$ by whatever tools are at one’s disposal. The multiplicity of roles for probability enters exactly when one starts to notice that there are randomized *algorithms* for the estimation of $R(p)$. This is just one example where there are several roles for probability in the context of a single problem.

There are even dicier instances where the role of probability in the design and analysis of algorithms starts to offer some ambiguity. For example, close cousins of the randomized algorithms are the algorithms that (a) assume that the input follows some stochastic model and (b) exploit that assumption in the computational choices that it makes. A natural example of this design is Karp’s algorithm for the Euclidean traveling salesman problem which we take up in Section 4. Such algorithms are fairly called *probabilistic algorithms*, but in the absence of internally generated random choices the best practice is to preserve the distinction made above and to avoid calling them randomized algorithms; though, admittedly, there is no reason to press for a rigid nomenclature.

The central aim of this chapter is to engage at least some aspect of each of the major roles for probability in the theory of network algorithms. When choices must be made, an emphasis is placed on those ideas one can expect to continue to be used and developed in the future. In Section 2 we engage the probability theory of network characteristics, where one mainly sees probability in either of the first two roles described above, as elements of either a physical or an idealized stochastic model. The section first develops the background for several inequalities that have evolved in the area of percolation theory. The FKG inequality is the best known and most widely used of these; but, as applications in percolation theory

have shown, the much newer BK inequality is also an instrument that belongs in every tool kit. The second part of Section 2 then looks at the computational problems of associated with more physical models of networks. In Section 3 we engage randomized algorithms in the context of several problems of concern to the basic themes of network theory. The first paradigm discussed there is one initiated in Karp & Luby [1985], which remains essential to the current technology of randomized algorithms. In Section 4 we focus on problems of geometric network theory. This is the area of network theory that seems to have progressed most extensively from the viewpoint of probability theory, but it also offers practical algorithmic insights on issues that have been of interest and concern even before there evolved an extensive theory of algorithms. The classic problems here include the behavior of traveling salesman tours, minimum spanning and Steiner minimum trees, and matchings.

2. Probability theory of network characteristics

There are three substantial probabilistic theories with lives of their own, yet which are intimately intertwined with the probability theory of networks. The most immediate of these is *network reliability*. This subject provides extensive investigation of the problem of calculating and bounding the probability of the existence of (s, t) -paths. Because network reliability is dealt with in a separate chapter of this volume and because the book of Colbourn [1987] provides an extensive treatment, we do not give many details of the subject. Still, in many probabilistic investigations of networks, one needs to keep in mind the existence and highlights of the large body of results provided by reliability theory, and several of the results reviewed here owe their motivation to the concerns of network reliability theory. A second theory that is closely connected to the theory of random networks is the theory of random graphs, which deals extensively with questions like the existence of long paths, connectedness, the existence of cycles, and many other issues that are of importance to the theory of networks. Since Bollobás [1985] provides an extensive treatment of the theory of random graphs, we do not go deeply into that subject here. A third closely-related field is percolation theory, and in many ways this subject has a claim on being the deepest of the three related fields. It certainly has been pursued extensively by a large number of mathematicians and physicists over a number of years.

2.1. Tools from percolation theory

In this section, we first recall what the aims of percolation theory have been over the years of its development. We then suggest some ways in which the theory may help researchers who are concerned with questions that are more at the heart of network theory. We will develop three elementary but central tools of percolation theory: the FKG inequality, the BK inequality, and Russo's formula. These powerful tools are the workhorses of percolation theory, yet they seem not

to be well known to researchers in the more general areas of stochastic network theory.

Percolation theory evolved from questions in physics that are themselves of many different flavors, including the magnetization of materials, the formation of crystals, the transport of electrons in special materials, sustenance of chemical reactions, and the flow of fluids. The latter offers perhaps the least compelling physics, but it provides the easiest metaphor and is often called upon for illustration.

We consider the classical d -dimensional rectangular lattice, Z^d , and for each vertex v of the lattice we join v by an edge to each of its $2d$ nearest neighbors (in the sense of the usual Euclidean metric). If we use the traditional language of percolation theory, these edges are called ‘bonds’. Bonds are viewed as being either ‘open’ or ‘closed’, and it is here where the probability modeling appears. To each bond e is associated an independent Bernoulli random variable X_e such that $P(X_e = 1) = p$ for some fixed $0 \leq p \leq 1$. The bonds for which $X_e = 1$ are regarded as being *open*, and the fundamental questions of the theory concern the components of lattice vertices connected by open edges. Among the main features that distinguish percolation theory from the theory of random graphs are the attention that is focused on subgraphs of the lattice and the interest that is focused on graphs with infinitely many edges.

A central quantity of interest in percolation theory is the *percolation probability* $\theta(p)$, defined as the probability that the origin is contained in an infinitely large connected component. One reason that $\theta(p)$ receives considerable attention is that it exhibits interesting critical phenomena that have close analogies with physical phenomena like the freezing of fluids. In particular, one can prove, for each dimension d , that there is a *critical constant* $p_c = p_c(d)$ depending on the dimension d such that $\theta(p) > 0$ if $p > p_c$ but $\theta(p) = 0$ if $p < p_c$. The work of Kesten [1980] culminated the efforts of a great many investigations and established the long conjectured result that $p_c(2) = 1/2$. This deep result required the development of techniques that would seem to offer useful insights for researchers in the theory of networks, and a well motivated exposition of Kesten’s theorem can be found in Chapter 9 of Grimmett [1989].

The FKG inequality

The first tool we consider is named the FKG inequality, in respect of the work of Fortuin, Kasteleyn & Ginibre [1971]. Even though we will not call on the full generality of their result, it is worth noting that the FKG inequality has a beautiful generalization due to Ahlswede & Daykin [1978], and the full-fledged FKG inequality has already found elegant applications to problems of interest in network theory. In particular, one should note the articles by Shepp [1982] and Graham [1983].

The version of the inequality that we develop is actually a precursor of the FKG inequality due to Harris [1960], but Harris’s inequality has the benefit of having very low overhead while still being able to convey the qualitative essence of its more sophisticated relatives. To provide a framework for Harris’s inequality, we

suppose that G is any graph and $\{X_e\}$ are identically distributed Bernoulli random variables associated with the edges of G . We think of the variables X_e as labels marking the edges of G that would be regarded in percolation theory as open edges. The random variables of interest in Harris's inequality are those that can be obtained as monotone non-decreasing functions of the variables $\{X_e\}$. In detail, if in a realization of the $\{X_e\}$ we change some of the $\{X_e\}$ that have value zero to have a value of one, then we require that the value of the function does not decrease. The classic example of such a variable is the indicator of an (s, t) -path of edges marked with ones. Harris's inequality confirms the intuitive fact that any pair X and Y of such monotone variables are positively correlated. Specifically, if X and Y are any non-decreasing random variables defined as functions of the edge variables X_e of G , then one has

$$E(XY) \geq E(X)E(Y).$$

This inequality is most often applied in the case of indicator functions. Since we will refer to this case later, we note that we can write Harris's inequality as

$$P(A \cap B) \geq P(A)P(B)$$

for all events A and B that are non-decreasing functions of the edge variables.

One can prove Harris's inequality rather easily by induction. If we write $X = f(\eta_1, \eta_2, \dots, \eta_n)$ and $Y = g(\eta_1, \eta_2, \dots, \eta_n)$, where f and g are monotonic and the $\{\eta_i\}$ are independent Bernoulli random variables, then by conditioning on η_n , we see that it suffices to prove Harris's inequality just in the case of $n = 1$. In this case we see that, for $q = 1 - p$,

$$\begin{aligned} EXY - EX \cdot EY &= \\ f(1)g(1)p + f(0)g(0)q - (f(1)p + f(0)q)(g(1)p + g(0)q) \end{aligned}$$

and since this factors as

$$pq\{f(1) - f(0)\}\{g(1) - g(0)\} \geq 0,$$

we obtain Harris's inequality.

One of the nice consequences of Harris's inequality is the fact that if m non-decreasing events A_1, A_2, \dots, A_m with equal probability have a *union* with large probability, then, *all* the events A_i must have fairly large probability. This so-called 'square root trick' noted in Cox & Durrett [1988] formally says that for each $1 \leq i \leq m$, we have

$$P(A_i) \geq 1 - \{1 - P(\bigcup_{j=1}^m A_j)\}^{1/m}.$$

The proof of this inequality requires just one line where Harris's inequality provides the central step:

$$1 - P(\bigcup_{j=1}^m A_j) = P(\bigcap_{j=1}^m A_j^c) \geq \prod_{j=1}^m P(A_j^c) = (1 - P(A_1))^m.$$

To appreciate the value of this inequality, one should note that without the assumption that the $\{A_i\}$ are monotone, one could take the $\{A_i\}$ to be a partition of the entire sample space, making the left side equal to $1/m$, while the right side equals one. We see therefore that the FKG inequality helps us extract an important feature of monotone events. As a point of comparison with a more combinatorial result, one should note that the square root trick and the local LYM inequality of Bollobás and Thomason [cf. Bollobás, 1986] both address the way in which probabilities of non-decreasing sets (and their ideals) can evolve. For further results that call on the FKG and Harris inequalities one should consult Graham [1983] and Spencer [1993].

The BK inequality

The insights provided by the FKG and its sibling inequalities are valuable, but they are limited. The inequalities often just provide rigorous confirmation of intuitive results that one can justify by several means. A much deeper problem arises when one needs an inequality that goes in a direction opposite that of the FKG inequality. For this problem, the progress is much more recent and less well known.

As one can show by considering any dependent, non-decreasing events A and B , there is no hope of simply reversing the FKG inequality. In fact, the same examples can show that additional assumptions on A and B that fall short of independence are of no help, so some sort of additional structure, or some modification is needed for $A \cap B$. Van den Berg & Kesten [1985] discovered that the key to a useful reversal of the FKG inequality rests on a strengthening of the notion of $A \cap B$. The essence of their idea is that the event of A and B both occurring needs to be replaced with that of ' A and B both occurring, but for different reasons' or, as we will shortly define, A and B occurring disjointly.

The technical definition of disjoint occurrence takes some work, but it is guided by a canonical example. If A corresponds to the existence of an (s, t) -path and B corresponds to the existence of an (s', t') -path, then $A \cap B$ needs to be replaced by the event corresponding to the existence of (s, t) - and (s', t') -paths that *have no edge in common*. To make this precise in a generally applicable way, we have to be explicit about the underlying probability space. To keep ourselves from straying too far from network applications, we let Ω denote the set of $(0, 1)$ -vectors (x_1, x_2, \dots, x_m) , where m is the number of elements in a set S of edges that are sufficient to determine the occurrence of A . In many problems m cannot be bounded by anything sharper than the number of edges of G , but the bound can be useful even in such cases.

We define a measure on Ω via the Bernoulli edge variables X_e taken in some fixed order, so Ω taken with our probability measure P give us a product measure space $\{\Omega, P\}$. We now define the set $A \circ B$, the *disjoint occurrence of non-decreasing events A and B* , as follows:

$$A \circ B = \{ \omega : \text{there exists } \omega_a \in A \text{ and } \omega_b \in B \\ \text{such that } \omega_a \cdot \omega_b = 0, \text{ and } \omega \geq \omega_a \text{ and } \omega \geq \omega_b \}.$$

Here, we use $\omega_a \cdot \omega_b$ to denote the usual inner product between vectors, so the combinatorial meaning of the last condition is that ω_a and ω_b share no 1's in their representation. In other words, for non-decreasing events A and B , ω_a and ω_b are able to bear respective witness that A and B occur, but they can base their testimony on disjoint sets of edges.

The BK Inequality. *If A and B are non-decreasing events in $\{\Omega, P\}$, then*

$$P(A \circ B) \leq P(A)P(B).$$

The systematic use of the BK inequality is just now becoming widespread even in percolation theory proper. In Grimmett [1989] one finds many proofs of older results of percolation theory that are rendered much simpler via the BK inequality.

Russo's formula

The last of the percolation theory tools that we will review is a formula due to Russo [1981] that tells how the probability of a non-decreasing event changes as one changes the probability of the events $\{X_e = 1\}$. To state the formula, we suppose as before that we have a graph G with edges that are 'open' with probability p in such a way that the indicator variables X_e are independent and identically distributed. In this context we will require that G is finite, and, to emphasize the use of p as a parameter, we will denote the governing probability measure by P_p .

Now, if A is any non-decreasing event, we introduce a new random variable N_A that we call 'the number of edges that are pivotal for A '. Formally, we define $N_A(\omega)$ as follows: (a) If $\omega \notin A$, then $N_A(\omega)$ is zero, and (b) if $\omega \in A$, then $N_A(\omega)$ equals the number of edges e such that, in the representation of ω as a $(0, 1)$ -vector of edge indicators $\omega = (x_1, x_2, \dots, x_m)$, we have $x_e = 1$, but, if we change x_e to 0 to get a new vector ω' , then $\omega' \notin A$. In the latter case, we say that e is *pivotal* for A .

Russo's formula. *If A is any non-decreasing event defined on the Bernoulli process associated with a finite graph G , and if N_A denotes the number of edges that are pivotal for A , then*

$$\frac{d}{dp} P_p(A) = E_p(N_A).$$

This beautiful and intuitive formula can be used in many ways, but it is often applied to show that $P_p(A)$ cannot increase too rapidly as p increases. To see how one such bound can be obtained in a crude but general context, we first note that the differential equation of Russo's formula can be rewritten in integrated form for $0 \leq p_1 < p_2 \leq 1$ as

$$P_{p_2}(A) = P_{p_1}(A) \exp\left(\int_{p_1}^{p_2} \frac{1}{p} E_p(N_A | A) dp\right).$$

If there is a set $S = \{e_1, e_2, \dots, e_m\}$ of m edges such that the occurrence of A can always be determined by knowledge of S , then the integral representation and the trivial bound

$$P_p(e \text{ is pivotal for } A \mid A) \leq 1$$

provide a general inequality that bounds the rate of growth of $P_p(A)$ as a function of p :

$$P_{p_2}(A) \leq \left(\frac{p_2}{p_1}\right)^m P_{p_1}(A).$$

2.2. Distributional problems of random networks

In percolation theory the random variables associated with the edges are invariably Bernoulli, but in the network models that aim to model physical systems the network ingredients often are modeled by random variables with more general distributions, and the central questions in such models concern the distributions of larger network characteristics. Sadly, many of these distributional questions are analytically infeasible. Moreover, in many cases of practical interest these same questions are computationally intractable as well. We will illustrate some of the technology that has been developed for such problems by considering the problem of determining the distribution of the minimum-weight path from source to sink in a network with random edge weights.

Calculation of the distribution of the shortest paths

Formally, we let $G = (V, E)$ be an acyclic network with source vertex s and sink t , where edge weights are represented by independent random variables W_e for all $e \in E$. The stochastic quantity of interest is the distribution of the random variable $L(G)$, denoting the length of a shortest (s, t) -path in G . Valiant [1979] showed that the problem of determining the distribution of $L(G)$ is in general *NP*-hard, so at a minimum one must look to approximation methods. One natural approach to the distribution problem is to try to exploit the independence of the edge weights through the use of cut sets. This idea forms the basis of the simulation method of Sigal, Pritsker & Solberg [1979, 1980]. To describe their method for building a simulation estimate for $P(L(G) \geq t)$, we first let $C = e_1, e_2, \dots, e_k$ be an *exact cut* in G , that is, we take C to be a set of edges such that every (s, t) -path in G shares exactly one edge with C . Such a cut always exists, and it offers us a natural method for exploiting the independence of the X_e . The key observation is that the edges of C induce a natural partition of the (s, t) -paths of G .

For each $1 \leq i \leq k$ and each $e_i \in C$ we let p_i be the set of all (s, t) -paths that contain e_i . Now, for any $t \in \mathbb{R}$, we consider the random variable defined by the conditional probability

$$R = P(L(G) \geq t \mid W_e, e \in E - C).$$

Since R satisfies $ER = P(L(G) \geq t)$, if we let r be the sample value of R based on a realization $\{w_e\}$ of $\{W_e : e \in E - C\}$, then by independence we have

$$\begin{aligned} r &= P(L(G) \geq t \mid w_e, e \in E - C) \\ &= P\left(\sum_{\substack{e \in p \\ e \neq e_i}} w_e + W_{e_i} \geq t \text{ for all } p \in p_i \text{ and for all } e_i \in C\right) \\ &= \prod_{i=1}^k P\left(W_{e_i} \geq t - \min_{p \in p_i} \left(\sum_{\substack{e \in p \\ e \neq e_i}} w_e\right)\right). \end{aligned}$$

Since the right hand side can be computed from the known distribution of the W_{e_i} , an estimate of $P(L(G) \geq t)$ is given by $n^{-1} \sum_{1 \leq i \leq n} r_i$, where the r_i are given by independent simulations of r . For some simple networks this process gives an effective method for the simulation estimation of the distribution of $L(G)$, and, at a minimum, the use of r is substantially more efficient than the completely naive estimation of $P(L(G) \leq t)$ via independent simulations of $L(G)$.

This method is reasonably crude, but with proper implementation it can provide answers in some situations of importance. Certainly, the computation of $\min_{p \in p_i} (\sum_{\substack{e \in p \\ e \neq e_i}} w_e)$ for $1 \leq i \leq k$ should not be conducted naively since there can be an exponential number of (s, t) -paths in G . There are even moderately sized networks for which an exhaustive evaluation of the required sums is computationally prohibitive. One does much better to note that once the edges in $E - C$ have been sampled, the lengths of the shortest paths from s to e_i and from e_i to t can be computed for each i by using an appropriate deterministic single-source shortest path algorithm, such as that of Dijkstra [1959], or more recent refinements. Dijkstra's algorithm is easy to implement, has low overhead, and takes $O(|V|^2)$ steps in the worst case to compute all the required path lengths. Having the lengths of the shortest paths from s to the e_i and from the e_i to t allows the computations of the minima in the representation for r to be obtained in at most $O(|C|^2)$ additional steps.

The problem of estimating the distribution of $L(G)$ offers a typical instance of the trade-off one often meets in simulation estimations. First, there is a desire to have an efficient estimate of ER for which we would like a cut that provides for a low variance of R . This is a kind of efficiency that helps us minimize the number of independent realizations one must take in the simulation. Second, we would like to have efficiency in the computation of the estimate in the sense of computing the shortest (s, e_i) - and (e_i, t) -paths. The trade-off that faces us is that as $|C|$ increases, the variance of R decreases, but as the cut size increases so does the cost of computing shortest paths and minima. There are often many different exact cuts on which one can base the simulation estimation of $P(L(G) \geq t)$ and the proper choice of the cut is an important design consideration.

Other distribution problems of random networks

Other studies have undertaken the difficult task of determining distributions of flows in networks with random capacities. Among these is the paper of Grimmett & Welsh [1982], which considered maximum flows in networks with independent and identically distributed capacities. Grimmett and Welsh found limit theorems for the cases where the networks are either complete graphs or branching trees. In subsequent work, Frieze & Grimmett [1985] looked at the shortest path problem under general independent models, and Kulkarni [1986] studied shortest paths in networks with exponentially distributed edge lengths. One point that emerges from these works is that the probability theory of network characteristics offers many individual problems of considerable challenge. So far there seems to have not been any attempt at providing the framework for the general theory of such characteristics. With the insights of several special problems in hand, perhaps it is time that work on a more general investigation was begun.

3. Probabilistic network algorithms

In this section we first provide an introduction to a general approach of Karp & Luby [1985] for the design of randomized algorithms. We then illustrate their method by showing how one can put the problem estimation of multiterminal network reliability into their framework. We then review briefly two recent randomized algorithms for maximum network flow. Finally, we review some of the work on randomized algorithms for perfect matching and maximal matching in graphs.

3.1. Karp–Luby structures for randomized algorithms

Karp & Luby [1985] provided a framework for randomized algorithms that is useful in a broad range of applications and which specifically offers an effective approach to some problems of network reliability. Their approach begins abstractly with a set S and a weight function $a : S \rightarrow \mathbb{R}^+$ which we then use to provide a weight for any $A \subset S$ by taking $a(A) = \sum_{x \in A} a(x)$. Clearly there are many important problems that can be framed in terms of the calculation of $a(A)$ for appropriate choices of a , S , and A ; but, as one must suspect, we will have to impose some additional structures before this framework can show its value. We call (S, R, a) a *Karp–Luby Monte Carlo structure* if $R \subset S$ and we have the following three properties: (1) the ‘total weight’ $a(S)$ is known, (2) there is a ‘sampling algorithm’ that selects an item x at random from S according to the probability $a(x)/a(S)$ with independent selections at each invocation of the sampling algorithm, and (3) there is a ‘recognition algorithm’ that can test if a given element x of S is also an element of R .

For any such structure (S, R, a) one can estimate the weight $a(R)$ in the most straightforward way imaginable. One just selects n independent random elements X_1, X_2, \dots, X_n of S by the sampling algorithm. Letting Y_i be 1 or 0

accordingly as $X_i \in R$ or not, one then takes as an estimator of $a(R)$ the value $\bar{Y} = a(S)(Y_1 + Y_2 + \dots + Y_n)/n$. As a consequence of the traditional Bernstein tail estimates of the binomial distribution, for any $\epsilon > 0$ and $\delta > 0$ we have

$$P\left(\frac{|\bar{Y} - a(R)|}{a(R)} > \epsilon\right) \leq \delta$$

provided that

$$n > \frac{9}{5} \epsilon^{-2} \log\left(\frac{2}{\delta}\right) \frac{a(S)}{a(R)}.$$

The punch line here is that once we are able to frame a problem in terms of a Karp–Luby structure, we can determine a δ - ϵ approximation in the sense of the preceding probability bound. Moreover, we can bound the expected computational cost of the algorithm by a polynomial in the parameters ϵ^{-1} , $\log(1/\delta)$, and the sensitivity ratio $a(S)/a(R)$.

3.2. Karp–Luby structures for network reliability

The multiterminal network reliability problem is a stylized model for communication reliability that has been studied from many perspectives, and it offers a good example of how one fits a natural problem into the framework of Karp–Luby structures. Given a connected graph $G = (V, E)$ and a special set of ‘terminal vertices’ $T = \{t_1, t_2, \dots, t_k\} \subset V$, the motivating issue of multiterminal network reliability is to model the possibility of communication between the elements of T under random degradation of the network links. The probability modeling calls for a function $p : E \rightarrow [0, 1]$ that is viewed as giving for each $e \in E$ the probability $p(e)$ that the edge e is ‘bad.’ Under the assumption that the edges are made good or bad according to independent choices governed by p , the key problem is to determine the probability that for all pairs of elements of the set of terminals there is a path between them that consists only of good edges.

More formally, we consider the set of all mappings $s : E \rightarrow \{0, 1\}$ as the elements of our probability space, and we take the interpretation of this function as an assignment of a label of 0 on the good edges and 1 on the bad edges. The probability of a specific state s thus is given by $P(s) = \prod_{e \in E} p(e)^{s(e)} (1 - p(e))^{1-s(e)}$. The computational challenge is to calculate the probability that there is some pair of terminal vertices for which there does not exist a path between them in the graph consisting of the vertex set V and the set of all edges of G which are labeled ‘good’. We call a state for which this event occurs a *failing state*, and we let F denote the set of all states s which have failure.

To provide a Karp–Luby structure so that we can use the strategy discussed in the preceding section, we first need the notion of a *canonical cut*. Let $s \in F$ be any failing state, and let $G(s) = (V, E(s))$ where $E(s)$ is the set of good edges for the state s . For any $1 \leq i \leq k$ we then let $C_i(s)$ denote the connected component of $G(s)$ that contains the terminal t_i . Since $s \in F$ there is some i for which $C_i(s)$ is not all of G and, further, because of the assumption that the full graph

$G = (V, E)$ is connected, there is at least one such $C_i(s)$ for which the graph induced by the removal of all the vertices of $C_i(s)$ from $G = (V, E)$ is connected. We let $i^*(s)$ denote the least such index i , and finally we let $g(s)$ denote the set of edges that have exactly one endpoint in $C_{i^*}(s)$. The set $g(s)$ is a T -cut in that it separates two terminals of T in the graph $G(s) = (V, E(s))$, and we call $g(s)$ the *canonical cut* for the state s .

We now have the machinery to specify the Karp–Luby structure for the multi-terminal reliability problem. Let S be the set of all pairs (c, s) where $s \in F$ is a failing state and c is a T -cut for which each edge of c fails in state s . The weight function associated with a pair $(c, s) \in S$ is taken to be the probability of the state s , so $a((c, s)) = P(s)$. Although this weight function ignores the first component of (c, s) , the presence of the first component turns out to be an essential in providing an effective sampling process. This choice of a and S permits us to write down a simple formula for $a(S)$. Since $a(S)$ is equal to the sum of all the probabilities of the states s where s fails for the cut c , we have

$$a(S) = \sum_{(c,s) \in S} a(c, s) = \sum_c \prod_{e \in c} p(e),$$

where the last sum is over all T -separating cut sets of $G = (V, E)$. The target set R is given by the set of all pairs $(g(s), s)$ where $s \in F$, and (S, R, a) will serve as our candidate for a Karp–Luby structure for the multiterminal network problem. To see the interest in this triple we first note that

$$a(R) = \sum_{s \in F} a(g(s), s) = \sum_{s \in F} P(s),$$

so the weight $a(R)$ corresponds precisely to the probability of interest.

For the effective use of (S, R, a) , it would be handiest if we had at our disposal a list L of all the T -separating cut sets of $G = (V, E)$. When the list is not too large, the formula given above provides a way to calculate $a(S)$. Similarly, we also have at hand an easy way to test if $s \in F$ by examining the failure of each of the cuts. To complete our check that the (S, R, a) leads to a Karp–Luby structure in this nice case, it only remains to check that sampling from S is not difficult.

To choose an element $(c, s) \in S$ according to the required distribution, we first choose at random a $c \in L$ according to the probability distribution $\prod_{e \in c} p(e)/a(S)$. We then select a state function s such that $s(e) = 1$ for all $e \in c$ and by letting $s(e) = 1$ or $s(e) = 0$ with probability $p(e)$ or $1 - p(e)$, respectively.

We have completed the verification that (S, R, a) satisfies the constraints required of a Karp–Luby structure, but for it to serve as the basis for an effective randomized algorithm we also need to have a bound on the sensitivity ratio $a(S)/a(R)$. In many multiterminal reliability problems a sufficiently powerful bound is provided by the following inequality of Karp & Luby [1985]:

$$\frac{a(S)}{a(R)} \leq \prod_{e \in E} (1 + p(e)).$$

Thus far we have given a reasonably detailed view of the Karp–Luby structure and how it can be applied to a problem of computational interest in network theory. The development recalled here so far has the shortcoming that it seems to require an explicit list of the T -cuts of the network, and that list must be reasonably short. Karp & Luby [1985] go further and show that there are cases where this requirement can be avoided; in particular they show that if G is a planar graph, then the program still succeeds even without explicitly listing the cut sets.

3.3. Randomized max-flow algorithms

The theory of network flows is to many people what the theory of networks is all about, and there are two recent contributions of randomized algorithms to this important topic that have to be mentioned here, even though this survey cannot dig deeply enough into them to do real justice. The first of these is the algorithm of Cheriyan & Hagerup [1989] for maximum flow in the context where all the arc capacities are deterministic. The Cheriyan–Hagerup algorithm produces a maximum flow for any (non-random) single-source, single-sink input network. The algorithm takes $O(|V||E|\log|V|)$ time in the worst case, although this happens with probability no more than $|V|^{-\alpha|V|^2}$, where α is any constant. Most important is that the Cheriyan–Hagerup algorithm takes $O(|V||E|+|V|^2(\log|V|)^3)$ expected time, which, being $O(|V||E|)$ for all but relatively sparse networks, compares favorably with all known strongly polynomial algorithms. The algorithm is also strongly polynomial in the sense that the running time bound does not depend on the edge-capacity data.

The Cheriyan–Hagerup algorithm builds on some of the best deterministic algorithms and takes a step forward by introducing randomization at key stages. The algorithm calls on scaling techniques in the spirit of Gabow [1985], Goldberg & Tarjan [1988], and Ahuja & Orlin [1987] and also employs *pre-push labeling*, another device of the Goldberg and Tarjan max-flow algorithm [cf. Ahuja, Magnanti & Orlin, 1991]. The randomness of the Cheriyan and Hagerup algorithm arises in how the network is represented at a given moment during the course of the algorithm. The model used for network representation is the adjacency list model, in which the neighbors of each $v \in V$ are maintained by a list associated with v . One of the key ideas of the Cheriyan–Hagerup algorithm is to randomly permute each adjacency list at the outset of the algorithm, then randomly permute the adjacency list of vertex v whenever the label of v is updated. The net effect of the permutation is to lower the expected number of relabeling events that the algorithm must carry out during each phase, lowering the expected running time. One further interesting aspect of the Cheriyan–Hagerup algorithm is that Alon [1990] has provided a device that derandomizes the algorithm in a way that is effective for a large class of graphs.

A more recent contribution of a randomized algorithm for max-flow has been provided in Karp, Motwani & Nisan [1993]. Given a realization of an undirected network with independent identically distributed random capacities, the algorithm finds a network flow that is equal in value to the optimum flow value with high

probability. The algorithm runs in linear time, which is significantly faster than the best known algorithms that are guaranteed to find an optimal flow.

The algorithm of Karp, Motwani, and Nisan is not simple, but at least some flavor for the design can be appreciated independently of the details. In the first stage of the algorithm, the max-flow problem on G is transformed to an instance of a probabilistic version of the transportation problem. The instance of the transportation problem is constructed so that its solution flow is forced to yield (1) a maximum flow that can be immediately transformed to a max-flow in G and (2) a flow that saturates the $(S, V - S)$ cut in G , where S is the set of sources. The second stage of the max-flow algorithm is a routine that attempts to solve the transportation problem. Here Karp, Motwani & Nisan [1993] introduce their so-called *mimicking method* which they outline in four steps: (1) before considering the realization of the random graph, consider instead the graph formed by replacing each random variable X_i with EX_i ; (2) solve the resulting deterministic problem; (3) consider now the realization of the random graph, and attempt to solve the problem by ‘mimicking’ the solution from (2); and (4) fine-tune the result to get the optimum solution. Even though these steps have engaging and evocative descriptions, there is devil in the details which in the end leads to delicate analyses for which we must refer to the original.

3.4. Matching algorithms of several flavors

Information about matchings has a useful role in many aspects of the theory of networks. Moreover, some of most effective randomized algorithms are those for matching, so this survey owes the reader at least a brief look at randomized matchings for algorithms and related ideas.

The key observation of Lovász [1979] was that one can use randomization to test effectively for the positivity of a determinant, and this test can be used in turn to test for the existence of a perfect matching in a graph. To sketch the components of the method we first recall that with the graph $G = (V, E)$ we can associate an adjacency matrix D by taking $d_{ij} = 1$ if $(i, j) \in E$ and zero otherwise. From the adjacency matrix we can construct the Tutte matrix T for G by replacing the above-diagonal elements d_{ij} by the indeterminants x_{ij} and the below-diagonal elements d_{ij} by the indeterminants $-x_{ij}$. The construction of T is completed by putting zeros along the diagonal. The theorem of Tutte, for which he introduced this matrix, is that G has a perfect matching if and only if $\det T \neq 0$.

The core of the idea for testing if G has a perfect matching is then quite simple. One chooses random numbers for the values of the x_{ij} and then computes the determinant numerically, a process that is not more computationally difficult than matrix inversion. The savings come here from the fact that the determinant in the indeterminant variables x_{ij} can have exponentially many terms, but to test that the polynomial is not identically zero we only have to see that it is non-zero at a point. Naturally, to be true to the values of honest computational complexity theory one cannot rely on computation with real numbers, but by working over a finite field one comes quickly to the conclusion that there is merit to the idea.

Lovász [1979] generalized Tutte's theorem and went on to provide an algorithm that takes advantage of the idea just outlined in order to find maximal matchings in a general graph. Rabin and Vazirani [1989] pressed this idea further and provided an algorithm that is faster than that of Lovász. A computational virtue of both the Lovász and Rabin-Vazirani algorithms is that they are readily implemented as parallel algorithms.

Another development from the theory of matching that has had wide-ranging impact on the theory of combinatorial algorithms is the introduction of the method of rapidly mixing Markov chains. The development evolving from Jerrum & Sinclair [1986, 1989] calls on the idea that if one runs a Markov chain for a long time, then its location in the state space is well approximated by the stationary distribution of the Markov chain. This idea can be used to estimate the number of elements in a complicated set, say, the set of all matchings on a graph, if one can find a chain on a set of states that includes the set of matchings and for which a Markov chain can be constructed that converges rapidly to stationarity. This idea has undergone an extensive development over the last few years. For a survey of this field we defer to the recent volume of Sinclair [1993].

The final observation about matching in random graphs that deserves space in the awareness of researchers in network theory is that algorithms based on augmenting paths are likely to perform much better than their worst-case measures of performance would indicate. These algorithms, which exhibit the fastest worst-case running times, are also fast in expectation, sometimes outperforming even the best heuristic algorithms. Many of the algorithms, including the algorithms of Even & Kariv [1975] and Micali & Vazirani [1980], run in linear expected time if the input graph is chosen uniformly from the set of all graphs. The reason behind this observation seems to be the expander properties of random graphs and the fact that in expander graphs one has a short path connecting any two typical points [cf. Motwani, 1989].

The proofs of these results come from an analysis of the lengths of augmenting paths. It is shown that, with high probability, every non-perfect matching in a random graph has augmenting paths that are relatively short. Since the bulk of augmenting path algorithms is spent carrying out augmentations, bounds on the lengths of augmenting paths translate to fast running times.

4. Geometric networks

One of the first studied and most developed parts of the theory of networks concerns networks that are embedded in Euclidean space. A *geometric network* is defined by the finite point set $S \subset \mathbb{R}^d$, with $d \geq 2$, and an associated graph, which is usually assumed to be the complete graph on S . The costs associated with the edges of the graph are typically the natural Euclidean lengths, though sometimes it is useful to consider functions of such lengths, for example, to take the cost of an edge to equal the square of its length. The central questions of the theory of geometric networks focus on the lengths of subgraphs; so for example, in the

traveling salesman problem, we are concerned with the length of the shortest tour through the points of S . Also of central interest in this development is the theory of minimum spanning trees, Steiner trees, and several types of matchings. The key result in initiating the probabilistic aspects of this developments is the classic Beardwood, Halton, and Hammersley theorem.

Theorem [Beardwood, Halton & Hammersley, 1959]. *If X_i , $1 \leq i < \infty$ are independently and identically distributed random variables with bounded support in \mathbb{R}^d , then the length L_n under the usual Euclidean metric of the shortest path through the points $\{X_1, X_2, \dots, X_n\}$ satisfies*

$$\frac{L_n}{n^{(d-1)/d}} \rightarrow \beta_{\text{TSP},d} \int_{\mathbb{R}^d} f(x)^{(d-1)/d} dx \text{ almost surely.}$$

Here, $f(x)$ is the density of the absolutely continuous part of the distribution of the X_i .

In addition to leading to algorithmic applications, the Beardwood, Halton, and Hammersley (BHH) theorem has led to effective generalizations, as well as new analytical tools. In this section, we review these tools, including the theory of subadditive Euclidean functionals, bounds on tail probabilities, related results in the theory of worst-case growth rates, and bounds on limit constants such as $\beta_{\text{TSP},d}$.

One elementary point that may help avoid confusion in the limit theory offered by the Beardwood, Halton, Hammersley theorem is the observation that it is of a much deeper character than $\Theta(n^{(d-1)/d})$ results for L_n , which only require that there exist positive constants a and b such that $an^{(d-1)/d} \leq L_n \leq bn^{(d-1)/d}$. The latter results are sometimes useful, but they are almost trivial in comparison, unless one presses for very good values for a and b . The stronger asymptotic result that $L_n/n^{(d-1)/d}$ converges to a constant requires entirely different techniques and typically leads to much different applications.

A second comment concerns uses to which one can put results such as the Beardwood, Halton, Hammersley theorem and its relatives. The use of the BHH theorem in the polynomial-time probabilistic TSP algorithm of Karp [1976, 1977] is one of the primary reasons results like the BHH theorem are studied today. Part of the charm of TSP is that it is NP -hard, and it has been studied from many heuristic and approximation perspectives. Karp's algorithm has a special place in the theory of algorithms because given any $\epsilon > 0$, its expected running time is almost linear and with probability one it produces a tour of length no more than $(1 + \epsilon)$ times the optimal tour length. Karp's algorithm played an important role in launching the field of probabilistic algorithms, and it certainly stimulated interest in the development of theorems that extend that of Beardwood, Halton, and Hammersley. Since then, theorems like the BHH theorem have been proved for other quantities, like the length of the minimum spanning and Steiner minimum trees, greedy and semi-matchings, and others. For further information on some

of these developments one can consult Halton & Terada [1982], Karp & Steele [1985], or Steele[1990a, b].

4.1. Subadditive Euclidean functionals and non-linear growth

The length of the traveling salesman tour has a few basic properties that are shared with a large number of problems of combinatorial optimization in Euclidean space. By abstracting some of the simplest of these properties, it is possible to suggest a very general result that provides information comparable to that given by the Beardwood, Halton, Hammersley theorem.

Let L be a function that maps the collection of finite subsets $\{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^d$ to the real numbers \mathbb{R} . To spell out the most innocent properties of L that mimic the behavior of the TSP, we first note that for the TSP, L exhibits *homogeneity* and *translation invariance*, i.e.,

$$L(\alpha x_1, \alpha x_2, \dots, \alpha x_n) = \alpha L(x_1, x_2, \dots, x_n) \text{ for all } \alpha > 0, \quad (4.1)$$

and

$$L(x_1 + y, x_2 + y, \dots, x_n + y) = L(x_1, x_2, \dots, x_n) \text{ for all } y \in \mathbb{R}^d. \quad (4.2)$$

The TSP's total length also has some strong smoothness and regularity properties, but these turn out not to be of essential importance, and for the generalization we consider we will need to call on the smoothness of L only to require that, for each n , the function L viewed as a function from \mathbb{R}^{nd} to \mathbb{R} is Borel measurable. This condition is almost always trivial to obtain, but it is nevertheless necessary in order to be able to talk honestly about probabilities involving L .

Functions on the finite subsets of \mathbb{R}^d that are measurable in the sense just described and that are homogeneous of order one and translation invariant are called *Euclidean functionals*. These three properties are somewhat bland, and one should not expect to be able to prove much in such a limited context, but with the addition of just a couple other structural features, one finds a rich and useful theory.

The first additional property of the TSP functional that we consider is that it is *monotone* in the sense that

$$L(x_1, x_2, \dots, x_n) \leq L(x_1, x_2, \dots, x_n, x_{n+1}) \text{ for } n \geq 1, \text{ and } L(\emptyset) = 0. \quad (4.3)$$

A second additional and final feature of the TSP functional that we abstract is the only truly substantial one. It expresses both the geometry of the space in which we work and the fundamental suboptimality of one of the most natural TSP heuristics, the partitioning heuristic. The *subadditive property* we require is that there exists a constant B such that

$$L(\{x_1, x_2, \dots, x_n\} \cap [0, t]^d) \leq \sum_{i=1}^{m^d} L(\{x_1, x_2, \dots, x_n\} \cap Q_i) + Btm^{d-1} \quad (4.4)$$

for all integers $m \geq 1$ and real $t \geq 0$, where $\{Q_i\}_{i=1}^{m^d}$ is a partition of $[0, t]^d$ into generally smaller cubes of edge length t/m .

Euclidean functionals that satisfy the last two assumptions will be called *monotone subadditive Euclidean functionals*. This class of processes seems to abstract the most essential features of the TSP that are needed for an effective asymptotic analysis of the functional applied to finite samples of independent random variables with values in \mathbb{R}^d .

To see how subadditive Euclidean functionals arise naturally and to see how some closely-related problems can just barely elude this framework, it is useful to consider two basic examples in addition to the TSP. The first is the *Steiner minimum tree*, which is a monotone subadditive Euclidean functional. For any finite set $S = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^d$, a Steiner minimum tree for S is a tree T whose vertex set contains S such that the sum of the lengths of the edges in T is minimal over all such trees. Note that the vertex set of T may contain points not in S ; these are called *Steiner points*. If $L_{ST}(x_1, x_2, \dots, x_n)$ is the length of a Steiner tree of x_1, x_2, \dots, x_n and if we let $l(e)$ be the length of an edge e , another way of defining L_{ST} is just

$$L_{ST}(S) = \min_T \left\{ \sum_{e \in T} l(e) : T \text{ is a tree containing } S \subset \mathbb{R}^d, S \text{ finite} \right\}.$$

A closely-related example points out that the innocuous monotonicity property of the TSP and Steiner minimum tree can fail in quite natural problems. The example we have in mind is the *minimum spanning tree*. For $\{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^d$, let $L_{MST}(x_1, x_2, \dots, x_n) = \min \sum_{e \in T} l(e)$, where the minimum is over all spanning trees of $\{x_1, x_2, \dots, x_n\}$. The functional L_{MST} is easily seen to be homogeneous, translation invariant, and properly measurable; one can also check without much trouble that it is subadditive in the sense required above. Still, by considering the sets $S = \{(0, 0), (0, 2), (2, 0), (2, 2)\}$ and $S \cup \{(1, 1)\}$, we see that L_{MST} fails to be monotone as required. One should suspect that this failure is of an exceptional sort that should not have great influence on asymptotic behavior, and it can be shown that this suspicion is justified. The example, however, puts us on warning that non-monotone functionals can require delicate considerations that are not needed in cases that mimic the TSP more closely.

Subject to a modest moment condition, the properties (4.1) through (4.4) are sufficient to determine the asymptotic behavior of $L(X_1, X_2, \dots, X_n)$, where the X_i are independent and identically distributed.

Theorem 1 [Steele, 1981a]. *Let L be a monotone subadditive Euclidean functional. If $\{X_i\}$, $i = 1, 2, \dots$, are independent random variables with the uniform distribution on $[0, 1]^d$, and $\text{Var } L(X_1, X_2, \dots, X_n) < \infty$ for each $n \geq 1$, then as $n \rightarrow \infty$*

$$\frac{L(X_1, X_2, \dots, X_n)}{n^{(d-1)/d}} \rightarrow \beta_{L,d}$$

with probability one, where $\beta_{L,d} \geq 0$ is a constant depending only on L and d .

The restrictions that this theorem imposes on a Euclidean functional are as few as one can reasonably expect to yield a generally useful limit theorem, and because of this generality the restriction to uniformly distributed random variables is palatable. Moreover, since many of the probabilistic models studied in operations research and computer science also focus on the uniformly distributed case, the theorem has immediate applications. Still, one cannot be long content with a theory confined to uniformly distributed random variables. Fortunately, with the addition of just a couple of additional constraints, the limit theory of subadditive Euclidean functionals can be extended to quite generally distributed variables.

4.2. Tail probabilities for the TSP and other functionals

The theory just outlined has a number of extensions and refinements. The first of these that we consider is the work of Rhee & Talagrand [1989] on the behavior of the tail probabilities of the TSP and related functionals under the model of independent uniformly distributed random variables in the unit d -cube. In Steele [1981b], it was observed that $\text{Var } L_n$ for $d = 2$ is bounded independently of n . This somewhat surprising result motivated the study of more detailed understanding of the tail probabilities $P(L_n \geq t)$, particularly the issue of determining if these probabilities decayed at the Gaussian rate $\exp(-cx^2/2)$. After introducing new methods from martingale theory and interpolation theory which led to interesting intermediate results, Rhee & Talagrand [1989] provided a remarkable proof that in $d = 2$, the TSP and many related functionals indeed have Gaussian tail bounds. The formal result can be stated as follows.

Theorem [Rhee & Talagrand, 1989]. *Let f be a Borel measurable function that assigns to each finite subset $F \subset [0, 1]^2$ a real value $f(F)$ such that*

$$f(F) \leq f(F \cup x) \leq f(F) + \min(d(x, y) : y \in F).$$

If X_i are independent and uniformly distributed in $[0, 1]^2$, then the random variable defined by $U_n = f(\{X_1, X_2, \dots, X_n\})$ is such that there exists a constant K for which, for all $t \geq 0$,

$$P(|U_n - E(U_n)| > t) \leq \exp\left(\frac{-t^2}{K}\right).$$

4.3. Worst-case asymptotics

The probabilistic rates of growth just surveyed are replicated in worst-case settings. In this section, we survey some of the work that has been done on worst-case growth rates and draw parallels with the probabilistic rates.

Let $l(e)$ be the usual Euclidean length $|e|$ of the edge e . As a primary example of a worst-case growth rate, consider the *worst-case length* of an optimal traveling

salesman tour in the unit d -cube:

$$\rho_{\text{TSP}}(n) = \max_{\substack{S \subset [0,1]^d \\ |S|=n}} \min_T \left\{ \sum_{e \in T} l(e) : T \text{ is a tour of } S \right\}. \quad (4.5)$$

In words, $\rho_{\text{TSP}}(n)$ is the maximum length, over all point sets in $[0, 1]^d$, that an optimal traveling salesman tour can attain. The minimized quantity in (4.5) is just the length of an optimal traveling salesman tour of the point set S , and the maximum is taken over all point sets S of size n . We note that there is no probability theory here, for the point sets and tours are deterministic. Steele & Snyder [1989] showed that, despite this, one obtains a rate of growth for ρ_{TSP} that is identical to the probabilistic growth rate in Theorem 1.

Theorem 3 [Steele & Snyder, 1989]. *As $n \rightarrow \infty$,*

$$\rho_{\text{TSP}}(n) \sim \alpha_{\text{TSP},d} n^{(d-1)/d},$$

where $\alpha_{\text{TSP},d} > 0$ is a constant depending only on the dimension d .

4.4. Progress on the constants

Estimation of the limiting constants has a long history in both the worst-case and stochastic contexts. Few [1955] improved some very early work to provide the bound $\alpha_{\text{TSP},2} \leq \sqrt{2}$ and gave a dimension- d bound of $\alpha_{\text{TSP},d} \leq d\{2(d-1)\}^{(1-d)/2d}$, where $d \geq 2$. After other intermediate work, Karloff [1989] broke the $\sqrt{2}$ barrier in dimension two by showing that $\alpha_{\text{TSP},2} \leq 0.984\sqrt{2}$. The best bounds currently available in higher dimensions are those of Goddyn [1990].

Bounds on the worst-case constants are also available for other Euclidean network problems. Of particular note is the bound on $\alpha_{\text{RST},d}$, the constant associated with the worst-case length of a rectilinear Steiner minimum tree in the unit d -cube. Chung & Graham [1981] proved that $\alpha_{\text{RST},2} = 1$, which is significant in that it is the only non-trivial worst-case constant for which we have an exact expression. The problem of determining $\alpha_{\text{RST},d}$ in higher dimensions is still open, with the current best-known bounds being $\max\{1, d/(4e)\} \leq \alpha_{\text{RST},d} \leq d4^{(1-d)/d}$, for $d \geq 1$ [Snyder, 1991, 1992; Salowe, 1992].

In the case of the probabilistic models, there is recent progress due to Bertsimas and van Ryzin [1990], where asymptotic expressions as d gets large were obtained for the probabilistic minimum spanning tree and matching constants $\beta_{\text{MST},d}$ and $\beta_{\text{M},d}$. Specifically, they showed that $\beta_{\text{MST},d} \sim \sqrt{d/2\pi e}$ and $\beta_{\text{M},d} \sim (1/2)\sqrt{d/2\pi e}$ as $d \rightarrow \infty$. Still, the most striking progress on probabilistic constants was the determination of an exact expression for $\beta_{\text{MST},d}$ for all $d \geq 2$ by Avram & Bertsimas [1992]. Their expression for $\beta_{\text{MST},d}$ comes in the form of series expansion in which each term requires a rather difficult integration. The representation is still an effective one, and the first few terms of the series in dimension two have been computed to yield a numerical lower bound of $\beta_{\text{MST},2} \geq 0.599$, which agrees well

with experimental data. The proof of the series representation for $\beta_{\text{MST},d}$ relies strongly on the fact that a greedy construction is guaranteed to yield an MST, and unfortunately these constructions are not possible for many objects of interest, including the TSP.

5. Concluding remarks

The theory of probabilistic networks and associated algorithms is rapidly evolving, but it is not yet a well consolidated field of inquiry. In surveying the literature, one finds relevant contributions growing in many separate areas, including the theory of random graphs, subadditive Euclidean functionals, stochastic networks, reliability, percolation, and computational geometry. Many of these areas make systematic use of tools and methodologies that remain almost unknown to the other areas, despite compelling relevance. The aim here has been to provide a view of part of the cross-fertilization that seems possible, but of necessity our focus has been on topics that allow for reasonably brief or self-contained description. Surely one can — and should — go much further. For more general information concerning probability theory applied to algorithms, one can consult the surveys of Karp [1977, 1991], Rinnooy Kan [1987], Hofri [1987], and Stougie [1990], as well as the bibliography of Karp, Lenstra, McDiarmid, and Rinnooy Kan [1985]. For more on percolation theory, Grimmett [1989] is a definitive reference.

Acknowledgements

This research was supported in part by Georgetown University 1991 Summer Research Award, by Georgetown College John R. Kennedy, Jr. Faculty Research Fund, and by the following grants: NSF DMS92-11634, NSA MDA904-91-H-2034, AFOSR-91-0259, and DAAL03-89-G-0092.

References

- Ahlswede, R., and D.E. Daykin (1978). An inequality for the weights of two families of sets, their unions and intersections. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* 43, 183–185.
- Ahuja, R.K., T.L. Magnanti and J.B. Orlin (1991). Some recent advances in network flows. *SIAM Rev.* 33, 175–219.
- Ahuja, R.K., and J.B. Orlin (1987). A fast and simple algorithm for the maximum flow problem. *Oper. Res.* 37, 748–759.
- Alon, N. (1990). Generating pseudo-random permutations and maximum flow algorithms. *Inf. Process. Lett.* 35, 201–204.
- Avram, F., and D. Bertsimas (1992). The minimum spanning tree constant in geometric probability and under the independent model: a unified approach. *Ann. Appl. Probab.* 2, 118–130.
- Beardwood, J., J.H. Halton and J. Hammersley (1959). The shortest path through many points. *Proc. Camb. Philos. Soc.* 55, 299–327.

- Bertsimas, D., and G. Van Ryzin (1990). An asymptotic determination of the minimal spanning tree and minimal matching constants in geometric probability. *Oper. Res. Lett.* 9, 223–231.
- Bollobás, B. (1985). *Random Graphs*, Academic Press, New York, N.Y.
- Bollobás, B. (1986). *Combinatorics*, Cambridge University Press, New York, N.Y.
- Cheriyian, J., and T. Hagerup (1989). A randomized max-flow algorithm, *Proc. 30th IEEE Foundations of Computer Science*, IEEE, pp. 118–123.
- Chung, F.R.K., and R.L. Graham (1981). On Steiner trees for bounded point sets. *Geom. Dedicata* 11, 353–361.
- Colbourn, C.J. (1987). *The Combinatorics of Network Reliability*, Oxford University Press, New York, N.Y.
- Cox, J.T., and R. Durrett (1988). Limit theorems for the spread of epidemics and forest fires. *Stochastic Process. Appl.* 30, 2, 171–191.
- Dijkstra, E.W. (1959). A note on two problems in connexion with graphs. *Numer. Math.* 1, 269–271.
- Even, S., and O. Kariv (1975). An $O(n^{2.5})$ algorithm for maximum matching in general graphs, *Proc. 16th IEEE Symp. Foundations of Computer Science*, IEEE, pp. 100–112.
- Few, L. (1955). The shortest path and the shortest road through n points in a region. *Mathematika* 2, 141–144.
- Fortuin, C.M., P.N. Kasteleyn and J. Ginibre (1971). Correlation inequalities on some partially ordered sets. *Commun. Math. Phys.* 22, 89–103.
- Frieze, A.M., and G.R. Grimmett (1985). The shortest-path problem for graphs with random arc-lengths. *Discrete Appl. Math.* 10, 57–77.
- Gabow, H.N. (1985). Scaling algorithms for network problems. *J. Comput. Systems Sci.* 31, 148–168.
- Goddyn, L. (1990). Quantizers and the worst-case Euclidean traveling salesman problem. *J. Comb. Theor., Ser. B* 50, 65–81.
- Goldberg, A.V., and R.E. Tarjan (1988). A new approach to the maximum-flow problem. *J. ACM* 35, 921–940.
- Graham, R.L. (1983). Applications of the FKG inequality and its relatives, in: A. Bachem, M. Grötschel, and B. Korte (eds). *Mathematical Programming: The State of the Art, Bonn 1982*, Springer-Verlag, New York, NY, pp. 115–131.
- Grimmett, G.R. (1989). *Percolation*, Springer-Verlag, New York, N.Y.
- Grimmett, G.R., and D.J.A. Welsh (1982). Flow in networks with random capacities. *Stochastics* 7, 205–229.
- Halton, J.H., and R. Terada (1982). A fast algorithm for the Euclidean traveling salesman problem, optimal with probability one. *SIAM J. Comput.* 11, 28–46.
- Harris, T.E. (1960). A lower bound for the critical probability in a certain percolation process, *Proc. Camb. Philos. Soc.* 56, 13–20.
- Hofri, M. (1987). *Probabilistic Analysis of Algorithms: On Computing Methodologies for Computer Algorithms Performance Evaluation*, Springer-Verlag, New York, N.Y.
- Jerrum, M., and A. Sinclair (1986). The approximation of the permanent, *Proc. 18th Symp. on Theory of Computing Association for Computing Machinery*, pp. 235–243.
- Jerrum, M., and A. Sinclair (1989). The approximation of the permanent. *SIAM J. Comput.* 18, 1149–1178.
- Karloff, H.J. (1989). How long can a Euclidean traveling salesman tour be? *SIAM J. Disc. Math.* 2, 91–99.
- Karp, R.M. (1976). The probabilistic analysis of some combinatorial search algorithms, in: J.F. Traub (ed.), *Algorithms and Complexity: New Directions and Recent Results*, Academic Press, New York, N.Y., pp. 1–19.
- Karp, R.M. (1977). Probabilistic analysis of partitioning algorithms for the traveling salesman problem in the plane. *Math. Oper. Res.* 2, 209–224.
- Karp, R.M. (1991). An introduction to randomized algorithms. *Discrete Appl. Math.* 34, 165–201.
- Karp, R.M., J.K. Lenstra, C.J.H. McDiarmid and A.H.G. Rinnooy Kan (1985). Probabilistic analysis, in: M. O’heigeartaigh, J.K. Lenstra and A.H.G. Rinnooy Kan (eds.), *Combinatorial Optimization: Annotated Bibliographies*, John Wiley and Sons, Chichester, pp. 52–88.

- Karp, R.M., and M. Luby (1985). Monte Carlo algorithms for planar multiterminal network reliability, *J. Complexity* 1, 45–64.
- Karp, R.M., R. Motwani and N. Nisan (1993). Probabilistic analysis of network flow algorithms. *Math. Oper. Res.* 18, 71–97.
- Karp, R.M., and J.M. Steele (1985). Probabilistic analysis of heuristics, in: E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan and D.B. Shmoys (eds.), *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*, John Wiley and Sons, New York, N.Y., pp. 181–206.
- Kesten, H. (1980). The critical probability of bond percolation on the square lattice equals $\frac{1}{2}$. *Commun. Math. Phys.* 74, 41–59.
- Kulkarni, V.G. (1986). Shortest paths in networks with exponentially distributed arc lengths. *Networks* 16, 255–274.
- Lovász, L. (1979). On determinants, matchings, and random algorithms, in: L. Budach (ed.), *Fundamentals of Computing Theory*, Akademie-Verlag, Berlin.
- Micali, S., and V.V. Vazirani (1980). An $O(|V|^{0.5}|E|)$ algorithm for finding maximum matchings in general graphs, *Proc. 21st IEEE Symp. on Foundations of Computer Science*, IEEE, pp. 17–27.
- Motwani, R. (1989). Expanding graphs and the average-case analysis of algorithms for matchings and related problems, *Proc. 21st Symp. on Theory of Computing Ass. Comput. Mach.*, pp. 550–561.
- Rabin, M.O., and V.V. Vazirani (1989). Maximum matching in general graphs through randomization. *J. Algorithms* 10, 557–567.
- Rhee, W.T., and M. Talagrand (1989). A sharp deviation inequality for the stochastic traveling salesman problem. *Ann. Probab.* 17, 1–8.
- Rinnooy Kan, A.H.G. (1987). Probabilistic analysis of algorithms. *Ann. Discrete Math.* 31, 365–384.
- Russo, L. (1981). On the critical percolation probabilities. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* 56, 129–139.
- Salowe, J.S. (1992). A note on lower bounds for rectilinear Steiner trees. *Inf. Proc. Lett.* 42, 151–152.
- Shepp, L.A. (1982). The XYZ conjecture and the FKG inequality. *Ann. Probab.* 10, 824–827.
- Sigal, E.C., A.A.B. Pritsker and J.J. Solberg (1979). The use of cutsets in Monte Carlo analysis of stochastic networks. *Math. Comput. Simulat.* 21, 376–384.
- Sigal, E.C., A.A.B. Pritsker and J.J. Solberg (1980). The stochastic shortest route problem. *Oper. Res.* 28, 1122–1130.
- Sinclair, A. (1993). *Algorithms for Random Generation and Counting: A Markov Chain Approach*, Birkhäuser Publishers, Boston, Mass.
- Snyder, T.L. (1992). Minimal rectilinear Steiner trees in all dimensions. *Discr. Comp. Geometry* 8, 73–92.
- Snyder, T.L. (1991). Lower bounds for rectilinear Steiner trees in bounded space, *Inf. Process. Lett.* 37, 71–74.
- Spencer, J. (1993). The Janson Inequality, in: D. Miklos, V.T. Sos and T. Szonyi (eds.), *Combinatorics, Paul Erdős is Eighty*, Vol. 1, Bolyai Mathematical Studies, Keszthely (Hungary), pp. 421–432.
- Steele, J.M. (1981a). Subadditive Euclidean functionals and non-linear growth in geometric probability. *Ann. Probab.* 9, 365–376.
- Steele, J.M. (1981b). Complete convergence of short paths and Karp's algorithm for the TSP. *Math. Oper. Res.* 6, 374–378.
- Steele, J.M. (1990a). Probabilistic and worst-case analyses of classical problems of combinatorial optimization in Euclidean space. *Math. Oper. Res.* 15, 749–770.
- Steele, J.M. (1990b). Seedlings in the theory of shortest paths, in: J. Grimmett and D. Welsh (eds.), *Disorder in Physical Systems: A Volume in Honor of J.M. Hammersley*, Cambridge University Press, London, pp. 277–306.
- Steele, J.M., and T.L. Snyder (1989). Worst-case growth rates of some classical problems of combinatorial optimization. *SIAM J. Comput.* 18, 278–287.

- Stougie, L. (1990). Applications of probability theory in combinatorial optimization, Class Notes, University of Amsterdam.
- Valiant, L.G. (1979). The complexity of enumeration and reliability problems. *SIAM J. Comput.* 12, 777–788.
- Van den Berg, J., and H. Kesten (1985). Inequalities with applications to percolation and reliability. *J. Appl. Prob.* 22, 556–569.