

Minimal Spanning Trees for Graphs with Random Edge Lengths

J. Michael Steele

ABSTRACT: *The theory of the minimal spanning tree (MST) of a connected graph whose edges are assigned lengths according to independent identically distributed random variables is developed from two directions. First, it is shown how the Tutte polynomial for a connected graph can be used to provide an exact formula for the length of the minimal spanning tree under the model of uniformly distributed edge lengths. Second, it is shown how the theory of local weak convergence provides a systematic approach to the asymptotic theory of the length of the MST and related power sums. Consequences of these investigations include (1) the exact rational determination of the expected length of the MST for the complete graph K_n for $2 \leq n \leq 9$ and (2) refinements of the results of Penrose (1998) for the MST of the d -cube and results of Beveridge, Frieze, and McDiarmid (1998) and Frieze, Ruzinkó, and Thoma (2000) for graphs with modest expansion properties. In most cases, the results reviewed here have not reached their final form, and they should be viewed as part of work-in-progress.*

1 Introduction and Main Results

Consider a finite, connected, simple graph G with vertex set $v(G)$, and for each element of the edge set $e(G)$ let ξ_e denote a nonnegative random variable that one views as the length of the edge e . The random variables $\{\xi_e : e \in e(G)\}$ are assumed to be independent with a common distribution F , and the quantities that are of central concern here are the total length of the minimal spanning tree (MST) of G ,

$$L_{\text{MST}}(G) = \sum_{e \in G} \xi_e \mathbb{I}(e \in \text{MST}(G)),$$

and the associated sums for power weighted edges

$$L_{\text{MST}}^\alpha(G) = \sum_{e \in G} \xi_e^\alpha \mathbb{I}(e \in \text{MST}(G)).$$

The first of these sums has been studied extensively since Frieze (1985) showed that for edge lengths with the uniform distribution on $[0, 1]$ that one has

$$\mathbb{E}[L_{\text{MST}}(K_n)] \rightarrow \zeta(3) = \sum_{k=1}^{\infty} \frac{1}{k^3} = 1.202 \dots \quad \text{as } n \rightarrow \infty \quad (1)$$

where K_n is the complete graph on n vertices.

In particular, this result has now been refined or extended by numerous investigations. There are relaxations of the distributional assumption by Steele (1987), extensions to the bipartite MST expectations $\mathbb{E}[L_{\text{MST}}(K_{n,n})]$ by Frieze and McDiarmid (1989), and even the development of a central limit theorem for $L_{\text{MST}}(K_n)$ by Janson (1995). More recently, the basic limit (1) has been extended

to larger classes of graphs, including an extension to the d -cube Q_d by Penrose (1998) and extensions to general classes of “modestly expansive” regular graphs by Beveridge, Frieze, and McDiarmid (1998) and Frieze, Ruzinkó, and Thoma (2000).

The path taken here diverges from this earlier work in several respects, but one key difference comes from the focus on exact calculations, rather than asymptotic relations. Specifically, we provide a formula for $\mathbb{E}[L_{\text{MST}}(G)]$ that permits one to determine the exact rational value of $\mathbb{E}[L_{\text{MST}}(G)]$ for many concrete choices of G .

We also pursue exact calculations for a certain infinite graph \mathcal{T} that is in a sense the universal limit for any sequence of randomly rooted independently weighted finite graphs whose vertex degrees go to infinity. This calculation then permits us to provide a necessary and sufficient conditions for the determination of the asymptotic behavior of $\mathbb{E}[L_{\text{MST}}^\alpha(G_n)]$ for a large class of sequences of graphs.

After framing our main results more fully in the next few paragraphs, we turn to the proofs. In particular, Section 2 develops an exact formula for $\mathbb{E}[L_{\text{MST}}(G)]$ finite G , and then in Section 3 we calculate the expected length per vertex of a special subgraph of \mathcal{T} that holds the key to many of the limit theorems for the MST. Section 4 then addresses some foundational results that connect calculations on \mathcal{T} to calculations for sequences of finite randomly rooted graphs, and these results are subsequently applied to complete the proof of the basic limit theorem for $\mathbb{E}[L_{\text{MST}}^\alpha(G_n)]$. Section 4 also examines a critical example that serves to illustrate the role of uniform integrability in the limit theory of the MST. The final section reviews some open problems and briefly speculates on the possibilities for further development.

A Formula for $\mathbb{E}[L_{\text{MST}}(G)]$

Theorem 1.1 *If G is a finite connected graph and the Tutte polynomial¹ of G is $T(G; x, y)$, then for independent edge lengths that are uniformly distributed on $[0, 1]$, one has*

$$\mathbb{E}[L_{\text{MST}}(G)] = \int_0^1 \frac{(1-p)}{p} \frac{T_x(G; 1/p, 1/(1-p))}{T(G; 1/p, 1/(1-p))} dp, \quad (2)$$

where $T_x(x, y)$ denotes the partial derivative of $T(x, y)$ with respect to x .

We illustrate the efficacy of this formula by providing what we believe to be the first explicit computations for $\mathbb{E}[L_{\text{MST}}(K_n)]$ for finite values of n that go beyond the trivial $n = 2$ and the easy $n = 3$. Specifically, we use this formula to calculate $\mathbb{E}[L_{\text{MST}}(K_n)]$ for $2 \leq n \leq 9$, and these calculations lead to several compelling conjectures.

Asymptotic Consequences of an Exact Calculation

If $d_n, n = 1, 2, \dots$ is a sequence of integers such that $d_n \rightarrow \infty$ as $n \rightarrow \infty$, we say that the sequence of graphs $G_n, n = 1, 2, \dots$ is *nearly regular* provided that the maximum $\Delta(G_n)$ degree and the minimum degrees $\delta(G_n)$ satisfy the degree conditions

$$\Delta(G_n) \sim d_n \quad \text{and} \quad \delta(G_n) \sim d_n \quad \text{as } n \rightarrow \infty. \quad (3)$$

¹Subsection 2.2 provides a brief but friendly development of the necessary background on the Tutte polynomial — beginning with its definition.

We also relax our assumption on $F(x) = P(\xi_e \leq x)$, and instead of requiring that the ξ_e be uniformly distributed on $[0, 1]$ we only require

$$F(0) = 0 \quad \text{and} \quad F(x) = x + o(x) \quad \text{as} \quad x \rightarrow 0, \quad (4)$$

a condition that simultaneously covers the uniform distribution on $[0, 1]$ and the exponential distribution with mean one — our two leading cases.

Next, we consider the power-weighted analog to the MST,

$$L_{\text{MST}}^\alpha(G_n) = \sum_e \xi_e^\alpha \mathbb{I}(e \in \text{MST}(G_n)),$$

and we introduce a new sequence

$$Y_\alpha(G_n) = \sum_e (d_n \xi_e)^\alpha \mathbb{I}(e \in \text{MST}(G_n) \text{ and } R(G_n) \in e), \quad (5)$$

where $R(G)$ denotes an element of the vertex set $v(G)$ that is chosen independently according to the uniform distribution. While $Y_\alpha(G_n)$ may not seem natural at first, we will see shortly that its expectation determines the expectation of L_{MST}^α ; moreover, there are major technical benefits to working with $Y_\alpha(G_n)$. In particular, $Y_\alpha(G_n)$ satisfies a limit law that requires nothing more of the graph sequence $\{G_n\}$ than those features that one needs for the definition of $Y_n(G_n)$ and the statement of the limit. As an easy consequence of the general theory of local weak convergence and an exact calculation on a special infinite tree, one obtains the asymptotic behavior of $\mathbb{E}[L_{\text{MST}}^\alpha(G_n)]$.

Theorem 1.2 *If G_n , $n = 1, 2, \dots$ is a sequence of connected graphs that are nearly regular in the sense of (3), then for any $0 < \alpha < \infty$, one has*

$$\mathbb{E}[L_{\text{MST}}^\alpha(G_n)] \sim \Gamma(1 + \alpha) \zeta(2 + \alpha) |v(G_n)| d_n^{-\alpha} \quad \text{as} \quad n \rightarrow \infty, \quad (6)$$

if and only if the sequence

$$\{Y_\alpha(G_n) : n = 1, 2, \dots\} \quad \text{is uniformly integrable.}$$

Thus, one finds that even a crude qualitative measure of the good behavior of the sequence $\{Y_\alpha(G_n) : n = 1, 2, \dots\}$ is enough to guarantee the regular asymptotic behavior of $\mathbb{E}[L_{\text{MST}}^\alpha(G_n)]$. Moreover, the good behavior of $\{Y_\alpha(G_n)\}$ turns out to be necessary, so one finds a strong hint that this sequence may be more fundamental to the theory of the MST than first impressions might suggest.

As a quick illustration of this last result, we should note that if one takes $\alpha = 1$ and takes G_n to be Q_n , the n -cube, then it implies

$$\mathbb{E}[L_{\text{MST}}(Q_n)] \sim \frac{2^n}{n} \zeta(3) \quad \text{as} \quad n \rightarrow \infty, \quad (7)$$

a limit which was found by Penrose (1998) by different means. A more novel consequence of the limit (6) comes from taking $\alpha = 1/2$ and $\alpha = 3/2$ to find that the limit (7) is nicely sandwiched between

$$\mathbb{E}[L_{\text{MST}}^{1/2}(Q_n)] \sim \frac{1}{2} \zeta(5/2) 2^n \sqrt{\frac{\pi}{n}} \quad \text{and} \quad \mathbb{E}[L_{\text{MST}}^{3/2}(Q_n)] \sim \frac{3}{4} \zeta(7/2) \sqrt{\pi n}^{-3/2} 2^n.$$

Despite the large swath of ground that Theorem 1.2 covers, one should not lose sight of the fact that it really is a simple corollary of more general result from the theory of local weak convergence that has its roots in Aldous (1992) and Aldous (2001). In particular, local weak convergence to the PWIT is a fundamental part of those papers, and the modest generalization PWIT Limit Theorem developed here in Theorem 4.2 is best viewed as part of a longer term effort to make the techniques introduced in Aldous (1992) and Aldous (2001) more easily accessible and more readily applied.

2 Exact Calculations for Finite Graphs

The program begins with the derivation of an exact formula for the expectation $\mathbb{E}[L_{\text{MST}}(G)]$ under the uniform model for the edge lengths. Here the first step is to derive a relationship between the random variable $L_{\text{MST}}(G)$ and an integral of another random variable that measures the connectedness of G when one just uses edges length not greater than $0 \leq p \leq 1$. Versions of this relationship go back at least to Avram and Bertsimas (1992), and in some way or another it has had a role in most recent investigations of the MST, including the central limit theorem of Janson (1999) and the general graph MST results of Beveridge, Frieze, and McDiarmid (1998) and Frieze, Ruzinkó, and Thoma (2000).

2.1 Length of the MST as an Integral

For any finite graph G and any subset A of the edge set $e(G)$, we write $k(G, A)$ for the number of connected components of the graph with vertex set $v(G)$ and edge set A . If each edge $e \in G$ is assigned length ξ_e , then we also write

$$e_t(G) = \{e \in e(G) : \xi_e \leq t\},$$

and we let

$$N_{\text{MST}}(G, t) = \sum_{e \in \text{MST}(G)} \mathbb{I}(\xi_e \leq t),$$

so $N_{\text{MST}}(G, t)$ denotes the number of edges of the MST of G that are elements of $e_t(G)$. Now, if G is a connected graph, then by counting the number of elements of $e_t(G)$ in each connected component of $(G, e_t(G))$ one finds

$$N_{\text{MST}}(G, t) + k(G, e_t(G)) = n,$$

so we can simply compute

$$\begin{aligned} L_{\text{MST}}(G) &= \sum_{e \in G} \xi_e \mathbb{I}(e \in \text{MST}(G)) = \sum_{e \in G} \int_0^1 \mathbb{I}(t < \xi_e, e \in \text{MST}(G)) dt \\ &= \int_0^1 \sum_{e \in G} (1 - \mathbb{I}(\xi_e \leq t, e \in \text{MST}(G))) dt \\ &= \int_0^1 (n - 1 - N_{\text{MST}}(G, t)) dt = \int_0^1 \{k(G, e_t(G)) - 1\} dt. \end{aligned}$$

In other words, for any connected graph we have the rather pleasing random variable representation

$$1 + L_{\text{MST}}(G) = \int_0^1 k(G, e_t(G)) dt. \quad (8)$$

Thus, our main task is to understand the expectation of $k(G, e_t(G))$, and this provides a natural roll for the Tutte polynomials.

2.2 The Tutte Polynomial

To define the Tutte polynomial, one needs to go outside the familiar class of simple graphs and to consider graphs that may have loops or parallel edges. Given such a graph G , the Tutte polynomial $T(G; x, y)$ is then defined by a set of four devilishly simple rules:

1. If G has no edges, then $T(G; x, y) = 1$.
2. If e is an edge of G that is neither a loop nor an isthmus, then

$$T(G; x, y) = T(G'_e; x, y) + T(G''_e; x, y),$$

where G'_e is the graph G with the edge e *deleted* and G''_e is the graph G with the edge e *contracted*.

3. If e is an isthmus, then $T(G; x, y) = xT(G'_e; x, y)$.
4. If e is a loop, then $T(G; x, y) = yT(G''_e; x, y)$.

To confirm the understanding of these rules, one might want to check that they imply that the Tutte polynomial of K_2 is just x ; indeed, by successive applications of Rule 3 one finds that the Tutte polynomial of any tree with n vertices is just the monomial x^{n-1} .

The rules are more amusing when one needs to use contractions, and here the basic exercise is to show that the Tutte polynomial of K_3 is $x + x^2 + y$. Finally, one might want to check that the Tutte polynomial of a *bow tie* (made by two copies of K_3 joined at a vertex) is just $(x + x^2 + y)^2$. The last exercise naturally suggests a general principle for finding the Tutte polynomial for the graph built by joining two arbitrary graphs at a single vertex; one can then recapture Rule 3 as a special case of the general principle.

Much of the usefulness of the Tutte polynomial comes from its relation to the rank function $r(\cdot)$ that associates to each $A \subset e(G)$ the integer $r(A)$ given by

$$r(A) = |v(G)| - k(G, A),$$

where, as before, $k(G, A)$ is the number of connected components of the graph with vertex set $v(G)$ and edge set A . The rank function provides a measure of the extent to which the graph $(v(G), A)$ is connected, and it permits one to express the Tutte polynomial as a large — but informative — sum:

$$T(G; x, y) = \sum_{A \subset e(G)} (x-1)^{r(G)-r(A)} (y-1)^{|A|-r(A)}, \quad (9)$$

where $r(G)$ is shorthand for the more pedantic $r(e(G))$.

One immediate consequence of this formula is that it shows the Tutte polynomial does not depend on the order in which one deletes the edges of G in the recursive definition of $T(G; x, y)$, a fact that may not seem particularly evident from the rules themselves. To return the favor, the defining rules make it evident the coefficients of $T(G; x, y)$ are nonnegative, while this is not so easily seen from the sum.

One obvious consequence of the sum formula (9) is that

$$T(G; 2, 2) = 2^m \quad \text{where } m = |e(G)|, \quad (10)$$

and a natural use of this triviality is to provide a quick feasibility check on a candidate Tutte polynomial. In fact, the evaluations of the Tutte polynomial at special choices of x and y provide a rich buffet of combinatorial interpretations (cf. Welsh (1999)), and in principle each such evaluation can be used as a check. In practice, the evaluation (10) is the easiest to use; it catches many blunders and offers many hints.

2.3 Connection to the Probability Model

Any sum over all of the subsets of $e(G)$ can be interpreted as an expectation

$$\sum_{A \subseteq e(G)} p^{|A|} (1-p)^{m-|A|} f(A) \quad (11)$$

for an appropriate choice of f , and when one recalls that

$$r(A) = |v(G)| - k(G, A) = n - k(G, A) \quad \text{and} \quad r(G) = n - 1$$

for a connected graph G , then the sum formula (9) is simply

$$T(G; x, y) = \frac{1}{(x-1)(y-1)^n} \sum_{A \subseteq e(G)} (y-1)^{|A|} ((x-1)(y-1))^{k(G,A)},$$

which obviously may be written in expectation form as

$$\frac{y^m}{(x-1)(y-1)^n} \sum_{A \subseteq e(G)} \left(\frac{y-1}{y}\right)^{|A|} \left(\frac{1}{y}\right)^{m-|A|} ((x-1)(y-1))^{k(G,A)}, \quad (12)$$

provided that we set $m = |e(G)|$ and make the natural identifications

$$p = \frac{y-1}{y} \quad \text{and} \quad 1-p = \frac{1}{y}. \quad (13)$$

This kind of reinterpretation of the Tutte polynomial is bread-and-butter to the theory of the correlated percolation model (cf. Fortuin and Kasteleyn (1972)), and this specific form of the Tutte polynomial has also been useful in the study of the computational complexity of the Tutte polynomial (cf. Welsh (1999) and especially Lemma 1 of Alon, Frieze, Welsh (1994)).

On the other hand, the application of this formula to the problem of calculating the minimal spanning tree for uniformly distributed edges seems to be novel,

though admittedly easy and natural. We first note that the first factors under the sum provide the probability under the uniform model that one has $\xi_e \leq p$ for exactly those edges in the set A . If one then takes

$$A = e_p(G) \equiv \{e : e \in e(G), \xi_e \leq p\}$$

then one can write the moment generating function

$$\varphi(t) \equiv \mathbb{E}[\exp(tk(G, e_p(G)))]$$

in terms of $T(G; x, y)$ as

$$\varphi(t) = p^{n-1}(1-p)^{m-n+1} e^t T\left(G; 1 + e^t \frac{1-p}{p}, \frac{1}{1-p}\right), \quad (14)$$

and this formula gives us a natural way to calculate the expectation of $k(G, e_p(G))$.

Specifically, if we retain the abbreviations (13), we have

$$\varphi'(t) = \varphi(t) \left\{ 1 + e^t \frac{1-p}{p} \frac{T_x(G; x, y)}{T(G; x, y)} \right\},$$

so, when we let $t = 0$, we find for $x = 1/p$ and $y = 1/(1-p)$ that

$$\mathbb{E}[k(G, e_p(G))] = 1 + \frac{1-p}{p} \frac{T_x(G; x, y)}{T(G; x, y)}. \quad (15)$$

Finally, when we expand the abbreviations for x and y and recall the representation (8) for $L_{\text{MST}}(G)$ in as an integral of $k(G, e_p(G))$, we find

$$\mathbb{E}[L_{\text{MST}}(G)] = \int_0^1 \frac{(1-p)}{p} \frac{T_x(G; 1/p, 1/(1-p))}{T(G; 1/p, 1/(1-p))} dp, \quad (16)$$

just as we needed to complete the proof of Theorem 1.1.

2.4 Illustrations and Applications

There are some natural and easy checks one can make to familiarize the formula (16). If we recall that for $G = K_2$ we have $T(G; x, y) = x$, then the integral (16) easily works out to be $1/2$, just as it should. More generally, if G is a tree with n vertices, then $T(G; x, y) = x^{n-1}$ and the integral work out to be $(n-1)/2$, and again this is obviously the correct value of $\mathbb{E}[L_{\text{MST}}(G)]$.

It is perhaps more informative to note that the form of the integrand as a logarithmic derivative is quite natural. If G and H are two graphs that share a common vertex, then the graph $G \cup H$ has Tutte polynomial $T(G; x, y)T(H; x, y)$ so the formula (16) recaptures the obvious fact that in this case one also has

$$\mathbb{E}[L_{\text{MST}}(G \cup H)] = \mathbb{E}[L_{\text{MST}}(G)] + \mathbb{E}[L_{\text{MST}}(H)].$$

For the complete graph on three vertices we have already seen that one has $T(K_3) = x + x^2 + y$, and for this polynomial the integral (16) yields $3/4$, and yet again one can check independently that $\mathbb{E}[L_{\text{MST}}(K_3)] = 3/4$. Nevertheless, for K_4

n	$\mathbb{E}[L_{\text{MST}}(K_n)]$	Numerical Value	Forward Difference
2	1/2	0.50000	0.250000
3	3/4	0.75000	0.135714
4	31/35	0.88571	0.080735
5	893/924	0.96645	0.051864
6	278/273	1.01832	0.035400
7	30739/29172	1.05372	0.025342
8	199462271/184848378	1.07906	0.018843
9	126510063932/115228853025	1.09790	—————

Table 1: The exact expected values of the MST of K_n for $n = 2$ to $n = 9$ under the model of independent $U[0, 1]$ edge lengths.

the situation is much more interesting. Hand computations become tedious, but they still suffice for one to show

$$T(K_4; x, y) = 2x + 2y + 3x^2 + 3y^2 + 4xy + x^3 + y^3.$$

When this polynomial is used in the integral formula (16), one then finds

$$\mathbb{E}[L_{\text{MST}}(K_4)] = \frac{31}{35},$$

and now we are on new ground. This appears to be the first time $\mathbb{E}[L_{\text{MST}}(K_4)]$ has been computed, and one may be hard pressed to provide an independent calculation that not pass through some integral like that provided by our basic representation (16).

Naturally one can go further, but beyond $n = 4$ it would be masochistic not to use symbolic calculation to determine the Tutte polynomials and to perform the required integrations. In fact, a table of the Tutte polynomials $T(K_n; x, y)$ for the values $n = 2, 3, \dots, 8$ is included in Gessel and Sagan (1996), and with help from Maple this table has been extended by Gessel (personal communication) to include all values up to $n = 15$. For convenience of display, we use just the first nine of these polynomials in the construction of Table 1.

The numerical evaluations in the table and their successive differences suggest two compelling conjectures; it seems inevitable that $\mathbb{E}[L_{\text{MST}}(K_n)]$ is monotone increasing and concave. This evidence is new and not fully digested, so it is possible that these conjectures will follow from our basic formula (16) and the known properties of the Tutte polynomial for K_n . On the other hand, if such an approach is not successful, the conjectures may prove to be difficult. After all, the analogous monotonicity conjecture for the assignment problem (cf. Steele (1997), p. 94) has resisted all attempts for more than fifteen years.

A final feature of Table 1 worth noting is that the rate of convergence is perhaps slower than one might guess. By the result of Frieze (1985) mentioned in the introduction, we know that $\mathbb{E}[L_{\text{MST}}(K_n)]$ converges to $\zeta(3) = 1.202\dots$, and one might hope that the behavior of $\mathbb{E}[L_{\text{MST}}(K_n)]$ would parallel that of the partial sums of $\zeta(3)$ given by $s_n = 1 + 1/2^3 + \dots + 1/n^3$. Sadly, s_n reaches 1.19 when $n = 6$ and reaches 1.20 when $n = 16$, while $\mathbb{E}[L_{\text{MST}}(K_n)]$ lags far behind. By analogy with the Parisi conjecture for the assignment problem (cf. Parisi (1998) and Aldous and Steele (2002)), one suspects that under the exponential model the corresponding expected values $\mathbb{E}[L_{\text{MST}}(K_n)]$ will indeed be closer to s_n . Nevertheless, such an exploration will have to wait for another day.

3 An Exact Calculation for an Infinite Tree

We now take up a second exact calculation, but this time it will be for a special infinite graph. To explain why this graph deserves to be singled out requires some background on the theory of the Poisson weighted infinite tree and the attending theory of local weak convergence. This background is developed more fully in Aldous (2001) and Aldous and Steele (2002), so the next two subsections recall just the most essential facts.

3.1 An Infinite Tree of Special Significance

The Poisson weighted infinite tree — or the PWIT — is a simple object. Nevertheless, it provides one with a direct and effective understanding of many of the problems of combinatorial optimization for large graphs with edge lengths that are given by independent random variables.

Formally, a PWIT is a rooted tree that one defines recursively. One starts with a single vertex r called the *root*, and one gives the root a countably infinite number of children. The set of these children is called *generation one*, and the edges from the root to the children are then labeled by the realizations of a Poisson process on $[0, \infty)$ that has constant intensity $\mu > 0$. That is, each edge from the root is assigned a unique element of the set

$$\mathcal{P}(\mu) = \{\xi_k : k = 1, 2, \dots\} \quad \text{where} \quad \xi_k = Y_1 + Y_2 + \dots + Y_k$$

and the random variables $\{Y_j : j = 1, 2, \dots\}$ are independent and

$$\mathbb{P}(Y_j > x) = \exp(-\mu x) \quad \text{for all } j = 1, 2, \dots \text{ and } x \in [0, \infty).$$

After generation k has been defined, one defines generation $k + 1$ by taking each element of generation k and applying the same construction that we applied to the root to get the first generation. At each stage the Poisson process that is used to label the edges is taken to be independent of all of the other Poisson processes that have been introduced. This construction is then continued until there is a well defined generation for each of the natural numbers $k = 1, 2, \dots$

A tree \mathcal{T} that is produced by this construction is said to be a PWIT with intensity $\mu > 0$, and, as shorthand, we will write

$$\mathcal{T} \stackrel{d}{=} \text{PWIT}(\mu),$$

whenever \mathcal{T} has the same distribution as the PWIT we have just constructed. Shortly, we will be more precise about the metric space in which one understands this distributional equality to take place.

3.2 Components of the PWIT

If G is any graph with a real number associated to each edge of G , then G is called a weighted graph, and the numbers on the edges are called the edge lengths. Given such a graph, we let $G(x)$ denote the graph that one obtains when all of the edges of length x or greater are delete, and if G is a rooted graph we also let $G_\star(s)$ denote the component of $G(s)$ that contains the root. If \mathcal{T} is a PWIT with intensity μ and root r , then we may again view $\mathcal{T}_\star(s)$ as a rooted graph with root

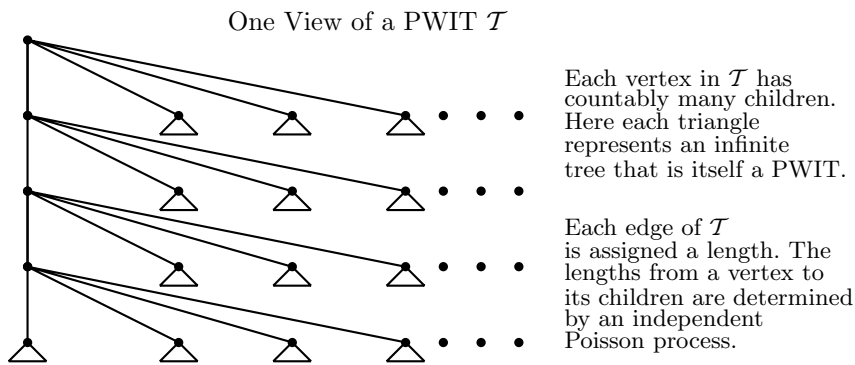


Figure 1: The PWIT is arguably the most fundamental limit object in the theory of randomly weighted graphs. It is the local weak limit of many different sequences, and it offers a unified approach to limit theorems for matching, spanning trees, and many other problems of combinatorial optimization.

r , and this graph turns out to be an old friend. It is nothing more than a Poisson Galton-Watson tree.

More precisely, if $PGW(s)$ denotes the distribution of the random tree determined Galton-Watson branching process with a single progenitor and an offspring distribution that is Poisson with mean s , then we have

$$\mathcal{T} \stackrel{d}{=} \text{PWIT}(\mu) \Rightarrow \mathcal{T}_*(s) \stackrel{d}{=} \text{PGW}(s\mu).$$

Many pleasing computations may be based on this simple observation.

In particular, we will need a qualitative understanding of the size of $\mathcal{T}_*(s)$ when $\mu = 1$, but everything we need has been known for 120 years or more. Specifically, the probability $p = p(s)$ that a $PGW(s)$ branching process is finite (the so-called *extinction* probability) is one if $0 \leq s \leq 1$ and for $s > 1$ the value of p is given by the unique root in $(0, 1)$ of the equation.

$$p = \exp(-s(1 - p)).$$

Lagrange-Bürmann inversion provides an explicit formula

$$p(s) = \frac{1}{s} \sum_{k=1}^{\infty} k^{-k} \frac{(se^s)^k}{k!} \quad \text{for } s > 1,$$

but, despite its charm, this sum does not always provide the best way to understand $p(s)$, or the complementary probability $q(s) = 1 - p(s)$. Here we will rely more directly on the fact that $q(s)$ is the unique strictly positive solution of

$$1 - q(s) = \exp(-sq(s)) \quad \text{when } s > 1, \tag{17}$$

and the obvious inverse relationship

$$s(q) = -\frac{\log(1 - q)}{q}. \tag{18}$$

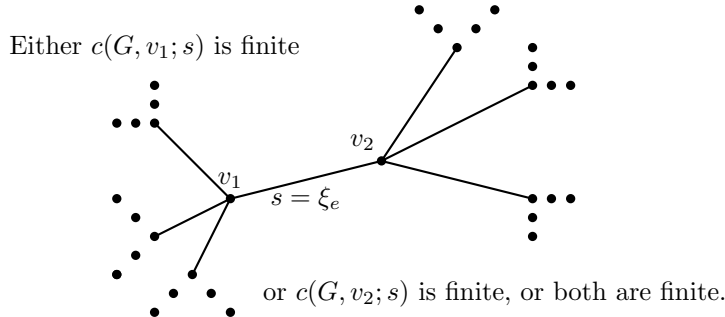


Figure 2: The edge $e = (v_1, v_2)$ of G is in the minimal spanning forest if and only if at least one of the trees $c(G, v_1; s)$ and $c(G, v_2; s)$ is finite when $s = \xi_e$.

that gives us the value of s for which we have we have probability $0 < q < 1$ that the total PGW(s) population is infinite.

3.3 Minimal Spanning Forests

The *minimal spanning forest* of an infinite graph G that has all distinct edge lengths is the subgraph $\text{MSF}(G)$ of G with the same vertex set as G and with an edge set that contains each edge $e = (v_1, v_2)$ of G for which

- (1) $c(G, v_1; s)$ and $c(G, v_2; s)$ are disjoint, and
- (2) $c(G, v_1; s)$ and $c(G, v_2; s)$ are *not both infinite*

when s is taken to be the length of the edge $e = (v_1, v_2) \in G$. An illustration of this definition is given in Figure 2 from Aldous and Steele (2002).

The real utility of this definition can only be brought out by the PWIT Limit Theorem (Theorems 4.2), but a good exercise with the definition is first to show that each component of $\text{MSF}(G)$ must be infinite and then to argue that $\text{MSF}(G)$ is indeed free of cycles.

3.4 Zeta Meets a PWIT

Let \mathcal{T} be a PWIT with intensity $\mu = 1$, and let r denote its root. If $\text{MSF}(\mathcal{T})$ is the minimal spanning forest of \mathcal{T} , then by a natural extension of our earlier notation we denote the sum of the edges incident to the root by

$$Y(\mathcal{T}) = \sum_{e:r \in e} \xi_e \mathbb{I}(e \in \text{MSF}(\mathcal{T})),$$

and denote the associated power sum by

$$Y_\alpha(\mathcal{T}) = \sum_{e:r \in e} \xi_e^\alpha \mathbb{I}(e \in \text{MSF}(\mathcal{T})).$$

The next lemma exploits the method of Lemma 4 of Aldous and Steele (2002) to obtain a slightly more general result. Although the innovation is minor, there do seem to be genuine benefits to having the parameter α at one's disposal. At a minimum, the joint presence of the gamma and zeta functions is amusing.

Lemma 3.1 *The sum of the α th powers of edges of the PWIT that are incident to the root has expectation*

$$\mathbb{E}[Y_\alpha(\mathcal{T})] = 2\Gamma(1 + \alpha)\zeta(2 + \alpha) \quad \text{for } \alpha \in (-1, \infty), \quad (19)$$

and by analytic continuation the same formula holds for all complex α for which the left-hand side is well defined; consequently, one has the Mellin integral representation

$$\mathbb{E}[Y_\alpha(\mathcal{T})] = \frac{2}{1 + \alpha} \int_0^\infty \frac{x^{\alpha+1}}{e^x - 1} dx \quad \text{for all } \operatorname{Re} \alpha > 0. \quad (20)$$

Proof: If one conditions a Poisson process \mathcal{P} on $[0, \infty)$ to have a point at s , then $\mathcal{P} \setminus \{s\}$ is again a Poisson process, so, if we condition on the event that there is an edge e of length s incident to the root, then the subtrees obtained by cutting that edge are again independent $\text{PGW}(s)$ trees.

Now, since the probability that at least one of these is finite is equal to $1 - q^2(s)$, we see that this is also the probability that the edge e is in the minimal spanning forest of T and we have

$$\mathbb{E}[Y_\alpha(\mathcal{T})] = \int_0^\infty s^\alpha (1 - q^2(s)) ds. \quad (21)$$

To compute the integral, we apply integration-by-parts, the implicit formula for (17) for $q(s)$, and the fact that $q(s)$ vanishes on $[0, 1]$ to find

$$\begin{aligned} (1 + \alpha)\mathbb{E}[Y_\alpha(\mathcal{T})] &= 2 \int_0^\infty s^{1+\alpha} q(s) q'(s) ds = 2 \int_1^\infty s^{1+\alpha} q(s) q'(s) ds \\ &= 2 \int_0^1 \frac{\log^2(1 - q)}{q^\alpha} dq. \end{aligned}$$

A good table would now suffice, but it is as easy to substitute $u = -\log(1 - q)$ to find

$$\begin{aligned} (1 + \alpha)\mathbb{E}[Y_\alpha(\mathcal{T})] &= 2 \int_0^\infty u^{1+\alpha} \frac{e^{-u}}{1 - e^{-u}} du = 2 \int_0^\infty u^{1+\alpha} \sum_{k=1}^\infty e^{-ku} du \\ &= 2 \sum_{k=1}^\infty \frac{1}{k^{2+\alpha}} \Gamma(2 + \alpha) = 2\Gamma(\alpha + 2)\zeta(\alpha + 2). \end{aligned}$$

Since one has $(1 + \alpha)\Gamma(1 + \alpha) = \Gamma(2 + \alpha)$, the proof of formula (19) is complete.

Finally, the analytic continuation of the identity (19) follows from the general principles of function theory, and the validity of the Mellin integral representation (20) is embedded in our calculations. Alternatively, one can note that the representation (20) follows from formula (19) and the well known formula

$$\Gamma(z)\zeta(z) = \int_0^\infty \frac{x^{z-1}}{e^x - 1} dx \quad \operatorname{Re} z > 1,$$

which one can prove by expanding $(e^x - 1)^{-1}$ as a geometric series. \square

4 Local Weak Convergence Theory

We now need to recall (and to modestly extend) some basic facts from the theory of local weak convergence. The main results in this section are the PWIT Convergence Theorem (Theorem 4.2) and the MST Convergence Theorem (Theorem 4.4). The first of these is implicit in Aldous (1992) and Aldous (2001), and, although Theorem 4.2 is nominally more general than the results that were needed in Aldous (1992) and Aldous (2001), no essentially new ideas are needed. Finally, the MST Convergence Theorem is a direct import from Aldous and Steele (2002).

The real benefit of the present development of the PWIT Limit Theorem is that it is reasonably self-contained. Thus, with very little overhead, one gains direct access to the single most important fact about the PWIT.

4.1 A Poisson Convergence Lemma

We begin with a lemma that is surely part of classic folklore, but the snappy proof via Rényi's characterization of the Poisson process appears to be new. At a minimum, this proof draws the straightest possible line between the hypotheses on F and the required Poisson limit.

Lemma 4.1 *Let F denote a distribution function such that*

$$F(0) = 0 \quad \text{and} \quad F(x) = \mu x + o(x) \quad \text{as } x \rightarrow 0. \quad (22)$$

If the random variables of the triangular array $\{\xi_{i,n} : 1 \leq i \leq d_n\}$ are independent within each row and if one has

$$P(\xi_{i,n} \leq x) = F(x/d_n) \quad \text{for all } 1 \leq i \leq d_n,$$

where $d_n \rightarrow \infty$ as $n \rightarrow \infty$, then one has the weak convergence

$$S_n = \{\xi_{i,n} : 1 \leq i \leq d_n\} \xrightarrow{d} \mathcal{P}(\mu) \quad \text{as } n \rightarrow \infty$$

in the sense of point processes.

Proof: By Rényi's characterization of the Poisson process (cf. Rényi (1967) or Kingman (1993), pp. 34–37), it suffices to show that for each union of disjoint intervals $A = (a_1, b_1] \cup (a_2, b_2] \cup \dots \cup (a_n, b_n]$ one has

$$P(|S_n \cap A| = 0) \rightarrow e^{-\lambda(A)\mu}, \quad (23)$$

where $\lambda(A)$ denotes the Lebesgue measure of A . By our hypothesis on F and the independence of the $\{\xi_{i,n} : 1 \leq i \leq d_n\}$, we have

$$\begin{aligned} P(|S_n \cap A| = 0) &= \left(1 - \sum_{i=1}^k \{F(b_i/d_n) - F(a_i/d_n)\}\right)^{d_n} \\ &= (1 - \mu\lambda(A)/d_n + o(\mu\lambda(A)/d_n))^{d_n}, \end{aligned}$$

so the limit (23) follows instantly. \square

There is a sense in which Rényi's criterion is modestly magical; it provides us with independence of a different sort than we assume at the beginning. Also

one should note that the only sly aspect of Rényi's Theorem is the requirement that one deal with all A that can be written as finite unions of disjoint intervals; in fact, Moran (1967) shows by example that one cannot get by with less. Finally, there is one small technical point; we have used Rényi's *characterization* of the Poisson process to provide *convergence criterion* for of a sequence of processes. Naturally, one only needs to apply the usual subsequence argument to pass from the characterization to the convergence criterion.

4.2 Local Weak Convergence Defined

We now need to extend the classical notion of weak convergence for point processes to a larger domain that is more directly connected with the convergence of weighted graphs. The treatment given here follows the exposition of Aldous and Steele (2002) which was designed in part to systematize the basic constructions used in Aldous (1992) and Aldous (2001).

To begin, we consider a graph G with a vertex set $v(G)$ that may be finite or countable. We further suppose there is a function ℓ from the edge set $e(G)$ to $(0, \infty]$, and we call $\ell(e)$ the length of the edge e . We then use ℓ to define a metric on $v(G)$ by taking the distance from u to v as the infimum over all paths between u and v of the sum of the lengths of the edges in the path. Naturally, the distance from any vertex v to itself is taken to be zero.

Now, if G is a connected graph with a countable or infinite vertex set and if ℓ is an edge length function that makes G locally finite in the sense that for each vertex v and each real $\rho < \infty$ the number of vertices within distance ρ from v is finite, then we call G a *geometric graph*. Also, when there is a distinguished vertex v , we say that G is a rooted geometric graph with *root* v , and to save space, we denote the set of geometric graphs by \mathcal{G} and the set of rooted geometric graphs by \mathcal{G}_* .

The key issue is to say what one means for a sequence $\{G_n\}$ of elements of \mathcal{G}_* to converge to a G in \mathcal{G}_* . The driving idea is that for large n , the rooted geometric graph G_n should look very much like G in a neighborhood of the root of G that is as large as we like.

Formally, we take $\rho > 0$ and let $N_\rho(G)$ denote the graph whose vertex set $V_\rho(G)$ is the set of vertices of G that are at a distance of at most ρ from the root of G and whose edge set consists of just those edges of G that have both vertices in $V_\rho(G)$. One again views $N_\rho(G)$ as an element of \mathcal{G}_* with edge length function and root given by those of G . Also, $\rho > 0$ is called a *continuity point* of G if no vertex of G is exactly at a distance ρ from the root of G .

Now, at last, we say that G_n *converges* to G_∞ in \mathcal{G}_* provided that for each continuity point ρ of G_∞ there is an $n_0 = n_0(\rho, G_\infty)$ such that for all $n \geq n_0$ there exists a isomorphism² $\gamma_{n,\rho}$ from the rooted geometric graph $N_\rho(G_\infty)$ to the rooted geometric graph $N_\rho(G_n)$ such that for each edge e of $N_\rho(G_\infty)$ the length of $\gamma_{n,\rho}(e)$ converges to the length of e as $n \rightarrow \infty$.

This definition determines a topology that makes \mathcal{G}_* into a complete separable metric space. As a consequence, and it gives us access to the usual tools of weak convergence theory. Here, if $\{X_n\}$ is a sequence of \mathcal{G}_* -valued random variables and X is a \mathcal{G}_* -valued random variable we write

$$X_n \xrightarrow{d} X \quad \text{to mean that} \quad \mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)]$$

²Graphs G and G' are isomorphic provided that is a bijection $\phi : v(G) \rightarrow v(G')$ such that $(\phi(u), \phi(v)) \in e(G')$ if and only if $(u, v) \in e(G)$.

for each bounded continuous function $f : \mathcal{G}_* \rightarrow \mathbb{R}$. This is just plain vanilla weak convergence \mathcal{G}_* -valued random variables, but to emphasize the special attention that is paid to the neighborhood of the root we also say that we have the *local weak convergence* of X_n to X .

From examples one finds that local weak convergence is a perfectly natural notion, despite the fact that it takes a while to make precise. In fact, the only subtle feature about local weak convergence is the way in which it force one to focus so myopically on the neighborhoods of the root.

4.3 The PWIT Limit Theorem

We now have the background in place to prove the theorem that makes us interested in PWIT; it shows that the PWIT arises as the limit of a very natural sequence of geometric graphs. As noted earlier, this particular version of the PWIT limit theorem is intended to make the PWIT limit ideas from Aldous (1992) and Aldous (2001) more explicit, more accessible, and modestly more general

Theorem 4.2 (The PWIT Limit Theorem) *Let $G_n, n = 1, 2, \dots$, denote a sequence of graphs such that the vertex set $v(G_n)$ has cardinality n for each $n = 1, 2, \dots$ and such that the maximum and minimum degrees satisfy the degree conditions*

$$\Delta(G_n) \sim d_n \quad \text{and} \quad \delta(G_n) \sim d_n \quad \text{as} \quad d_n \rightarrow \infty. \quad (24)$$

Also, let F denote a distribution function that satisfies the conditions (22) and associate to each $e \in v(G_n)$ an independent edge length ξ_e with distribution

$$P(\xi_e \leq x) = F(x/d_n) \quad \text{for all } x \geq 0.$$

Next, independently choose an element of $v(G_n)$ according to the uniform distribution, and let \mathbf{G}_n denote the rooted geometric graph produced by this construction. One then has

$$\mathbf{G}_n \xrightarrow{d} \text{PWIT}(\mu).$$

As one often does in the theory of weak convergence, we prove this limit theorem by passage to an equivalent characterization theorem. Specifically, one first argues (in a step that we leave as an exercise) that the sequence of \mathcal{G}_* -valued random variables $\{\mathbf{G}_n\}$ is tight. Then we consider an arbitrary subsequence, say $\{n_k : k = 1, 2, \dots\}$, and we note by tightness that there must exist a further subsequence $\{m_k : k = 1, 2, \dots\}$ and a \mathcal{G}_* -valued random variable \mathbf{G} such

$$\mathbf{G}_{m_k} \xrightarrow{d} \mathbf{G} \quad \text{as} \quad n \rightarrow \infty.$$

Next, we observe by Skorohod's theorem (cf. Dudley (1989), pp. 325–327) that one can assume without loss of generality that

$$\mathbf{G}_{m_k} \rightarrow \mathbf{G} \quad \text{almost surely as } n \rightarrow \infty,$$

and now all we have to do is to prove that \mathbf{G} is actually a PWIT.

From the definition of the topology of local weak convergence, we know automatically that \mathbf{G} is connected, so \mathbf{G} will be a tree provided that we show that it has no cycles. This will follow from the next lemma.

For the statement of the lemma, we note that a *path-plus-cycle* is a graph that can be written as a path plus one additional edge that makes a cycle by joining two vertices on the path. Also, to anticipate the application of the lemma, one should recall that the root of \mathbf{G}_n is randomly chosen uniformly from the vertex set $v(G_n)$, so the lemma immediately implies that with probability one the limit graph \mathbf{G} has no cycles in any ρ neighborhood of its root.

Lemma 4.3 *Let $S(n, \rho)$ the set of all vertices $v \in v(\mathbf{G}_n)$ for which there exists a path-plus-cycle $H = H(v) \subset \mathbf{G}_n$ such that*

$$v \in H \quad \text{and} \quad \sum_{e \in e(H)} \xi_e \leq \rho.$$

One then has

$$|S(n, \rho)|/n \xrightarrow{p} 0 \quad \text{as} \quad n \rightarrow \infty.$$

Proof: We first note that the number of path-plus-cycle subgraphs of \mathbf{G}_n with k vertices cannot be larger than

$$n \cdot \Delta(G_n)^{k-1} \binom{k}{2} < \frac{1}{2} n k^2 \Delta(G_n)^{k-1}.$$

Also, by our hypothesis on F , we know there is an x_0 such that $F(x) \leq 2\mu x$ for all $0 \leq x \leq x_0$, and from this bound, integration by parts, and induction one finds a corresponding bound for the k -fold convolution is given by

$$F^{(k)}(x) \leq (2\mu)^k \frac{x^k}{k!} \quad \text{for all } 0 \leq x \leq x_0.$$

Thus, for any k edges e_1, e_2, \dots, e_k of \mathbf{G}_n we find from the distributional assumption $\mathbb{P}(\xi_e \leq x) = F(x/d_n)$ that

$$P(\xi_{e_1} + \xi_{e_2} + \dots + \xi_{e_k} \leq \rho) \leq (2\mu)^k \frac{\rho^k}{d_n^k k!} \quad \text{for all } 0 \leq \rho/d_n \leq x_0.$$

The expected number vertices of \mathbf{G}_n that are contained in path-plus-cycle subgraph \mathbf{G}_n with k vertices and total length bounded by ρ is therefore no larger than

$$(2\mu)^k n k^2 \frac{\Delta(G_n)^{k-1} \rho^k}{d_n^k k!} \quad \text{provided that } 0 \leq \rho \leq x_0 d_n.$$

Now, since $\Delta(G_n) \sim d_n$, we may chose a constant $C = C(\mu, \rho)$ such that this bound is not larger than $nC^k/d_n k!$, and, thus, one finds

$$\mathbb{E}(|S(n, \rho)|) \leq e^C n/d_n, \quad \text{for all } n \text{ such that } d_n \geq \rho/x_0.$$

Since we assume $d_n \rightarrow \infty$, this bound is more than one needs to complete the proof of the lemma. \square

Now that we know \mathbf{G} is a tree, the proof of Theorem 4.2 will be complete provided that we confirm that the edge lengths from each vertex to its children are given by the realization of an independent Poisson process. For the root of \mathbf{G} this

is quite easy. When we look at the edges incident to the root of \mathbf{G}_n for large n , we see by Lemma 4.1 that the lengths of these edges are approximately the points of a Poisson process, and consequently the lengths of the edges incident to the root of \mathbf{G} must exactly follow a Poisson process.

Now consider a fixed ρ and an n so large that the probability that \mathbf{G}_n contains a cycle in the ρ -neighborhood of the root of \mathbf{G}_n is small. We know that the distances to the children of the root approximately follow the initial segment of a Poisson process, and now we consider the second generation. Let c be a fixed child of the root r , and consider the set S of edges incident to c . The set of edge lengths $\{\xi_e : e \in S \text{ and } e \neq (r, c)\}$ again satisfy the assumptions of Lemma 4.1, so the distribution of the lengths of the descendants of c will again follow a Poisson process as closely as we like. This argument shows that the first two generations of \mathbf{G} are consistent with the construction of the PWIT. There is no change to the argument as one moves from the k th to the $k + 1$ st generations, so one finds that \mathbf{G} is indeed a PWIT and the proof of Theorem 4.2 is complete.

4.4 Convergence of MSTs

We now need a general result from Aldous and Steele (2002) that tells us that the local weak convergence of a sequence of randomly rooted graphs automatically gives us the local weak convergence of their associated MSTs.

Theorem 4.4 (MST Convergence Theorem) *Let G_∞ denote a \mathcal{G}_* -valued random variable such that with probability one G_∞ has infinitely many vertices and no two of the edges of G have the same length. Further, let $\{G_n : n = 1, 2, \dots\}$ denote a sequence of \mathcal{G}_* -valued random variables such that for each n the distribution of G_n is given by the standard construction and such that for each n the vertex set of G_n has cardinality n with probability one. If*

$$G_n \xrightarrow{d} G_\infty \quad \text{as } n \rightarrow \infty, \quad (25)$$

then one has the joint weak convergence in $\mathcal{G}_* \times \mathcal{G}_*$,

$$(G_n, \text{MST}(G_n)) \xrightarrow{d} (G_\infty, \text{MSF}(G_\infty)). \quad (26)$$

Further, if N_n denotes the degree of the root of $\text{MST}(G_n)$ and N denotes the degree of the root of $\text{MSF}(G_\infty)$

$$N_n \xrightarrow{d} N \quad \text{and} \quad \mathbb{E}[N_n] \rightarrow \mathbb{E}[N] = 2, \quad (27)$$

and, if L_n denotes the sum of lengths of the edges incident to the root of $\text{MST}(G_n)$ and L denotes the corresponding quantities for $\text{MSF}(G_\infty)$, then

$$L_n \xrightarrow{d} L. \quad (28)$$

4.5 Closing the Loops

Theorems 1.2 is now a remarkably easy corollary of the PWIT Limit Theorem, the MST Convergence Theorem, and the exact PWIT calculation developed in Section 3. The first step is simply to make the link between $L_{\text{MST}}^\alpha(\mathbf{G}_n)$ and $Y_\alpha(\mathbf{G}_n)$ more explicit.

Here it is useful to let $\mathbb{E}_R[f(\mathbf{G}_n)]$ denote the conditional expectation of $f(\mathbf{G}_n)$ given $\{\xi_e : e \in v(\mathbf{G}_n)\}$; in other words, we just average $f(\mathbf{G}_n)$ over the possible values of the random root R . We now just compute

$$\begin{aligned} L_{\text{MST}}^\alpha(\mathbf{G}_n) &= d_n^{-\alpha} \sum_e (d_n \xi_e)^\alpha \mathbb{I}(e \in \text{MST}(\mathbf{G}_n)) \\ &= d_n^{-\alpha} \frac{1}{2} \sum_v \sum_e (d_n \xi_e)^\alpha \mathbb{I}(e \in \text{MST}(\mathbf{G}_n)) \\ &= \frac{1}{2} d_n^{-\alpha} |v(\mathbf{G}_n)| \mathbb{E}_R \left[\sum_e (d_n \xi_e)^\alpha \mathbb{I}(e \in \text{MST}(\mathbf{G}_n) \text{ and } R(\mathbf{G}_n) \in e) \right] \\ &= \frac{1}{2} d_n^{-\alpha} |v(\mathbf{G}_n)| \mathbb{E}_R[Y_\alpha(\mathbf{G}_n)]. \end{aligned}$$

Finally, if one takes expectations in this representation one finds the basic identity

$$\mathbb{E}[L_{\text{MST}}^\alpha(\mathbf{G}_n)] = \frac{1}{2} d_n^{-\alpha} |v(\mathbf{G}_n)| \mathbb{E}[Y_\alpha(\mathbf{G}_n)]. \quad (29)$$

Now, by the PWIT limit theorem we already know that $\mathbf{G}_n \xrightarrow{d} \mathcal{T}$, and by the MST Convergence Theorem this automatically entails $\text{MST}(\mathbf{G}_n) \rightarrow \text{MSF}(\mathcal{T})$. From the defining topology of local weak convergence, we then have

$$\sum_{e \in e(\mathbf{G}_n)} \xi_e^\alpha \mathbb{I}(e \in \text{MST}(\mathbf{G}_n) \text{ and } R \in e) \xrightarrow{d} \sum_{e \in e(\mathcal{T})} \xi_e^\alpha \mathbb{I}(e \in \text{MSF}(\mathcal{T}) \text{ and } R \in e),$$

or in other words

$$Y_\alpha(\mathbf{G}_n) \xrightarrow{d} Y_\alpha(\mathcal{T}). \quad (30)$$

Now, if $\{Y(G_n) : n = 1, 2, \dots\}$ uniformly integrable, we can take expectations in the limit (30) and apply our earlier calculation of $\mathbb{E}[Y_\alpha(\mathcal{T})]$ to find

$$\mathbb{E}[Y_\alpha(\mathbf{G}_n)] \rightarrow 2\Gamma(1 + \alpha)\zeta(2 + \alpha),$$

but by the introductory identity (29), this is equivalent to

$$\mathbb{E}[L_{\text{MST}}^\alpha(\mathbf{G}_n)] \sim d_n^{-\alpha} |v(\mathbf{G}_n)| \Gamma(1 + \alpha)\zeta(2 + \alpha), \quad (31)$$

so the direct part of Theorem 1.2 is complete.

The converse now comes almost for free. One first notes that we may reverse the path from the limit (31) to the convergence of the expectations $\mathbb{E}[Y_\alpha(\mathbf{G}_n)]$, so when one pairs this fact with the limit (30), the loop is closed by applying the following simple lemma.

Lemma 4.5 *If a sequence of nonnegative random variables X_n , $n = 1, 2, \dots$ converges to X in distribution, then one has*

$$\mathbb{E}[X_n] \rightarrow \mathbb{E}[X] \quad \text{if and only if} \quad \{X_n : n = 1, 2, \dots\} \quad \text{is uniformly integrable.}$$

Proof: If we assume that the sequence $\{X_n : n = 1, 2, \dots\}$ is uniformly integrable then convergence in probability implies X_n convergence in L^1 and this certainly implies that one has the convergence of the expectations. For the converse, we first note that by the Skorohod embedding theorem (Dudley (1989), pp. 325–327), there is no loss of generality if we assume that X_n converges almost surely to X . In this case, the nonnegativity and convergence of the expectations implies that X_n converges to X in L^1 by Sheffé’s lemma (Williams (1991), p. 55). Since L^1 convergence is stronger than uniform integrability of $\{X_n : n = 1, 2, \dots\}$, the proof of the lemma is complete. \square

4.6 An Illustrative Example

Theorem 1.2 tells us that the limit behavior of $E[L_{\text{MST}}^\alpha(G_n)]$ is determined once one shows the uniform integrability of the sequence $\{Y_\alpha(G_n)\}$. The systematic treatment of this question will be left for another time, but one should note that this reasonably crude and qualitative property of $\{Y_\alpha(G_n)\}$ often follows from known results. Nevertheless, there are certainly many situations where uniform integrability fails or where the proof of uniform integrability can be subtle.

The example we consider here is illustrated in Figure 3, and it has also been used for illustration in Beveridge, Frieze, and McDiarmid (1998) and Frieze, Ruzinkó, and Thoma (2000). If one takes $G_n = C(m_n, K_n^{-1})$, then one can see just from Frieze’s $\zeta(3)$ theorem that under the model of uniformly distributed costs that one has

$$E[L_{\text{MST}}(G_n)] \sim m_n(\zeta(3) + 1) \quad \text{as } n \rightarrow \infty$$

for *any* choice of the sequence $2 \leq m_n < \infty$. On the other hand, uniform integrability of $\{Y(G_n) : n = 1, 2, \dots\}$ would imply that

$$E[L_{\text{MST}}(G_n)] \sim m_n \zeta(3),$$

so in this case we certainly know $\{Y(G_n) : n = 1, 2, \dots\}$ is not uniformly integrable. Nevertheless, one might want to check this directly, and, in fact, a moment’s thought about the impact of a random root is all one needs to show

$$\limsup_{n \rightarrow \infty} E[Y(G_n) \mathbb{I}(Y(G_n) \geq t)] \geq 1 \quad \text{for all } 0 \leq t < \infty.$$

The situation is more interesting in case one takes $0 < \alpha < 1$, in which case one may now show that

$$\lim_{t \rightarrow \infty} \limsup_{n \rightarrow \infty} E[Y_\alpha(G_n) \mathbb{I}(Y_\alpha(G_n) \geq t)] = 0,$$

so the sequence $\{Y_\alpha(G_n)\}$ is uniform integrable. As a consequence, one obtains a positive result for $C(m_n, K_n^{-1})$ that asserts

$$E[L_{\text{MST}}^\alpha(C(m_n, K_n^{-1}))] \sim \Gamma(\alpha + 1) \zeta(\alpha + 2) m_n n^{1-\alpha},$$

a fact which is perhaps more amusing when made more concrete. If one takes $m_n = \lfloor n^\alpha \rfloor$ and $\alpha = 0.99$, then $\Gamma(1.99) = 0.995 \dots$ and $\zeta(2.99) = 1.204 \dots$, so one finds

$$E[L_{\text{MST}}^{0.99}(C(n, K_n^{-1}))] \sim cn \quad \text{where } c = \Gamma(1.99) \zeta(2.99) = 1.1990 \dots$$

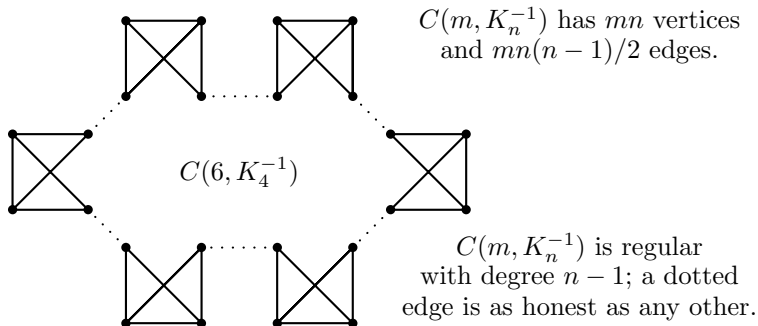


Figure 3: The graph $C(6, K_4^{-1})$ is built out of 6 copies of K_4 that have had one edge removed. These altered graphs — called K_4^{-1} s — are then chained together in a cycle. The result is a regular graph with degree 3 and $6 \cdot 4 = 24$ vertices; in grocer’s terms one has a cubic graph with two dozen vertices.

5 Concluding Observations

As noted earlier, this is a report on work-in-progress and there are many loose ends that time and diligence may suffice to resolve. Perhaps the most compelling questions that have been left open concern the monotonicity and concavity of $E[L_{\text{MST}}(K_n)]$ under the uniform model. Next on the list would be the possible analog of Parisi’s conjecture and an exploration of the relationship of $E[L_{\text{MST}}(K_n)]$ to $s_n = 1 + 1/2^3 + \dots + 1/(n-1)^3$ under the exponential model.

More generally, the exact formula (2) for $E[L_{\text{MST}}(G)]$ provides one with considerable motivation to work out detailed representations for the Tutte polynomials for those graphs that are of most interest in probability theory. One also suspects that interesting consequences may flow from the interpretation of formula (2) in the light of general results for the Tutte polynomial. Specifically, one might speculate that results like the Negami Splitting Formula (Negami (1987)) could lead to interesting inferences. Finally, the appearance of the logarithmic derivative of the Tutte polynomial in formula (2) suggests that this rational function may have informative properties beyond those it inherits from the Tutte polynomial.

The questions that are left open from the second part of the report are less well formed. Certainly, the question of uniform integrability of $\{Y_\alpha(G_n)\}$ deserves more systematic thought. Right now the easiest paths to uniform integrability freload on the efforts of the more direct approaches to $E[L_{\text{MST}}(G)]$, especially the recent arguments of Frieze, Ruzinkó, and Thoma (2000) that exploit the lovely bound of Karger (1999) on the number of approximately minimal cuts. Nevertheless, as the easy example of Subsection 4.6 suggests, the sequence $\{Y_\alpha(G_n)\}$ does have an independent character. One suspects that in time the direct investigation of its uniform integrability will lead to arguments that do not poach on other approaches.

Finally, the second part of the report suggests several overarching questions from the theory of local weak convergence. One major line of investigation that

surely deserves a sustained effort is the extension of the MST Convergence Theorem. There are many other classes of subgraphs from combinatorial optimization for which one expects an analogous result, and the class of confirmed examples is growing. Nevertheless, the final form of this theory is nowhere in sight.

ACKNOWLEDGEMENTS: It is a pleasure to thank David Aldous for continued coaching in the theory of local weak convergence and to thank Ira Gessel for providing an electronic extension to $n = 15$ of the table of Tutte polynomials for K_n given in Gessel and Sagan (1996).

References

- [1] D.J. Aldous (1992) *Asymptotics in the random assignment problem*, Probab. Th. Rel. Fields, 93, 507–534.
- [2] D.J. Aldous (2001) *The $\zeta(2)$ limit in the random assignment problem*, Random Structures Algorithms, 18, 381–418.
- [3] D.J. Aldous and J.M. Steele (2002) *The Objective Method and the Theory of Local Weak Convergence* in *Discrete and Combinatorial Probability* (ed. H. Kesten), Springer-Verlag, New York. [in press]
- [4] N. Alon, A. Frieze, and D. Welsh (1994) *Polynomial time randomized approximation schemes for the Tutte polynomial of dense graphs*, pp. 24–35 in the Proceedings of the 35th Annual Symposium on the Foundations of Computer Science (S. Goldwasser, ed.), IEEE Computer Society Press, 1994.
- [5] F. Avram and D. Bertsimas (1992) *The minimum spanning tree constant in geometric probability and under the independent model: a unified approach*, Annals of Applied Probability 2, 113–130.
- [6] A. Beveridge, A. Frieze, C. McDiarmid (1998) *Minimum length spanning trees in regular graphs*, Combinatorica 18, 311–333.
- [7] R.M. Dudley (1989) *Real Analysis and Probability*, Wadworth Publishing, Pacific Grove CA.
- [8] A.M. Frieze (1985) *On the value of a random minimum spanning tree problem*, Discrete Appl. Math. 10, 47–56.
- [9] A.M. Frieze and C.J.H. McDiarmid (1989) *On random minimum length spanning trees*, Combinatorica, 9, 363–374.
- [10] A. M. Frieze, M. Ruszinkó, and L. Thoma (2000) *A note on random minimum length spanning trees*, Electronic Journal of Combinatorics 7, Research Paper 5, (5 pp.)
- [11] I.M. Gessel and B.E. Sagan (1996) *The Tutte polynomial of a graph, depth-first search, and simplicial complex partitions*, Electronic Journal of Combinatorics 3, Research Paper 9, (36 pp.)
- [12] T.H. Harris (1989) *The Theory of Branching Processes*, Dover Publications, New York.

- [13] S. Janson (1995) *The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph*, Random Structures and Algorithms 7, 337–355.
- [14] D. Karger (1999) *A randomized fully polynomial time approximation scheme for the all-terminal network reliability problem*, SIAM Journal of Computing 29, 492–514.
- [15] J.F.C. Kingman (1993) *Poisson Processes*, Oxford University Press, New York, 1993.
- [16] P.A.P. Moran (1967) *A non-Markovian quasi-Poisson process*, Studia Scientiarum Mathematicarum Hungarica 2, 425–429.
- [17] S. Negami (1987) *Polynomial Invariants of Graphs*, Transactions of the American Mathematical Society 299, 601–622.
- [18] G. Parisi (1998) *A conjecture on random bipartite matching*, ArXiv Condmat 9801176.
- [19] M. Penrose (1998) *Random minimum spanning tree and percolation on the n -cube*, Random Structures and Algorithms 12, 63–82.
- [20] A. Rényi, (1967) *Remarks on the Poisson Process*, Studia Scientiarum Mathematicarum Hungarica 2, 119–123.
- [21] J.M. Steele (1987) *On Frieze’s $\zeta(3)$ limit for lengths of minimal spanning trees*, Discrete Applied Mathematics, 18 (1987), 99–103.
- [22] J.M. Steele (1997) *Probability Theory and Combinatorial Optimization*, NSF-CBMS Volume 69. Society for Industrial and Applied Mathematics, Philadelphia.
- [23] D. Welsh (1999) *The Tutte Polynomial*, Random Structures and Algorithms 15, 210–228.
- [24] D. Williams (1991) *Probability with Martingales*, Cambridge University Press, Cambridge.

J. Michael Steele

Department of Statistics, Wharton School
 University of Pennsylvania
 Philadelphia, PA 19104
 steele@wharton.upenn.edu