ABOUT SUBSPACES THE MOST DEVIATING FROM THE COORDINATE ONES

YURI NESTERENKO

Abstract. Taking the largest principal angle as the distance function between same dimensional nontrivial linear subspaces in $\mathbb{R}^n$, we describe the class of subspaces deviating from all the coordinate ones by at least $\arccos(1/\sqrt{n})$. This study compliments and is motivated by the long-standing hypothesis put forward in [6] and essentially stating that so-defined distance to the closest coordinate subspace cannot exceed $\arccos(1/\sqrt{n})$. In this context, the subspaces presented here claim to be the extremal ones.

Realized as the star spaces of all nontrivial 2-connected series-parallel graphs with certain edge weights and arbitrary edge directions, the given subspaces may be of interest beyond numerical linear algebra within which the original problem was formulated.

1. Introduction

In this paper we study the real case of the problem about linear subspaces the most deviating from the coordinate ones.

Problem. *Given $n > k > 0$, what is the largest possible deviation of a k-dimensional subspace of $\mathbb{R}^n$ from all the coordinate subspaces of the same dimension, if the deviation is measured as the largest principal angle?*

Initially formulated in [6] as a hypothesis about matrices and their submatrices, it contains the estimate.

Hypothesis. *For every $n > k > 0$ and an arbitrary real $n \times k$ matrix with orthonormal columns, such a $k \times k$ submatrix exists that the spectral norm of its inverse is bounded above by $\sqrt{n}$.*

According to this, the answer to the question posed in Problem 1 does not exceed $\arccos(1/\sqrt{n})$[1]. Our computer experiments described in the next section allow us to suggest that in reality the equality holds, and it is achieved on a finite number of subspaces. The study presented here began with counting these subspaces for various n and k, and resulted in their formal description as the star spaces of all nontrivial 2-connected series-parallel graphs on $k+1$ vertices and n edges equipped with certain positive weights and arbitrary directions.

---

[1]Since principle angles are arccosines of singular values of the corresponding submatrices.

2. Numerical experiments

In this section, we describe the numerical experiments we conducted to study the problem, and also give some definitions necessary to formulate the main result. Pseudocode of the corresponding algorithms is available in the Appendix.

In our first experiments we represented k-dimensional subspaces by $n \times k$ matrices. Starting with an initial guess subspace sampled uniformly distributed on the corresponding real Grassmannian and using successive random perturbations, we maximized the value of the *target function* measuring the least deviation of the given subspace from all the same-dimensional coordinate subspaces. Doing this for a large number of samples, we accumulated a set of the resulting subspaces corresponding to the greatest target function value. In all the experiments we have conducted for various n and k, this value was equal to $\arccos(1/\sqrt{n})$. Based on these empirical results, we will call such subspaces *extremal*.

In order to reduce the number of stored (unique) subspaces, our algorithm skips new extremal subspaces symmetric to an already found with respect to coordinate permutations and inversions. This approach is correct due to invariance of the target function under the given transformations, and reduces the number of iterations necessary for the size of the accumulated set to reach a plateau.

Having carried out our experiments for all $1 \le k < n \le 9$, we decided to search for the resulting triangle of numbers (see Table 1) in the On-Line Encyclopedia of Integer Sequences [8]. To our surprise, the search ended with a single match found: "Triangle read by rows: number of isomorphism classes of series-parallel matroids of rank d[2] on n elements" [11].

n, k	1	2	3	4	5	6	7	8
2	1							
3	1	1						
4	1	1	1					
5	1	2	2	1				
6	1	3	4	3	1			
7	1	4	8	8	4	1		
8	1	5	14	19	14	5	1	
9	1	6	23	42	42	23	6	1

Table 1. Number of symmetry classes of extremal subspaces for various n and k.

Let us give several definitions needed further.

Definition 2.1. *A graphical matroid is called series-parallel if it is associated with a series-parallel graph.*

[2] k in our notations

Definition 2.2. *A graph is called series-parallel if it can be obtained from a loop or a bridge*[3] *by a sequence of series and parallel extensions, i.e. subdivisions and duplications of edges.*

The series-parallel matroids from [11] are associated with so-called *nontrivial 2-connected series-parallel graphs*, or simply *2-sp-graphs.*

Definition 2.3. *A graph is called 2-connected, if no removal of a single vertex disconnects it.*

Loops and bridges are called trivial 2-connected graphs. Thus, 2-sp-graphs are those that can be obtained by series and parallel extensions from a 2-cycle.

The rank of a graphical matroid equals to the number of edges in every spanning tree of the associated graph, which in case of serial-parallel graphs is one less than the number of vertices.

Having this graph-theoretic context, we decided to check if the subspaces we found for arbitrary $n > k > 0$ are actually associated somehow with 2-sp-graphs on $k + 1$ vertices and n edges. The most natural idea was to search for them among subspaces given by matrices of the form $W^{\frac{1}{2}}B^T$, where $B \in \mathbb{R}^{(k+1)\times n}$ is an incidence matrix of a 2-sp-graph, and $W = diag(w_1, w_2, \dots, w_n)$ is a diagonal matrix of positive edge weights. Notion of incidence matrix implies certain choice of edge directions, which however doesn't affect values of our target function.

Remark. *Since $W^{\frac{1}{2}}B^T$ is not orthogonal and even not of full rank, before computing the principle angles we need to multiply it by $(L^+)^{\frac{1}{2}}$ from the right, where L^+ is the Moore-Penrose inverse of the graph Laplacian $L = BWB^T$.*

To realize this idea, we modified our algorithm so that it would iterate over 2-sp-graphs (setting for them some arbitrary edge directions) and tune the edge weights so as to maximize the target function.

Remark. *Our function of distance between subspaces is invariant under taking orthogonal complements (see matrix CS-decomposition in* [5]*). In terms of 2-sp-graphs this duality looks like taking a graph dual (which is also a 2-sp-graph) with the inverted edge weights. The symmetry in Table 1 reflects this fact.*

Eventually, our expectations were met – for each tested 2-sp-graph the optimizer reached the desired target function value. It was an interesting separate task to analyze the corresponding edge weights given as a floating point numbers, but we will not dwell on that here. In order to formally describe them, we will requite the alternative definition of series-parallel graphs equivalent to Definition 2.2, and which in particular justifies their second name – *2-terminal series-parallel graphs.*

Definition 2.4. *A series-parallel graph G with terminal vertices l and r is either an edge (l, r) or can be constructed as in 1) or 2):*

[3] A single-edge graph.

1) G is a parallel composition of series-parallel graphs $G_1, G_2, \ldots, G_p$ with the terminals $l_1, l_2, \ldots, l_p$ identified into vertex l and the terminals $r_1, r_2, \ldots, r_p$ identified into vertex r, and denoted as

$$G = G_1 || G_2 || \ldots || G_p \,,$$

2) G is a serial composition of series-parallel graphs $G_1, G_2, \ldots, G_s$ with the terminals identified as follows $l \sim l_1, r_1 \sim l_2, \ldots, r_{s-1} \sim l_s, r_s \sim r$, and denoted as

$$G = G_1 \circ G_2 \circ \ldots \circ G_s \,.$$

Hereinafter, we will use the agreement that series and parallel compositions in the given recurrent definition alternate. For 2-sp-graphs we deal with, the final composition (or, conversely, the first decomposition) is always parallel.

According to Lemma 9 from [4], if (l, r) is an edge in a 2-sp-graph, l and r can be treated as the terminals.

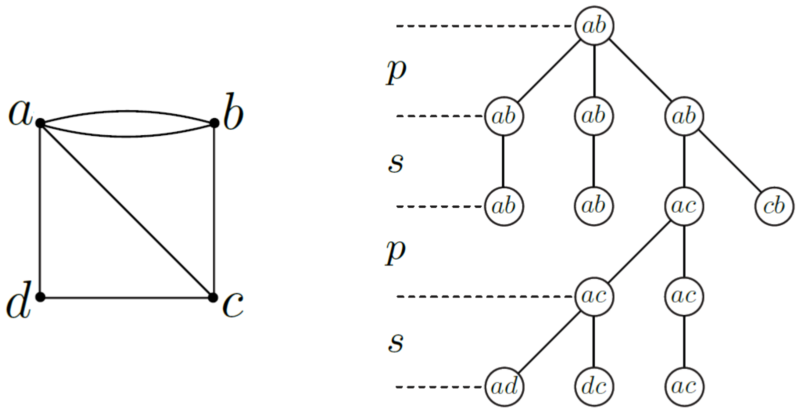

Figure 1. A series-parallel graph and its decomposition tree for terminal vertices a and b.

Every 2-terminal series-parallel graph can be fully described by its *decomposition tree*, which root corresponds to the whole graph and all the other nodes correspond to the subgraphs given by series-parallel decomposition. According to the agreement above, when moving away from the root to a leaf, the types of decompositions alternate starting from the parallel. Also, if any sequence of such decompositions ends with a parallel one, we will complete it by a dummy serial decomposition (see Figure 1).

We are now ready to formally describe the edge weights corresponding to the extremal subspaces.

3. Graph-induced edge weights

Hereinafter, $|\cdot|$ as applied to graphs, will denote the number of edges.

Definition 3.1. *Let G be a 2-sp-graph with a pair of vertices treated as terminals, and let*

$$G \supset G_1 \supset G_2 \supset \ldots \supset G_{2d-1} \supseteq G_{2d} = \{e\}$$

be a chain of nested subgraphs obtained by alternating parallel and serial decompositions. The edge weights given by the formula

$$w_e = \frac{\phi(|G_1|)}{\phi(|G_2|)} \cdots \frac{\phi(|G_{2d-1}|)}{\phi(|G_{2d}|)} \tag{3.1}$$

with $\phi(x) = x(|G| - x)$ will be referred to as G-induced.

Figure 2 illustrates formula 3.1 using decomposition tree.

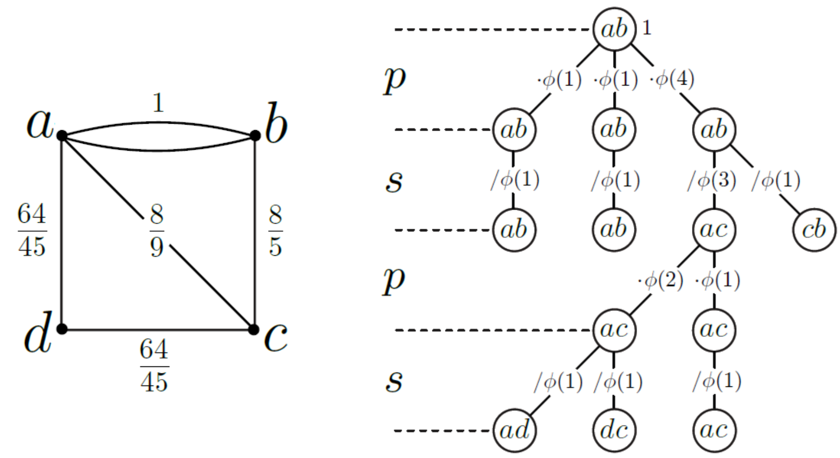

FIGURE 2. A series-parallel graph and its graph-induced edge weights.

Proposition 3.1. *Graph-induced edge weights are invariant to the choice of terminals up to a common factor.*

Proof. Let us consider two pairs of terminals in a 2-sp-graph and the corresponding decomposition trees t and $\tilde{t}$. Due to transitivity of the property being proven, it is sufficient to establish the invariance for the case when the second pair of terminals are vertices of some edge, further denoted as e. In terms of t, formula 3.1 defines the weight of edge e as

$$w_e = \frac{\phi(x_1)}{\phi(x_2)} \cdots \frac{\phi(x_{2d-1})}{\phi(x_{2d})}, \tag{3.2}$$

where $x_1 > x_2 > \ldots > x_{2d-1} \geq x_{2d} = 1$ are the leaf counts of the corresponding nested t's subtrees.

When we move from t to $\tilde{t}$, in terms of which the weight of e equals to 1, all other weights are divided by 3.2 as well. Indeed, let e' be an edge other than e with

$$w_{e'} = \frac{\phi(x_1)}{\phi(x_2)} \cdots \frac{\phi(x_{2q-1})}{\phi(x'_{2q})} \cdots \frac{\phi(x'_{2d'-1})}{\phi(x'_{2d'})}, \tag{3.3}$$

where x_{2q-1} is the number of leafs in the smallest common subtree of e and e' in t (the case of an even index is similar).

After division by 3.2 we obtain

$$\begin{aligned}\frac{w_{e'}}{w_e} &= \frac{\phi(x_{2d})}{\phi(x_{2d-1})}\cdots\frac{\phi(x_{2q})}{\phi(x'_{2q})}\cdots\frac{\phi(x'_{2d'-1})}{\phi(x'_{2d'})} = \\ &= \frac{\phi(n-x_{2d})}{\phi(n-x_{2d-1})}\cdots\frac{\phi(n-x_{2q})}{\phi(x'_{2q})}\cdots\frac{\phi(x'_{2d'-1})}{\phi(x'_{2d'})},\end{aligned} \tag{3.4}$$

where $n - x_{2d} > \ldots > n - x_{2q} > x'_{2q} > \ldots \geq x_{2d'}$ are the leaf counts of the corresponding nested subtrees, and which coincides with the formula for weight of e' in terms of $\tilde{t}$ (up to a formal extra multiplier equal to $\frac{\phi(n-1)}{\phi(n-1)}$ in case $x_{2d-1} = x_{2d} = 1$). □

Remark. *Since parallel composition of graphs corresponds to serial composition of their duals and vice versa, the graph-induced edge weights of any 2-sp-graph are reciprocals of the graph-induced edge weights of its dual.*

We are now ready to formulate the main result.

Theorem. *If G is a directed 2-sp-graph on $k+1$ vertices and n edges, and B is its incidence matrix and W is the diagonal matrix of G-induced weights, then the column space of $W^{\frac{1}{2}}B^T$ deviates from all the coordinate k-dimensional subspaces of $\mathbb{R}^n$ by at least* $\arccos(1/\sqrt{n})$.

Corollary. *The estimate from the Hypothesis cannot be improved for any n and k.*

Before we proceed to the proof, let us define another type of graph-induced coefficients required below.

Definition 3.2. *Let G be a directed 2-sp-graph with vertices l and r treated as terminals, and τ be its spanning tree. Let also $e \in \tau$ and*

$$G \supset G_1 \supset G_2 \supset \ldots \supset G_{2d-1} \supseteq G_{2d} = \{e\}$$

be the corresponding chain of nested subgraphs obtained by the alternating parallel and serial decompositions. The coefficients

$$y_e = \pm\frac{\psi(G_1)}{\psi(G_2)}\cdots\frac{\psi(G_{2d-1})}{\psi(G_{2d})}, \tag{3.5}$$

where the positive sign corresponds to e oriented from l to r and

$$\psi(\Gamma) = \begin{cases} |G| - |\Gamma|, & \text{if } \tau \cap \Gamma \text{ connects } \Gamma\text{'s terminals,} \\ -|\Gamma|, & \text{otherwise,} \end{cases}$$

will be referred to as (G,τ)-induced.

Figure 3 illustrates formula 3.5 using decomposition tree.

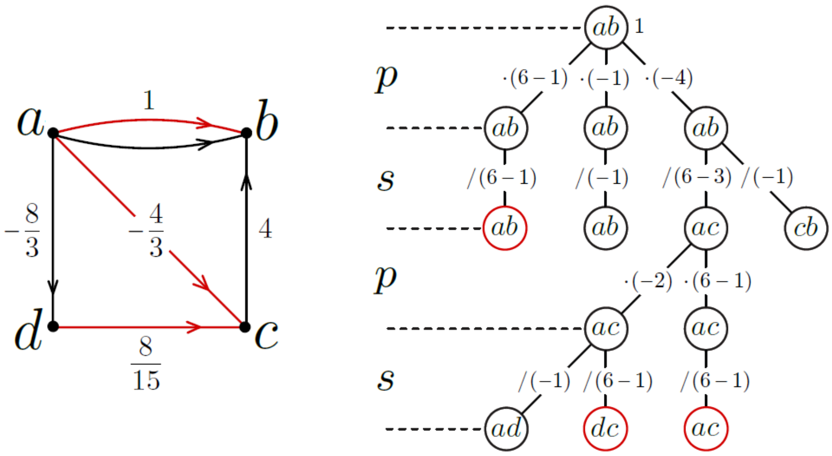

FIGURE 3. A directed series-parallel graph, its spanning tree (in red) and the induced edge coefficients.

We leave it as an exercise for the reader to prove the following property.

Proposition 3.2. *Graph-tree-induced edge coefficients are invariant to the choice of terminals up to a common factor.*

4. PROOF OF THE THEOREM

The matrix of orthogonal projection onto the column space of $W^{\frac{1}{2}}B^T$ is given by the formula

$$\mathcal{P} = W^{\frac{1}{2}}B^T L^+ B W^{\frac{1}{2}}.$$

To yield the desired claim, we will show that $\frac{1}{n}$ is the eigenvalue[4] of all submatrices of $\mathcal{P}$ corresponding to the spanning trees of G. All other submatrices correspond to subgraphs containing cycles, and hence are degenerate.

For convenience, instead of $\mathcal{P}$ we will work with the similar projection matrix

$$Y = W^{\frac{1}{2}}\mathcal{P}W^{-\frac{1}{2}} = WB^T L^+ B,$$

which submatrices are also similar to ones in $\mathcal{P}$. Known as the transfer current matrix (see [1]), Y has the following explicit form

$$Y = \frac{1}{T}\begin{array}{c} \begin{array}{ccccc} & e & & f & \end{array} \\ \left[\begin{array}{ccccc} & \vdots & & \vdots & \\ \cdots & T_{e+} & \cdots & T^{\rightrightarrows}_{e+f-} - T^{\rightleftarrows}_{e+f-} & \cdots \\ & \vdots & & \vdots & \\ \cdots & T^{\rightrightarrows}_{f+e-} - T^{\rightleftarrows}_{f+e-} & \cdots & T_{f+} & \cdots \\ & \vdots & & \vdots & \end{array}\right] \end{array} \begin{array}{c} \\ \\ e \\ \\ f \\ \\ \end{array}, \tag{4.1}$$

[4]While we observe that $\frac{1}{n}$ is also the least eigenvalue, we do not prove this here.

where T denotes the sum of weights of all spanning trees[5], T_{e+} denotes the sum of weights of those containing edge e, and $T^{\rightrightarrows}_{e+f-}$ and $T^{\rightleftarrows}_{e+f-}$ denote the similar sums for trees connecting f's vertices by the unique paths containing directed edge e and passing it in the positive and negative direction respectively (see Theorem 1.16 from [7]).

By Theorem 3 from [3], any two edges of a 2-sp-graph presented in a cycle as same oriented cannot be presented in a different cycle in the opposite way (and vice versa). This gives us the simpler form of matrix 4.1

$$Y = \frac{1}{T} \begin{array}{cl} \begin{array}{ccccc} \quad & e & \quad & f & \quad \end{array} & \\ \left[\begin{array}{ccccc} & \vdots & & \vdots & \\ \cdots & T_{e+} & \cdots & \pm T_{e+f-} & \cdots \\ & \vdots & & \vdots & \\ \cdots & \pm T_{f+e-} & \cdots & T_{f+} & \cdots \\ & \vdots & & \vdots & \end{array}\right] & \begin{array}{l} \\ e \\ \\ f \\ \\ \end{array} \end{array}, \tag{4.2}$$

where T_{f+e-} has the same meaning as given above but without taking into account edge directions. The latter affects the sign before the expression.

Now we fix arbitrary spanning tree τ and the corresponding submatrix $Y_\tau \in \mathbb{R}^{k\times k}$. Our next step is to show that $Y_\tau y_\tau = \frac{1}{n} y_\tau$, where y_τ is the restriction of the vector of (G,τ)-induced coefficients to the set of τ's edges. More specifically, we will show that

$$(TY_\tau y_\tau)(e_0) = \frac{T}{n} y_\tau(e_0), \tag{4.3}$$

where e_0 is an arbitrary edge of τ.

Let us denote e_0's vertices by l and r and treat them as the terminals. We will also assume without loss of generality that all G's edges are directed from l to r.

The given choice of terminals defines the following series-parallel decomposition

$$G = \{e_0\} \,||\, G_1 \,||\, \ldots \,||\, G_p =$$
$$= \{e_0\} \,||\, (G_{11} \circ \ldots \circ G_{1s_1}) \,||\, \ldots \,||\, (G_{p1} \circ \ldots \circ G_{ps_p}) = \ldots$$

Let $2d \geq 0$ be its maximum depth. For the corresponding edge counts we have

$$n = 1 + \sum_{i_1=1}^{p} n_{i_1} = 1 + \sum_{i_1=1}^{p} \sum_{j_1=1}^{s_{i_1}} n_{i_1 j_1} = \ldots =$$
$$= 1 + \sum_{i_1=1}^{p} \sum_{j_1=1}^{s_{i_1}} \cdots \sum_{i_d=1}^{p_{i_1 j_1 \ldots i_{d-1} j_{d-1}}} \sum_{j_d=1}^{s_{i_1 j_1 \ldots i_{d-1} j_{d-1} i_d}} n_{i_1 j_1 \ldots i_d j_d},$$

where the integer p's and s's denote the numbers of subgraphs in the corresponding parallel and serial decompositions.

[5]The weight of a spanning tree is the product of the weights of its edges.

Except the single-edge subgraph $\{e_0\}$, we will assume that the decomposition depth is uniform and equals to $2d$. This agreement can always be enforced by adding dummy parallel and serial decompositions of single-edge subgraphs (which doesn't affect neither G-induced edge weights nor (G, τ)-induced edge coefficients).

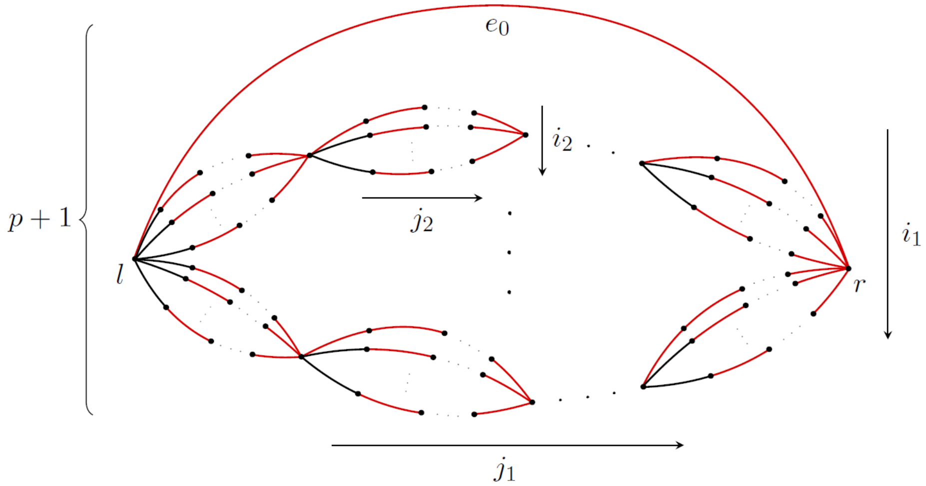

FIGURE 4. Graph G and the layout of τ (in red) for the case $d = 2$.

In order to write the recurrent formulas for such quantities as T and others similar to it, we will require the notion of 2-component forest – a spanning subgraph of a 2-terminal series-parallel graph composed of two disjoined trees rooted in its terminals[6]. Similar to trees, the weight of a 2-component forest is the product of the weights of its edges. The sum of such weights over all G's 2-component forests will be denoted as T'.

By Theorem 1 from [10], for arbitrary weighted graphs $\Gamma_1, \dots, \Gamma_m$

$$
\begin{aligned}
&T(\Gamma_1 \circ \ldots \circ \Gamma_m) = T(\Gamma_1)T(\Gamma_2)\ldots T(\Gamma_m),\\
&T(\Gamma_1 \,||\, \ldots \,||\, \Gamma_m) = \sum_{i=1}^{m} T'(\Gamma_1)\ldots T(\Gamma_i)\ldots T'(\Gamma_m),\\
&T'(\Gamma_1 \circ \ldots \circ \Gamma_m) = \sum_{i=1}^{m} T(\Gamma_1)\ldots T'(\Gamma_i)\ldots T(\Gamma_m),\\
&T'(\Gamma_1 \,||\, \ldots \,||\, \Gamma_m) = T'(\Gamma_1)T'(\Gamma_2)\ldots T'(\Gamma_m).
\end{aligned}
\tag{4.4}
$$

Based on recurrent formulas 4.4 and in order to avoid unnecessary additional notations, we will assume the following layout of τ in G:

1) in every nested parallel decomposition of G, the first subgraph contains τ's subtree,

[6]In particular, such 2-component forest can be obtained from τ by deletion of e_0.

2) in every nested serial decomposition of G, the first subgraph contains τ's 2-component subforest.

Figure 4 illustrates this agreement.

Based on 4.2, the left-hand side of 4.3 can be written as

$$(TY_\tau y_\tau)(e_0) = T_{0+}\, y_\tau(e_0) + \sum_{e\in\tau\setminus\{e_0\}} T_{e_0+e-}\, y_\tau(e). \tag{4.5}$$

By 3.5 and 4.4, the last sum denoted as P' satisfies the following recurrent equations

$$\begin{aligned}
P'_{..} &= \sum_{i=1}^{p_{..}}(-n_{..\,i})S'_{..\,i}\,, \quad \text{where } .. = i_1j_1\dots i_tj_t\,,\\
S'_{..} &= \frac{1}{(-n_{..\,1})}P'_{..\,1} + \sum_{j=2}^{s_{..}}\frac{1}{(n-n_{..\,j})}P_{..\,j}\,, \quad \text{where } .. = i_1j_1\dots i_{t+1}\,,\\
P_{..} &= (n-n_{..\,1})S_{..\,1} + \sum_{i=2}^{p_{..}}(-n_{..\,i})S'_{..\,i}\,, \quad \text{where } .. = i_1j_1\dots i_tj_t\,,\\
S_{..} &= \sum_{j=1}^{s_{..}}\frac{1}{(n-n_{..\,j})}P_{..\,j}\,, \quad \text{where } .. = i_1j_1\dots i_{t+1}\,,
\end{aligned} \tag{4.6}$$

for $t = 0,\dots,d-1$ and with the initial conditions $P'_{i_1j_1\dots i_dj_d} = 0$ and $P_{i_1j_1\dots i_dj_d} = T_{0+i_1j_1\dots i_dj_d-}$ for the finest refinement level corresponding to the single edge subgraphs. Here, $T_{0+i_1j_1\dots i_tj_t-}$ denotes the sum of weights of the spanning trees containing e_0 and connecting terminals of $G_{i_1j_1\dots i_tj_t}$ via e_0.

Excluding $S_{..}$ and $S'_{..}$ from 4.6 we obtain

$$\begin{aligned}
P'_{..} &= \sum_{i=1}^{p_{..}}(-n_{..\,i})\left(\frac{1}{(-n_{..\,i1})}P'_{..\,i1} + \sum_{j=2}^{s_{..\,i}}\frac{1}{(n-n_{..\,ij})}P_{..\,ij}\right)\,,\\
P_{..} &= (n-n_{..\,1})\sum_{j=1}^{s_{..\,1}}\frac{1}{(n-n_{..\,1j})}P_{..\,1j} +\\
&\quad + \sum_{i=2}^{p_{..}}(-n_{..\,i})\left(\frac{1}{(-n_{..\,i1})}P'_{..\,i1} + \sum_{j=2}^{s_{..\,i}}\frac{1}{(n-n_{..\,ij})}P_{..\,ij}\right) =\\
&= (n-n_{..\,1})\sum_{j=1}^{s_{..\,1}}\frac{1}{(n-n_{..\,1j})}P_{..\,1j} +\\
&\quad + n_{..\,1}\left(\frac{1}{(-n_{..\,11})}P'_{..\,11} + \sum_{j=2}^{s_{..\,1}}\frac{1}{(n-n_{..\,1j})}P_{..\,1j}\right) + P'_{..} =\\
&= n\sum_{j=1}^{s_{..\,1}}\frac{P_{..\,1j}}{(n-n_{..\,1j})} + n_{..\,1}\left(\frac{P'_{..\,11}}{(-n_{..\,11})} - \frac{P_{..\,11}}{(n-n_{..\,11})}\right) + P'_{..}\,.
\end{aligned} \tag{4.7}$$

Let us prove the following identities by induction on $t = d, \dots, 0$

$$
\begin{aligned}
P'_{..} &= \frac{T_{0+..-}}{n}\left(\frac{1}{w_{..}}\frac{T_{..}}{T'_{..}} - n_{..}\right),\\
P_{..} &= T_{0+..-} + P'_{..} = \frac{T_{0+..-}}{n}\left(\frac{1}{w_{..}}\frac{T_{..}}{T'_{..}} + n - n_{..}\right),
\end{aligned}
\tag{4.8}
$$

where $.. = i_1 j_1 \dots i_t j_t$ and

$$
w_{i_1 j_1 \dots i_t j_t} = \frac{\phi(|G_{i_1}|)}{\phi(|G_{i_1 j_1}|)} \cdots \frac{\phi(|G_{i_1 j_1 \dots i_t}|)}{\phi(|G_{i_1 j_1 \dots i_t j_t}|)}.
$$

For $t = d$ we have $T_{..} = w_{..}$, $T'_{..} = 1$ and $n_{..} = 1$. Hence $P'_{..} = 0$ and $P_{..} = T_{0+..-}$, which coincide with the initial conditions.

For $t = d-1, \dots, 0$ and $.. = i_1 j_1 \dots i_t j_t$ by the inductive hypothesis

$$
\begin{aligned}
P'_{..} = &\sum_{i=1}^{p_{..}} (-n_{..\,i}) \left[\frac{1}{(-n_{..\,i1})}\frac{T_{0+..\,i1-}}{n}\left(\frac{1}{w_{..\,i1}}\frac{T_{..\,i1}}{T'_{..\,i1}} - n_{..\,i1}\right) + \right.\\
&\left. + \sum_{j=2}^{s_{..\,i}} \frac{1}{(n - n_{..\,ij})}\frac{T_{0+..\,ij-}}{n}\left(\frac{1}{w_{..\,ij}}\frac{T_{..\,ij}}{T'_{..\,ij}} + n - n_{..\,ij}\right)\right] =\\
= &\frac{T_{0+..-}}{n}\sum_{i=1}^{p_{..}} (-n_{..\,i})\frac{T_{..\,i}}{T'_{..\,i}}\left[\frac{1}{(-n_{..\,i1})}\frac{T'_{..\,i1}}{T_{..\,i1}}\left(\frac{1}{w_{..}}\frac{n_{..\,i1}(n - n_{..\,i1})}{n_{..\,i}(n - n_{..\,i})}\frac{T_{..\,i1}}{T'_{..\,i1}} - n_{..\,i1}\right) - \right.\\
&- \frac{1}{(n - n_{..\,i1})}\frac{T'_{..\,i1}}{T_{..\,i1}}\left(\frac{1}{w_{..}}\frac{n_{..\,i1}(n - n_{..\,i1})}{n_{..\,i}(n - n_{..\,i})}\frac{T_{..\,i1}}{T'_{..\,i1}} + n - n_{..\,i1}\right) +\\
&\left. + \sum_{j=1}^{s_{..\,i}} \frac{1}{(n - n_{..\,ij})}\frac{T'_{..\,ij}}{T_{..\,ij}}\left(\frac{1}{w_{..}}\frac{n_{..\,ij}(n - n_{..\,ij})}{n_{..\,i}(n - n_{..\,i})}\frac{T_{..\,ij}}{T'_{..\,ij}} + n - n_{..\,ij}\right)\right] =\\
= &\frac{T_{0+..-}}{n}\sum_{i=1}^{p_{..}} (-n_{..\,i})\frac{T_{..\,i}}{T'_{..\,i}}\left[\frac{(-1)}{w_{..}}\frac{n}{n_{..\,i}(n - n_{..\,i})} + \sum_{j=1}^{s_{..\,i}}\left(\frac{1}{w_{..}}\frac{n_{..\,ij}}{n_{..\,i}(n - n_{..\,i})} + \frac{T'_{..\,ij}}{T_{..\,ij}}\right)\right] =\\
= &\frac{T_{0+..-}}{n}\sum_{i=1}^{p_{..}}\left(\frac{T_{..\,i}}{T'_{..\,i}}\frac{1}{w_{..}}\frac{n}{(n - n_{..\,i})} - \frac{T_{..\,i}}{T'_{..\,i}}\frac{1}{w_{..}}\frac{n_{..\,i}}{(n - n_{..\,i})} - n_{..\,i}\right) =\\
= &\frac{T_{0+..-}}{n}\left(\frac{1}{w_{..}}\frac{T_{..}}{T'_{..}} - n_{..}\right).
\end{aligned}
$$

Similarly, for $P_{..}$ we have

$$
\begin{aligned}
P_{..} = &\, n\sum_{j=1}^{s_{..\,1}} \frac{1}{(n - n_{..\,1j})}\frac{T_{0+..\,1j-}}{n}\left(\frac{1}{w_{..\,1j}}\frac{T_{..\,1j}}{T'_{..\,1j}} + n - n_{..\,1j}\right) -\\
&- \frac{n_{..\,1}}{n_{..\,11}}\frac{T_{0+..\,11-}}{n}\left(\frac{1}{w_{..\,11}}\frac{T_{..\,11}}{T'_{..\,11}} - n_{..\,11}\right) -\\
&- \frac{n_{..\,1}}{n - n_{..\,11}}\frac{T_{0+..\,11-}}{n}\left(\frac{1}{w_{..\,11}}\frac{T_{..\,11}}{T'_{..\,11}} + n - n_{..\,11}\right) + P'_{..} =
\end{aligned}
$$

$$
\begin{aligned}
&= T_{0+..-}\frac{T_{..1}}{T'_{..1}}\sum_{j=1}^{s_{..1}}\frac{1}{(n-n_{..1j})}\frac{T'_{..1j}}{T_{..1j}}\left(\frac{1}{w_{..}}\frac{n_{..1j}(n-n_{..1j})}{n_{..1}(n-n_{..1})}\frac{T_{..1j}}{T'_{..1j}}+n-n_{..1j}\right)-\\
&\quad - n_{..1}\frac{T_{0+..-}}{n}\frac{T_{..1}}{T'_{..1}}\left[\frac{1}{n_{..11}}\frac{T'_{..11}}{T_{..11}}\left(\frac{1}{w_{..}}\frac{n_{..11}(n-n_{..11})}{n_{..1}(n-n_{..1})}\frac{T_{..11}}{T'_{..11}}-n_{..11}\right)+\right.\\
&\quad \left.+\frac{1}{n-n_{..11}}\frac{T'_{..11}}{T_{..11}}\left(\frac{1}{w_{..}}\frac{n_{..11}(n-n_{..11})}{n_{..1}(n-n_{..1})}\frac{T_{..11}}{T'_{..11}}+n-n_{..11}\right)\right]+P'_{..}=\\
&= T_{0+..-}\frac{T_{..1}}{T'_{..1}}\left(\frac{1}{w_{..}}\frac{1}{n_{..1}(n-n_{..1})}n_{..1}+\frac{T'_{..1}}{T_{..1}}\right)-\\
&\quad -\frac{T_{0+..-}}{n}n_{..1}\frac{T_{..1}}{T'_{..1}}\left[\frac{1}{w_{..}}\frac{(n-n_{..11})}{n_{..1}(n-n_{..1})}-\frac{T'_{..11}}{T_{..11}}+\right.\\
&\quad \left.+\frac{1}{w_{..}}\frac{n_{..11}}{n_{..1}(n-n_{..1})}+\frac{T'_{..11}}{T_{..11}}\right]+P'_{..}=\\
&= T_{0+..-}\frac{T_{..1}}{T'_{..1}}\frac{1}{w_{..}}\frac{1}{n-n_{..1}}+T_{0+..-}-\frac{T_{0+..-}}{n}\frac{T_{..1}}{T'_{..1}}\frac{1}{w_{..}}\frac{n}{n-n_{..1}}+P'_{..}=\\
&= T_{0+..-}+P'_{..}\,.
\end{aligned}
$$

The identities 4.8 have been proved.

Applying the first of them (with $t=0$) to 4.5 we obtain

$$(TY_\tau y_\tau)(e_0)=T_{0+}+\frac{T_{0+}}{n}\left(\frac{T_{0-}}{T'}-(n-1)\right)=\frac{T_{0-}}{n}+\frac{T_{0+}}{n}=\frac{T}{n},$$

which completes the proof of the theorem.

5. Final remarks and speculations

The column space of Y (as well as WB^T) is also known as the star-space of graph G, which can be described as the linear space of currents – antisymmetric functions on edges – satisfying both Kirchhoff's laws, and arising while routing arbitrary demands – zero-sum functions on vertices. Therefore, the column-space of $W^{\frac{1}{2}}B^T$ we considered can be characterized as a star-space written in the scaled basis.

The results presented here may be also of interest in probability theory, because of the special role of the transfer current matrix in the uniform spanning tree model ([9]) and more specifically, because of the meaning of the eigenvalues of its submatrices in the given model ([2]).

The mention of series-parallel graphs/matroids in the context of the problem being considered seems to us very interesting but still leaves more questions than answers.

6. Acknowledgements

I would like to thank Igor Makhlin, Tomoyuki Shirai and the authors of the original paper [6] for fruitful discussions.

References

[1] W. Wu A. Kassel. Transfer current and pattern fields in spanning trees. *Probab. Theory Relat. Fields*, 163:89–121, 2015.

[2] R. B. Bapat. Mixed discriminants and spanning trees. *Sankhyā: The Indian Journal of Statistics, Series B (1960-2002)*, 54:49–55, 1992.

[3] R.J Duffin. Topology of series-parallel networks. *Journal of Mathematical Analysis and Applications*, 10:303–318, 1965.

[4] D. Eppstein. Parallel recognition of series-parallel graphs. *Information and Computation*, 98, 1:41–55, 1992.

[5] G.H. Golub and C.F. Van Loan. *Matrix Computations*, pages 84–85. Johns Hopkins U.P., 4th ed. edition, 2013.

[6] S.A. Goreinov, E.E. Tyrtyshnikov, and N.L. Zamarashkin. A theory of pseudo-skeleton approximations. *Linear Algebra Appl.*, 261:1–21, 1997.

[7] G. R. Grimmett. *Probability on Graphs, Random Processes on Graphs and Lattices*, pages 6–8. Cambridge University Press, 2nd ed. edition, 2018.

[8] OEIS. `https://oeis.org`.

[9] R. Pemantle R. Burton. Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances. *The Annals of Probability*, 21:1329–1371, 1993.

[10] YY. Lee RH. Jan, LH. Hsu. The most vital edges with respect to the number of spanning trees in two-terminal series-parallel graphs. *BIT Numerical Mathematics*, 32:403–412, 1992.

[11] D. E. Speyer. `https://oeis.org/A115594`, 2006.

7. Appendix

Algorithm 1 Subspace optimization

```
procedure OPTIMIZE(A)
    d ← Distance(A)                         ▷ Initial target function value
    M ← M_init                              ▷ Some initial perturbation magnitude
    repeat
        A_new ← Perturbate(A, M)            ▷ A new subspace candidate
        d_new ← Distance(A_new)             ▷ Its target function value
        if d_new > d then                   ▷ Update the current best solution
            A ← A_new
            d ← d_new
        else                                ▷ Otherwise, update the magnitude
            M ← Decrease(M)
        end if
    until certain stopping criteria are met
    return A                                ▷ An optimized subspace
end procedure
```

Algorithm 2 Extremal subspaces calculation

Input: $n > k > 0$ and additional heuristic parameters $N \gg 1$, $\varepsilon \ll 1$
Output: a set of extremal subspaces ES

```
ES ← ∅
i ← N                                       ▷ Attempts allocated to update the set
while i > 0 do
    A ← InitialGuess()                      ▷ An initial subspace
    A ← Optimize(A)                         ▷ An optimized subspace
    d ← Distance(A)                         ▷ Its target function value
    if d > 1/sqrt(n) + ε then
        Print( "The hypothesis is solved in negative" )   ▷ (Never seen this)
        return −1
    else if d > 1/sqrt(n) − ε and cl(A) ∩ ES = ∅ then
        ES ← ES ∪ {A}                       ▷ Update the set
        i ← N                               ▷ Reset number of attempts
    else
        i ← i − 1                           ▷ A suboptimal subspace found
    end if
end while
return ES
```

Email address: yuri.r.nesterenko@gmail.com