MINORS OF TREE DISTANCE MATRICES

HARRY RICHMAN, FARBOD SHOKRIEH, AND CHENXI WU

Abstract. We prove a formula for the determinant of a principal minor of the distance matrix of a weighted tree. This generalizes a result of Graham and Pollak.

Contents

1. Introduction

Suppose $G = (V, E)$ is a tree with n vertices. Let D denote the distance matrix of G. In graham-pollak [4], Graham and Pollak proved that

eq:full-det
$$\det D = (-1)^{n-1}2^{n-2}(n-1). \tag{1}$$

This identity is remarkable in that the result does not depend on the tree structure, beyond the number of vertices. This identity was motivated by a problem in data communication, and inspired much further research on distance matrices.

The main result of this paper is to generalize (eq:full-det 1) by replacing D with any principal submatrix. For $S \subset V(G)$, let $D[S]$ denote the principal minor consisting of the S-indexed rows and columns.

thm:main
Theorem 1. *Suppose G is a tree with n vertices, and distance matrix D. Let $S \subset V(G)$ be a subset of vertices. Then*

eq:main
$$\det D[S] = (-1)^{|S|-1}2^{|S|-2}\left((n-1)\,\kappa(G/S) - \sum_{\mathcal{F}_2(G/S)} k(F_*)^2\right). \tag{2}$$

where G/S denotes the quotient graph that identifies together vertices in S, $\kappa(G/S)$ is the number of S-rooted spanning forests of G, $\mathcal{F}_2(G/S)$ is the set of $(S,*)$-rooted spanning forests of G, F_* denotes the $*$-component of F, and

$$k(F_*) = 2 - \deg^o(F, *).$$

Date: v1, April 29, 2022 (Preliminary draft, not for circulation).

Note: the quantities $c(F,*)$, $k(F_*)$ satisfy

$$1 \le c(F,*) \le |S|, \qquad 2 - |S| \le k(F_*) \le 1.$$

When $S = V$ is the full vertex set, the quotient graph G/V consists of a single vertex with $n-1$ loop edges, so $\kappa(G/V) = 1$ and $\mathcal{F}_2(G/V) = \emptyset$.

Weighted version: A weighted version of (1) was proved by Bapat–Kirkland–Neumann [1].

eq:w-full-det

$$\det D_\alpha = (-1)^{n-1} 2^{n-2} \prod_{e\in E} \alpha_e \sum_{e \in E} \alpha_e. \tag{3}$$

thm:w-main

Theorem 2. *Suppose G is a finite, weighted tree, and $A \subset V(G)$ is a subset of vertices. Then*

eq:w-main

$$\det D[A] = (-1)^{|S|-1} 2^{|S|-2} \left(\sum_{E(G)} \ell(e) \sum_{\mathcal{T}(G/S)} w(T) - \sum_{\mathcal{F}_2(G/S)} w(F) k(F_*)^2 \right). \tag{4}$$

where $\mathcal{T}(G/A)$ denotes the set of A-rooted spanning forests of G, F_2 varies over all $(A,*)$-rooted spanning forests of G, F_* denotes the $*$-component of F.

It is worth observing that the distances appearing in $D[S]$ may ignore a large part of the ambient tree G. We could instead replace G by the subtree consisting of paths between vertices in S, which we call $\mathrm{conv}(S,G)$, the *convex hull* of $S \subset G$. To apply formula (2) "efficiently", we should replace G with this convex hull $\mathrm{conv}(S,G)$. However, the formula as stated is true even without this replacement due to cancellation of terms.

Corollary 3.

$$\frac{\det D[S]}{\mathrm{cof}\, D[S]} = \frac{1}{2} \left(\sum_{E(G)} \ell(e) - \frac{\sum_{\mathcal{F}_2(G/S)} w(F) k(F_*)^2}{\sum_{\mathcal{T}(G/S)} w(T)} \right). \tag{5}$$

Theorem 4 (Monotonicity of principal minor ratios)**.** *Suppose $G = (V,E)$ is a finite, weighted tree with distance matrix D.*

(1) If $S \subset V(G)$ is nonempty,

$$0 \le \frac{\det D[S]}{\mathrm{cof}\, D[S]} \le \frac{1}{2} \sum_{E(G)} \ell(e).$$

(2) If $\mathrm{conv}(S,G)$ denotes the subtree of G consisting of all paths between points of $S \subset V(G)$,

$$\frac{\det D[S]}{\mathrm{cof}\, D[S]} \le \frac{1}{2} \sum_{E(\mathrm{conv}(S,G))} \ell(e).$$

(3) If $A, B \subset V(G)$ are nonempty subsets with $A \subset B$, then

$$\frac{\det D[A]}{\mathrm{cof}\, D[A]} \le \frac{\det D[B]}{\mathrm{cof}\, D[B]}.$$

Theorem 5 (Nonsingular minors)**.** *Let G be a finite, weighted tree with distance matrix D, and let $S \subset V(G)$ be a subset of vertices. If $|S| \ge 2$ then $\det D[S] \neq 0$.*

1.1. **Previous work.** A formula for the inverse matrix D^{-1} was found by Graham and Lovász in [3].

1.2. **Notation.** G a finite graph, loops and parallel edges allowed, possibly disconnected

$E(G)$ edge set of G

$V(G)$ vertex set of G

2. Background

For background on enumeration problems for graphs and trees, see Moon moon [5].

The following theorem is due to Kirchhoff. For any graph G, let $\kappa(G)$ denote the number of spanning trees of G.

Theorem 6 (All-minors matrix tree theorem)**.** *Let $G = (V, E)$ be a finite graph. Let L denote the Laplacian matrix of G. Then for any nonempty vertex set $S \subset V(G)$,*

$$\det L[V \setminus S] = \kappa(G/S). \tag{6}$$

Note that $\kappa(G/S)$ is also the number of S-rooted spanning forests of G.

Theorem 7 (bapat-sivasubramanian [2])**.** *Let T be a tree with $m+1$ vertices and m edges. Let D be the distance matrix of T, and L the Laplacian matrix. Let $S \subset V(T)$ be a subset of vertices of T. Then*

$$\operatorname{cof} D[S] = (-2)^{|S|-1} \det L[V \setminus S].$$

2.1. **Trees and forests.**

3. Proofs

Outline of proof: given subset S and distance matrix minor $D[S]$, we will

(1) Find vector $\mathbf{m} \in \mathbb{R}^S$ such that $D[S]\mathbf{m} = \lambda \mathbf{1}$.

(2) Compute the sum of entries of $\mathbf{m}$, i.e. $\mathbf{1}^T\mathbf{m}$.

(3) Note the identity

$$\mathbf{1}^T\mathbf{m} = \lambda(\mathbf{1}^T D[S]^{-1}\mathbf{1}) = \lambda \frac{\operatorname{cof} D[S]}{\det D[S]}.$$

where $\operatorname{cof} D[S]$ is the sum of cofactors of $D[S]$.

(4) Use known expression for $\operatorname{cof} D[S]$ to compute

$$\det D[S] = \lambda(\operatorname{cof} D[S]) \left(\mathbf{1}^T\mathbf{m}\right)^{-1}.$$

The interesting part of this expression will be located in the constant λ.

Example 8. Suppose G is a tree consisting of three paths joined at a central vertex. Let S consist of the central vertex, and the three endpoints of the paths. The corresponding minor of the distance matrix is

$$D[S] = \begin{bmatrix} 0 & a & b & c \\ a & 0 & a+b & a+c \\ b & a+b & 0 & b+c \\ c & a+c & b+c & 0 \end{bmatrix}.$$

Following the steps outlined above:

(1) The vector $\mathbf{m} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ satisfies $D[S]\mathbf{m} = (a+b+c)\mathbf{1}$

(2) The sum of entries of $\mathbf{m}$ is $\mathbf{1}^T\mathbf{m} = 2$.

(3) We have
$$2 = \mathbf{1}^T\mathbf{m} = \lambda(\mathbf{1}^T D[S]\mathbf{1}) = \lambda\frac{\operatorname{cof} D[S]}{\det D[S]}.$$
(4) The cofactor sum is $\operatorname{cof} D[S] = -8abc$, so the determinant is
$$\det D[S] = \lambda\frac{\operatorname{cof} A}{\mathbf{1}^T\mathbf{m}} = (a+b+c)(-8abc)\frac{1}{2} = -4(a+b+c)abc.$$

Proposition 9. *Let $T = (V, E)$ a tree, and consider the vector $\mathbf{m} \in \mathbb{R}^V$ defined by*
$$\mathbf{m}(v) = 2 - \deg v,$$
where $\deg v$ denotes the degree of v in T. Then $\mathbf{1}^T\mathbf{m} = \sum_{v\in V}(2 - \deg v) = 2$.

Let $\mathbf{1}$ denote the all-ones vector. When we choose a subset $S \subset V(G)$, we no longer have a single "obvious" replacement for $\mathbf{m}$ inside $\mathbb{R}^S$. Instead, we can take an average over S-rooted spanning forests.

In the outline above, our first goal is to find a "special" vector $\mathbf{m} \in \mathbb{R}^S$ satisfying $D[S]\mathbf{m} = \lambda\mathbf{1}$. We can approach this first goal as follows: consider $\mathbb{R}^S$ inside the larger vector space $\mathbb{R}^V = \mathbb{R}^S \oplus \mathbb{R}^{V\setminus S}$, and we wish to find vectors $\mathbf{n}_i \in \mathbb{R}^V$ satisfying $\pi_S(D\mathbf{n}_i) = \lambda_i\mathbf{1}$. By finding sufficiently many such vectors $\mathbf{n}_i$, we can hope to find a linear combination that lies inside $\mathbb{R}^S \oplus \{0\}$.

Proposition 10. *Suppose $v \in V \setminus S$. For each $s \in S$, let $\mu(v, s) =$ current flowing to s when unit current enters G at v and G is grounded at S. Explicitly,*
$$\mu(v, s) = \frac{\#\text{ of } S\text{-rooted spanning forests of } G \text{ whose } s\text{-component contains } v}{\#\text{ of } S\text{-rooted spanning forests of } G}$$
$$= \frac{\sum_{\mathcal{T}(G/S)} \mathbb{1}(v \in T(s))}{\kappa(G/S)}$$
Consider the vector $\mathbf{n} \in \mathbb{R}^V$ defined by
$$\mathbf{n}(v) = 1, \qquad \mathbf{n}(s) = -\mu(v, s) \text{ if } s \in S, \qquad \mathbf{n}(w) = 0 \text{ if } w \notin S \cup v$$
Then $\pi_S(D\mathbf{n}) = \lambda\mathbf{1}$ for some λ.

Proof sketch. For any $s, s' \in S$, consider tracking value of $D\mathbf{n}$ along path from s to s'. The value of $D\mathbf{n}$ changes according to current flow in the corresponding system, i.e. $D\mathbf{n}$ records electrical potential. By assumption S is grounded, so $D\mathbf{n}$ takes the same value at s and s'. □

Note that we can express the tree distance $d(v, w)$ as a sum over edges
$$d(v, w) = \sum_{e\in E(G)} \delta(e; v, w) \qquad \text{where } \delta(e; v, w) = \begin{cases} 1 & \text{if } e \text{ lies on } v \sim w \text{ path}, \\ 0 & \text{otherwise.} \end{cases}$$

- If we fix v and w, then $\delta(-; v, w) : E(G) \to \{0, 1\}$ is the indicator function for the unique v w path in G.
- On the other hand if we fix e and v, then the deletion $G\setminus e$ has two connected components, and $\delta(e; v, -) : V(G) \to \{0, 1\}$ is the indicator function for the component of $G \setminus e$ not containing v.

Theorem 11. *Let G be a tree, S a subset of vertices, and $D[S]$ the corresponding minor of the distance matrix. Suppose $\mathbf{m}_S \in \mathbb{R}^S$ is defined by*

$$\mathbf{m}_S(v) = \sum_{T \in \mathcal{T}(G/S)} \sum_{w \in T_v} (2 - \deg w) = \sum_{T \in \mathcal{T}(G/S)} 2 - c(T, v).$$

Then $D[S]\mathbf{m} = \lambda \mathbf{1}$ for some constant λ.

Proof. Note that

$$\mathbf{m} = \kappa(G/S) \left(\sum + \sum_{v \in V \setminus S} (\deg v - 2)\mathbf{n}_v = \right)$$

□

Theorem 12. *With $\mathbf{m}$ defined as above, $D[S]\mathbf{m} = \lambda \mathbf{1}$ for*

$$\lambda = \sum_{\mathcal{T}(G/S)} w(T) \sum_{E(G)} \ell(e) - \sum_{\mathcal{F}_2(G/S)} w(F) k(F, *)^2. \tag{7}$$

where $c(T, w)$ is the "cut index" of the w-component of T (as a spanning forest).

Proof. We have

$$\begin{aligned}
(D[S]\mathbf{m})(v) &= \sum_{s \in S} d(v, s)\mathbf{m}(s) \\
&= \sum_{s \in S} \left(\sum_{e \in E(G)} \ell(e)\delta(e; v, s) \right) \left(\sum_{T \in \mathcal{T}(G/S)} w(T)(2 - \deg^o(T, s)) \right) \\
&= \sum_{T} \sum_{e} \ell(e) w(T) \sum_{s \in S} (2 - \deg^o(T, s))\delta(e; v, s) \\
&= \sum_{T} \sum_{e} \ell(e) w(T) \sum_{s \in S^*(e,v)} (2 - \deg^o(T, s)).
\end{aligned}$$

where

$$S^*(e, v) = \{s \in S : e \text{ lies on path from } v \text{ to } s\}.$$

If $e \in \operatorname{conv}(G, S)$, then $S^*(e, v)$ is nonempty and we have

$$\sum_{s \in S^*(e,v)} (2 - c(T, s)) = \begin{cases} 1 & \text{if } e \notin T, \\ 1 - (2 - c(T \setminus e, *)) & \text{if } e \in T(s'),\ s' \in S^*(e, v), \\ 1 + (2 - c(T \setminus e, *)) & \text{if } e \in T(s'),\ s' \notin S^*(e, v) \end{cases}$$

Here $c(T \setminus e, *)$ refers to the cut index of the "free" component of the spanning forest $T \setminus e$.

If $e \notin \operatorname{conv}(G, S)$ on the other hand, $S^*(e, v)$ is empty and we have

$$\sum_{s \in S^*(e,v)} 2 - c(T, w) = 0 = 1 - (2 - \deg^o(T \setminus e, *))^2,$$

since $\deg^o(T \setminus e, *) = 1$ in this case. From ($\overset{\text{above}}{\textbf{??}}$) we have

$$(D[S]\mathbf{m})(v) = \sum_{e} \sum_{T} \ell(e) w(T)(1 - f(v, e, T))$$

where

$$f(v,e,T) = \begin{cases} 0 & \text{if } e \notin T \\ 2 - \deg^o(T \setminus e, *) & \text{if } e \in T(s') \text{ for some } s' \notin S^*(e,v) \\ -(2 - \deg^o(T \setminus e, *) & \text{if } e \in T(s') \text{ for some } s' \in S^*(e,v) \end{cases}$$

(Note that $e \notin \operatorname{conv}(S,G)$ implies $e \in T$.)

$$\begin{aligned} (D[S]\mathbf{m})(v) - \sum_e \sum_T \ell(e)w(T) = &-\sum_T \sum_{e \in T \setminus T(S^*)} \ell(e)w(T)(2 - \deg^o(T \setminus e, *)) \\ &+ \sum_T \sum_{e \in T(S^*)} \ell(e)w(T)(2 - \deg^o(T \setminus e, *)) \end{aligned}$$

The deletion $T \setminus e$ is an $(S,*)$-rooted spanning forest of G, so we may rewrite the above expression in terms of $\mathcal{F}_2(G/S)$.

$$\begin{aligned} (1) = &-\sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F,*)) \sum_T \sum_{e \in T \setminus T(S^*)} \mathbb{1}(F = T \setminus e) \\ &+ \sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F,*)) \sum_T \sum_{e \in T(S^*)} \mathbb{1}(F = T \setminus e) \end{aligned}$$

Finally, we note that $F = T \setminus e$ is equivalent to $T = F \cup e$, and in particular this only occurs when we choose $e \in \partial F(*)$:

$$\begin{aligned} (1) = &-\sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F,*)) \sum_{e \in \partial F(*)} \mathbb{1}(\delta(e; v, F(*)) = 1) \\ &+ \sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F,*)) \sum_{e \in \partial F(*)} \mathbb{1}(\delta(e; v, F(*)) = 0) \end{aligned}$$

Finally, we observe that for any forest F,

$$\#\{e : e \in \partial F(*),\, \delta(e; v, F(*)) = 0\} = \deg^o(F,*) - 1$$

and

$$\#\{e : e \in \partial F(*),\, \delta(e; v, F(*)) = 1\} = 1$$

Thus

$$\begin{aligned} (1) &= -\sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F,*))(1) \\ &\quad + \sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F,*))(\deg^o(F,*) - 1) \\ &= -\sum_{F \in \mathcal{F}_2(G/S)} w(F)(2 - \deg^o(F,*))^2. \end{aligned}$$

The set $\mathcal{F}_2(G/S)$ of $(S,*)$-rooted spanning forests of G can be partitioned into two types: "active" and "inactive".

$$\mathcal{F}_2(G/S) = \mathcal{F}_2^{in}(G/S) \sqcup \mathcal{F}_2^{out}(G/S),$$

where

$$\begin{aligned} \mathcal{F}_2^{in}(G/S) &= \{F \in \mathcal{F}_2(G/S) \text{ such that } \deg^o(*,F) \geq 2\}, \\ \mathcal{F}_2^{out}(G/S) &= \{F \in \mathcal{F}_2(G/S) \text{ such that } \deg^o(*,F) = 1\}. \end{aligned}$$

Moreover, for a given spanning forest $F \in \mathcal{F}_2(G/S,*)$, there are exactly $c(F,*)$ choices of pairs $(T,e) \in \mathcal{T}(G/S) \times E(G)$ such that

$$F = T \setminus e.$$

Consider the map

$$E(G) \times \mathcal{T}(G/S) \to \mathcal{F}_2(G/S)$$

defined by ...

$$(e,T) \mapsto T \setminus e.$$

For a forest F in $\mathcal{F}_2(G/S)$, the preimage under this map has $c(F,*)$ elements.

Therefore (let $\mathcal{T} = \mathcal{T}(G/S)$)

$$\begin{aligned}
\sum_{T\in\mathcal{T}}\sum_{e\in E}(\cdots)) &= \sum_{T\in\mathcal{T}}\sum_{e\in E(G)} 1 + \sum_{T\in\mathcal{T}}\sum_{e\in E}(...) \\
&= \sum_{T\in\mathcal{T}}\sum_{e\in E} 1 - \sum_{T\in\mathcal{T}}\sum_{e\in E}\mathbb{1}(e\in T(v))(2-c(T\setminus e,*)) \\
&\qquad + \sum_{T\in\mathcal{T}}\sum_{e\in E}\mathbb{1}(e\in T\setminus T(v))(2-c(T\setminus e,*)) \\
&= \sum_{T\in\mathcal{T}}\sum_{e\in E} 1 - \sum_{T\in\mathcal{T}}\sum_{e\in E}\mathbb{1}(e\in T(v))\sum_{F\in\mathcal{F}_2}\mathbb{1}(F=T\setminus e)(2-c(F,*)) \\
&\qquad + \sum_{T\in\mathcal{T}}\sum_{e\in E}\mathbb{1}(e\in T\setminus T(v))\sum_{F\in\mathcal{F}_2}\mathbb{1}(F=T\setminus e)(2-c(F,*)) \\
&= \sum_{F\in\mathcal{F}_2}(2-c(F,*))\sum_{T\in\mathcal{T}}\sum_{e\in E}\mathbb{1}(F=T\setminus e)(\mathbb{1}(e\in T\setminus T(v))-\mathbb{1}(e\in T(v))) \\
&= \sum_{F\in\mathcal{F}_2}(2-c(F,*))\sum_{e\in E}\mathbb{1}(e\notin F)(\mathbb{1}(e\notin(F\cup e)(v))-\mathbb{1}(e\in(F\cup e)(v))) \\
&= \sum_{F\in\mathcal{F}_2}(2-c(F,*))((c(F,*)-1)-1) \\
&= -\sum_{F\in\mathcal{F}_2}(2-c(F,*))^2.
\end{aligned}$$

□

Proposition 13. *Let $G=(V,E)$ be a tree, and $S \subset V$. Suppose we label $S = \{s_1,\ldots,s_r\}$ and $V \setminus S = \{t_1,\ldots,t_{n-r}\}$. For each $t_i \in V\setminus S$, consider $\mathbf{f}_i \in \mathbb{R}^V$ defined by*

Example 14. If

$$D[S\cup t] = \begin{bmatrix} 0 & a & b & c \\ a & 0 & a+b & a+c \\ b & a+b & 0 & b+c \\ c & a+c & b+c & 0 \end{bmatrix}$$

then

$$\begin{bmatrix} 0 & a & b & c \\ a & 0 & a+b & a+c \\ b & a+b & 0 & b+c \\ c & a+c & b+c & 0 \end{bmatrix}\begin{bmatrix} ab+ac+bc \\ -bc \\ -ac \\ -ab \end{bmatrix} = \begin{bmatrix} -3abc \\ -abc \\ -abc \\ -abc \end{bmatrix}$$

4. Physical interpretation

If we consider G as a network of wires with each edge containing a unit resistor, which is grounded at all nodes in S, then $\mathbf{m}_S$ records the currents flowing to S when current is added on $V \setminus S$ in the amount $2 - \deg v$ for each $v \notin S$.

5. Examples

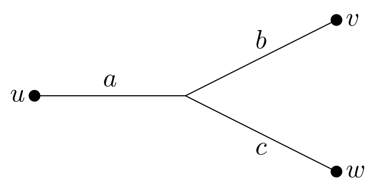
Example 15. Suppose G is a tree consisting of three paths joined at a central vertex. Let S consist of the central vertex, and the three endpoints of the paths. The corresponding minor of the distance matrix is

$$D[S] = \begin{bmatrix} 0 & a & b & c \\ a & 0 & a+b & a+c \\ b & a+b & 0 & b+c \\ c & a+c & b+c & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & a & b & c \\ a & -a & a & a \\ b & b & -b & b \\ c & c & c & -c \end{bmatrix} \sim \begin{bmatrix} 0 & a & b & c \\ a & -2a & 0 & 0 \\ b & 0 & -2b & 0 \\ c & 0 & 0 & -2c \end{bmatrix}.$$

The determinant is

$$\det D[S] = -4(a+b+c)abc.$$

Example 16. Suppose Γ is a tripod with lengths a, b, c and corresponding leaf vertices u, v, w.



Let $B = \{u, v, w\}$. Then

$$D[B] = \begin{bmatrix} 0 & a+b & a+c \\ a+b & 0 & b+c \\ a+c & b+c & 0 \end{bmatrix}.$$

and

$$\det D[B] = 2(a+b)(a+c)(b+c) = 2\left((a+b+c)(ab+ac+bc) - abc\right).$$

The "special vector" in this example is $\mathbf{m}^T = \begin{bmatrix} a(b+c) & b(a+c) & c(a+b) \end{bmatrix}$.

6. Further work

richman-shokrieh-wu

See [6].

Acknowledgements

The authors would like to thank Ravindra Bapat for helpful discussion.

References

bapat-kirkland-neumann [1] R. Bapat, S. J. Kirkland, and M. Neumann. On distance matrices and Laplacians. *Linear Algebra Appl.*, 401:193–209, 2005.

bapat-sivasubramanian [2] R. B. Bapat and S. Sivasubramanian. Identities for minors of the Laplacian, resistance and distance matrices. *Linear Algebra Appl.*, 435(6):1479–1489, 2011.

graham-lovasz [3] R. L. Graham and L. Lovász. Distance matrix polynomials of trees. *Adv. in Math.*, 29(1):60–88, 1978.

graham-pollak [4] R. L. Graham and H. O. Pollak. On the addressing problem for loop switching. *Bell System Tech. J.*, 50:2495–2519, 1971.

moon [5] J. W. Moon. *Counting labelled trees*. Canadian Mathematical Monographs, No. 1. Canadian Mathematical Congress, Montreal, Que., 1970. From lectures delivered to the Twelfth Biennial Seminar of the Canadian Mathematical Congress (Vancouver, 1969).

richman-shokrieh-wu [6] D. H. Richman, F. Shokrieh, and C. Wu. Capacity on metric graphs, 2022. in preparation.