Estimation from Relative Measurements in Sensor Networks: Asymptotic Error Bounds from Electrical Analogy

Prabir Barooah and João P. Hespanha

Center for Control, Dynamics and Computation, Univ. of California, Santa Barbara, CA 93106

{pbarooah,hespanha}@ece.ucsb.edu

Abstract

We consider the problem of estimating vector-valued variables from noisy "relative" measurements. The measurement model can be expressed in terms of a graph. The vertices, or nodes, of the graph correspond to the variables being estimated and the edges correspond to noisy measurements of the difference between the two variables associated with the corresponding nodes (i.e., their relative values). This type of measurement model appears in several sensor networks problems such as location estimation and time synchronization.

We take the value of one particular variable as a reference and consider the optimal linear estimate for the differences between the remaining variables and the reference. We establish asymptotic upper and lower bounds on the estimation error variance of a node's variable as a function of the Euclidean distance in a drawing of the graph between the node and the reference one. These bounds result in a classification of graphs : civilized and dense, based on how the variance grows asymptotically with distance: at a rate greater than or less than linearly, logarithmically, or bounded. These bounds, being true for the optimal estimate, serve as a benchmark against which algorithms devised for specific applications can be evaluated for large networks.

In deriving these results, we establish and exploit an analogy between the optimal estimator variance and the effective resistance in a generalized electrical network that is significant on its own.

CONTENTS

I. INTRODUCTION

We consider the problem of estimating a number of vector valued variables from a number of noisy "relative measurements", i.e., measurement of the difference between certain pairs of these variables. Consider n vector-valued variables $x_1, x_2, \ldots, x_n \in \mathbb{R}^k$, called node variables, one or more of which are known, and the rest are unknown. A number of noisy measurements of the difference $x_u - x_v$ are available for certain pairs of nodes (u, v). We can associate the variables with the nodes $\mathbf{V} = \{1, 2, \ldots, n\}$ of a directed graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ and the measurements with the edges $\mathbf{E}$ of it, consisting of ordered pairs (u, v) such that a noisy "relative" measurement between x_u and x_v is available:

$$\zeta_{uv} = x_u - x_v + \epsilon_{u,v}, \quad (u, v) \in \mathbf{E} \subset \mathbf{V} \times \mathbf{V}, \tag{1}$$

where the ϵ_{uv}'s are uncorrelated zero-mean noise vectors with known covariance matrices $P_{u,v} = \mathrm{E}[\epsilon_{u,v}\epsilon_{u,v}^T]$. $P_{u,v} = \mathrm{E}[\epsilon_{u,v}\epsilon'_{u,v}]$. Just with relative measurements, determining the x_u's is only possible up to an additive constant. To avoid this ambiguity, we assume that there is at least one *reference* variables $o \in \mathbf{V}$ whose node variable x_o is known. In general, there may be more than one node that knows their variables. The set of nodes that know their node variables are called reference nodes, and are denoted by $\mathbf{V}_r$ (where $\mathbf{V}_r \subset \mathbf{V}$). This particular problem is relevant to several sensor networks applications such as location estimation and time synchronization, as well as motion coordination and flocking, that are described in section I-A. Distributed algorithms to compute the optimal estimate using only local information were reported in [1, 2], where the *Optimal estimate* refers to the classical Best Linear Unbiased (BLU) Estimator, which achieves the minimum variance among all linear unbiased estimators.

However, one may wonder what are the *fundamental limitations in terms of accuracy* for estimation problems defined in truly large graphs. Reasons for concern arise from estimation problems such as the one associated with the simple graph shown in Figure 1. It is a chain of nodes with node 1 as the reference and with a single edge $(u, u + 1)$, $u \geq 1$ between
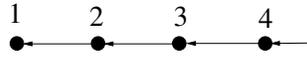

Fig. 1.   A graph where x_u's optimal estimate has a error variance that grows linearly with distance from the reference.

consecutive nodes u and $u + 1$. Without much difficulty, one can show that for such a graph the optimal estimate of x_u is given by

$$\hat{x}_u = \zeta_{u,u-1} + \cdots \zeta_{3,2} + \zeta_{2,1} + x_1,$$

and since each measurement introduces an additive error, the variance of the optimal estimation error $\hat{x}_u - x_u$ will increase linearly with u. This means that if u is very "far" from the reference node 1 then its estimate will necessarily be quite poor. Although the precise estimation error depends on the exact values of the variances of the measurements, for this graph the variance of the optimal estimate of x_u will grow essentially linearly with u.

In this paper we investigate how the structure of the graph $\mathbf{G}$ affects the "quality" of the optimal estimate $\hat{x}_u^*$ of x_u, measured in terms of the covariance of the estimation error $\Sigma_{u,o} := \mathrm{E}[(x_u - \hat{x}_u^*)(x_u - \hat{x}_u^*)^T]$. We focus on the case of a single reference node. Specifically, we raise and answer the following questions:

1) How does the error variance of a node variable's estimate scale with the node's distance from the reference node asymptotically, and
2) can one deduce these scaling laws from coarse structure of the measurement graph $\mathbf{G}$?

It seems reasonable that for any measurement graph the estimation error variance will increase with the distance to the reference nodes. We will see that the exact nature of the scaling of error variance with distance depends on intrinsic structural properties of the measurement graph, and that some graphs exhibit scaling laws far better than the linear scaling that we encountered in the graph shown in Figure 1. For a given maximum acceptable error, the number of nodes with acceptable estimation errors will be large if the graph exhibits a slow increase of variance with distance, but small otherwise. These scaling laws therefore determine how large a graph can be in practice. The answer to the questions raised here will also help us in designing and deploying very large networks for which accurate estimates are possible. In this paper we will describe a classification of graphs that determines how the variance grows with distance from the reference node. There are graphs where variance grows linearly with distance, as in the example of Figure 1. But there are also a large

class of graphs where it grows only logarithmically with distance. Most surprisingly, in certain graphs it can even stay below a constant value, no matter the distance.

The structural properties of a graph that determine the scaling of variance with distance are related to the "denseness" of the graph. However, the notion of denseness of a graph is not a easy one to define. In the sensor network literature, the average number of nodes in an unit area/volume of the deployed network or the average number of neighbors of a node are used as measures of denseness [3, 4] The former is not a good measure of density since it doesn't take into account the degree of interconnections between nodes. Even the latter, though it does take into account interconnections, fails to distinguish structural differences between graphs that have a large impact on the variance growth with distance. There are graphs with the same node degree but with very different variance growth rates, and there are graphs with different degrees that have the same growth rate of variance.In section VI-A we provide a few such examples to highlight this issue, that measures of "denseness" based on node degree or node/edge density etc. are somewhat naive and inadequate.

A suitable notion of graph "density/sparsity" can be developed based on the concept of graph drawing. The *drawing* of a graph is simply a function that maps its nodes to points in some Euclidean space $\mathbb{R}^d$, $d \geq 1$. We define a graph to be *dense in $\mathbb{R}^d$* if its nodes can be drawn in $\mathbb{R}^d$ such that there is a positive number γ so that (i) every ball in $\mathbb{R}^d$ with diameter γ contains at least one node of the graph and (ii) there is a nonzero minimum ratio of the Euclidean distance in the drawing to the graphical distance, between any two nodes of the graph. Intuitively, dense graphs in $\mathbb{R}^d$ have sufficiently many nodes to cover $\mathbb{R}^d$ without holes, and sufficiently many interconnecting edges so that two nodes with a small Euclidean distance in the drawing also has a small graphical distance. We define a graph to be sparse in $\mathbb{R}^d$ (for some d) if it can be drawn in $\mathbb{R}^d$ in a *civilized* manner. Graphs that can be drawn in a civilized manner appeared in [5] in connection with random walks, where it was defined as a graph that can be drawn in $\mathbb{R}^d$ such that (i) there is a minimum distance $s > 0$ between any two nodes and (ii) there is a maximum distance $r < \infty$ between nodes connected by an edge in the drawing. Intuitively, the nodes and edges of such a graph are sufficiently sparse to be drawn in $\mathbb{R}^d$ without too much clutter[1].

These two concepts of dense and sparse graphs allow one to characterize precisely how the variance of the BLU estimation error grows with distance from the reference. Table I summarizes the main results of this paper.

TABLE I

COVARIANCE MATRIX $\Sigma_{u,o}$ OF x_u'S OPTIMAL ESTIMATE FOR GRAPHS THAT ARE DENSE OR SPARSE IN $\mathbb{R}^d$.

Euclidean space	Covariance matrix $\Sigma_{u,o}$ of the estimation error of x_u in a *sparse graph*	Covariance matrix $\Sigma_{u,o}$ of the estimation error of x_u in a *dense graph*
$\mathbb{R}$	$\Sigma_{u,o} = \Omega\Big(d_f(u,o)\Big)$	$\Sigma_{u,o} = \mathcal{O}\Big(d_f(u,o)\Big)$
$\mathbb{R}^2$	$\Sigma_{u,o} = \Omega\Big(\log d_f(u,o)\Big)$	$\Sigma_{u,o} = \mathcal{O}\Big(\log d_f(u,o)\Big)$
$\mathbb{R}^3$	$\Sigma_{u,o} = \Omega\Big(1\Big)$	$\Sigma_{u,o} = \mathcal{O}\Big(1\Big)$

In the table, $d_f(u,o)$ denotes the Euclidean distance between node u and the reference node o, for any drawing f that establishes the graph's sparseness/denseness in the appropriate Euclidean space. For functions $\phi : \mathbb{R} \to \mathbb{R}^{k \times k}$ and $\psi : \mathbb{R} \to \mathbb{R}$, the notation $\phi(y) = \Omega(\psi(y))$ means that there exists a symmetric positive definite matrix P_L and a positive scalar y_L such that $\phi(y) \geq P_L \psi(y)$ for all $y > y_L$, and the notation $\phi(y) = \mathcal{O}(\psi(y))$ means that there exists a symmetric positive definite matrix P_U and a positive scalar y_U such that $\phi(y) \leq P_U \psi(y)$ for all $y > y_U$. For two matrices $A, B \in \mathbb{R}^{k \times k}$, $A \geq B$ implies $A - B$ is positive semidefinite.

A crucial step in proving the results in Table I, is to establish that the covariance matrix $\Sigma_{u,o}$ of the BLU estimation error is numerically equal to a effective resistance in a suitably defined electrical circuit. Such equivalence was first noted by Karp *et. al.* [8] for scalar measurements ($k = 1$). Here, we prove that this also holds for vector measurements, provided that one considers a *generalized electric circuits* in which currents, voltages and resistances are matrices. Tools for bounding the effective resistance in complicated graphs by that is simpler graphs are employed to establish the bounds on BLU estimator covariances.

A. Estimation from Relative Measurements in Sensor Networks

This estimation problem addressed in this paper has multiple applications in the area of sensor networks. Some of these are summarized below:

---

[1]We could have used the terminology "a graph is civilized in $\mathbb{R}^d$" instead of "sparse in $\mathbb{R}^d$", and in fact this is what we did initially in [6]. However, according to Peter Doyle "it is not the graph that is civilized but the manner in which can be drawn" [7]. Hence we decided to call a graph sparse if it can be drawn in a civilized manner.

1) Position estimation: A large number of sensors is deployed in a region of space. Sensors do not know their positions in a global coordinate system, but every sensor can measure its relative position with respect to a set of nearby sensors. These measurements could be obtained, for example, from range and bearing (angle) measurements (cf. Figure 2), provided nodes have on-board compasses so that their local coordinate systems have the same orientation. We are of course assuming that each node does not have, or does not want to rely upon, GPS. In this problem, $k = 2$ if the nodes are located in a plane and $k = 3$ if they are located in three-dimensional space.

PSfrag replacements

$$\zeta_{u,v} = \begin{bmatrix} r_{u,v} \cos \theta_{u.v} \\ r_{u,v} \sin \theta_{u,v} \end{bmatrix}$$
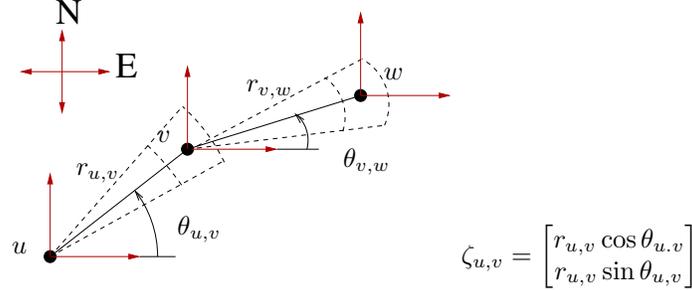
Fig. 2. Relative position measurement in a Cartesian reference frame from range and bearing measurement. A local compass at each sensor is needed to measure bearing with respect to a common "North."

Range can be estimated using RF or acoustics measurements (or a combination of the two) of RSS (Received Signal Strength) and/or ToA (Time of Arrival) [9–11]. When nodes do not have stringent space or power constraints, such as mobile robots or Unmanned Aerial Vehicles, they can be fitted with phased arrays of RF antennas or with laser range finders [12] which can provide measurement of bearing/AoA and range. However, To equip tiny sensors such as MICA [13] motes with direction sensing capability may be a challenge. A small form-factor node with AoA sensing capability has been realized in the Cricket system through the use of multiple ultrasound receivers [11]. Switched arrays of microstrip antennas, due to their small size, offer another attractive possibility in this case. Such antennas have been successfully mounted on MICA2 motes [14]. Kalis *et. al.* [15, 16] have also reported algorithms to provide coarse estimates of AoA using switched micostrip antennas. In [17], the authors describe a MAC protocol that can estimate AoA[2]. Thus, it is reasonable to believe that small sensor nodes with both range and direction measurement capability will be technologically feasible in the near future.

2) Time synchronization: A group of sensing nodes are part of a multi-hop communication network. Each node has a local clock but these local clocks are not precisely synchronized. However, nodes that can communicate directly can determine the offsets between their clocks. This is typically done by exchanging "hello" messages that are time-stamped with local clock times. In particular, two nodes u and v that can exchange messages, have access to the measurement

$$\zeta_{u,v} = t_u - t_v + \epsilon_{u,v} \in \mathbb{R},$$

where t_u and t_v are the local times of nodes u ad v, respectively and $\epsilon_{u,v}$ denotes measurement error that arises from various inaccuracies in the estimation of time offsets using "hello" messages [8]. The task is now to determine the clock times with respect to some reference, which could be the local time at a particular node or some arbitrary global time.

Motion consensus: Several mobile agents move in space. Every agent would like to determine its velocity with respect to the velocity of a "leader", which is either an agent or a global reference, but the agents can only measure their relative velocities with respect to nearby agents. These measurements could be obtained, for example, using vision-based sensors. In particular, two nearby agents u and v moving with velocities $\dot{p}_u$ and $\dot{p}_v$, respectively have access to the measurement

$$\zeta_{u,v} = \dot{p}_u - \dot{p}_v + \epsilon_{u,v} \in \mathbb{R}^3,$$

where $\epsilon_{u,v}$ denotes measurement error. The task is to determine the velocity of each agent with respect to the leader based solely on the available relative velocities between pairs of neighboring agents.

Among the applications cited above, position estimation has probably received the most attention in recent times. In sensor network literature, this problem is known as *localization*. Several localization algorithms have been designed assuming only relative range information [10, 18–20], and to a lesser degree assuming only angle measurement [21]. In recent times combining both range and bearing information has received some attention [11, 22–24]. It was reported in [22] that location estimation using both relative distance and angle measurements can substantially improve the accuracy of estimates. Still, to the best of our knowledge, no one has looked at the localization problem in terms of the noisy measurement model (1). The advantage of this formulation is that the effect of measurement noise can be explicitly accounted for and filtered out to the maximum extent possible by employing the classical Best Linear Unbiased

---

[2]more careful review needed

Estimator(BLUE). This estimator produces the minimum variance estimate, and hence is the most accurate on average. Location estimation techniques using only range measurement are usually not designed to handle measurement noises, which may introduce significant errors into the location estimate due to flip ambiguities [10]. A localization algorithm designed to work with noisy range measurement was proposed in [10]. However, the algorithm did not guarantee localization of all the nodes. The advantage of posing the localization problem as an estimation problem in Cartesian coordinates using the measurement model (1) is that the *optimal* (minimum variance) estimates all node positions in a connected network can be unambiguously determined when only one node that knows its position. A large number of well placed beacon nodes that know their position and broadcast that to the network – a usual requirement for many localization schemes – are not required.

B. Related Work

A very limited number of papers have attempted to look into the effect of network properties on the estimation error for the localization problem, for different localization algorithms. These papers typically investigate the problem by looking at the Cramer-Rao lower bound [3, 4, 18, 25]. However, even with the assumption of Gaussian measurements, the C-R bound takes a form so complex that no useful information could be extracted from it by analysis. Hence, in [4], the effect of graph properties, such as node density and network size, on the estimation accuracy was studied by numerically evaluating the Cramer-Rao bound for different values of these parameters. However, conclusions drawn from such studies suffer from the limitation that they are applicable only so long as one considers the same set of parameters that the study was conducted for. In fact, such studies frequently report that higher node degree improves estimation accuracy, whereas node degree is a very poor predictor of how estimation error scales with distance, as we will show in section VI-A. To the best of our knowledge, no one has looked at the localization problem in terms of the noisy measurement model (1), or investigated the question of how the estimation error of the optimal estimate scales with distance.

C. Statement of Contribution

We have systematically examined the question of how the optimal estimate's variance scales with distance from the reference, and how this scaling is affected by structural properties of the graph. We have developed a graph classification – dense and sparse – based on coarse structural properties of graphs that determine at what rate variance grows with distance, whether linear, logarithmic or bounded. The classification into dense and sparse graphs are based on how a graph looks when *drawn* on a suitable Euclidean space. Our results also help us construct examples of graphs that show the inadequacy of conventional measures of denseness such as node degree and number of nodes/edges per unit area.

For most problems that arise in the context of sensor networks there are "natural drawings" that are physically meaningful and for which one would expect the graphs to be sparse and/or dense in appropriate Euclidean spaces. Indeed, in all the examples discussed above, the variables x_i refer to properties of objects, so the measurement graph G could be "drawn" by associating to each node the Cartesian position of the corresponding object (either in $\mathbb{R}$, $\mathbb{R}^2$, or $\mathbb{R}^3$ depending on the problem).

Such a natural drawing would be *civilized* as long as (i) there is a minimum distance between every two objects and (ii) the sensors used to obtain the relative measurements ζ_{ij} have limited range. These are clearly true in any realistic network. In this case, the formulas in the second column of Table I provide explicit asymptotic lower-bounds on the variance of the estimation error for x_u as a function of the distance between the corresponding object and the reference one.

If the natural drawing considered above is *dense*, the formulas in the third column of Table I provide upper-bounds on the variances of the estimation errors. Formally, dense graphs must be infinite. However, in practice all that matters is that both the reference node and the nodes whose variable one wants to estimate fall spatially within the remaining objects.

Since these bounds are for the optimal estimate, they serve as a benchmark for any linear unbiased estimation algorithm designed for specific applications. Thus, a candidate estimation algorithm's performance can be compared against these bounds to see if it achieves the limiting performance possible.

D. Organization

The remainder of the paper is organized as follows: In section II, a few definitions and concepts from Graph theory that are relevant to our problem are presented. The reader may skip this section, since every term is explained in the main body of the paper the first time it appears. Section III provides the complete description of the problem of estimation from relative measurement and a review of the BLU estimator. In section IV we describe the analogy between the BLU estimator covariance matrix and the effective resistance in a suitably defined generalized electrical network. We also describe tools for bounding effective resistances in one graph by that in another. Section VI starts with a discussion of graph drawings and formally introduces the concepts of dense and sparse graphs. The main results of the paper are stated and proved in Section VI-D. Section VII contains a discussion of the results and directions for future work.

Fig. 3.   A directed graph.

II. Graph Theory Preliminaries and Terminology

This section need not be read before starting with the rest of the paper. We do define every non-trivial term in the main body of the paper the first time it appears. However, if a reader wishes to skip certain portions of the paper and then encounters a term in the subsequent sections that was defined in the parts he/she skipped, instead of searching for the term all over the paper, he/she can simply look for it in this section. That is the sole purpose of including this section - to serve an index of terminology used. Furthermore, the reader familiar with basic graph theory terminology will know most of the terms anyway.

A *graph* $\mathbf{G}$ consists of a set of vertices, or nodes, denoted by $\mathbf{V}$, and a set of edges $\mathbf{E}$, where each edge $e \in \mathbf{E}$ is a pair of nodes: $e = (u, v)$, $u, v \in \mathbf{V}$. An edge $(u, v)^3$ is said to be incident on the nodes u and v. If u and v have an edge between them, they are called *adjacent*. A graph is called a *directed graph* if every edge (u, v) is an ordered pair of nodes it is incident on. In this case, an edge (u, v) is thought of as having a direction, from u to v. If every edge is simply a set of two nodes, and edges do not have any directions associated with them, then the graph is said to be *undirected*. We assume that no edge is incident on a single vertex only, therefore there are no self-loops in the graph.

In this paper we will use the notation

$$e \sim u \; (\sim u, v)$$

to mean that the edge e is incident on node u (nodes u and v).

An *undirected path* $\mathcal{P}$ in an directed graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ from $u_1 \in \mathbf{V}$ to $u_\ell \in \mathbf{V}$ is an ordered sequence of alternating nodes and edges of the form

$$\mathcal{P} = \{u_1, e_1, u_2, e_2, \ldots, u_{\ell-1}, e_{\ell-1}, u_\ell\} \tag{2}$$

with $u_j \in \mathbf{V}$ and $e_j \in \mathbf{E}$ for $1 \le j \le \ell$ that satisfies the following properties:

1) each edge in $\mathcal{P}$ is incident on the vertices that are immediately before and after it in $\mathcal{P}$. That is $e_j = (u_j, u_{j+1})$ or (u_{j+1}, u_j).
2) $\mathcal{P}$ is open: $u_1 \neq u_\ell$,
3) No vertex in $\mathcal{P}$ is repeated, i.e., $u_j \neq u_i$ for $i \neq j$.

Figure 3 shows a directed graph with 4 nodes and 4 edges. $\mathcal{P}_{p,r} = \{p, e_1, q, e_2, r\}$ is a path from p to r. Note that a path in our definition is not a directed path as defined in traditional graph theory literature.

A directed graph is said to be *weakly connected* if there is an undirected path from any node to any other node.

For a path $\mathcal{P}$ from u_1 to u_ℓ of the form (2), we say that the *orientation* of an edge $e \in \mathcal{P}$ coincides with the *orientation* of the path if $e = (u_j, u_{j+1})$ and u_j, e, u_{j+1} appear in that order in $\mathcal{P}$. The orientations are opposite if $e = (u_{j+1}, u_j)$ and u_j, e, u_{j+1} appear in that order in $\mathcal{P}$. In the example shown in figure 3, the orientation of the path $\mathcal{P}_{p,r}$ is the same as the orientation of the edge e_1 but opposite to that of the edge e_2.

A *circuit* $\mathcal{C}$ is a sequence of edges of the form (2), with $u_1 = u_\ell$, but satisfies all the other properties of a path. In figure 3, $\mathcal{C}_p := \{p, e_1, q, e_2, r, e_3, w, e_4, p\}$ is a circuit. Orientation of a circuit is defined in the same way as that for a path. For example, the orientation of $\mathcal{C}_p$ in figure 3 is the same as that of e_1 but opposite to that of e_3.

The graphical distance $d_{\mathbf{G}}(u, v)$ between two nodes u and v in an undirected graph $\mathbf{G}$ is the length of the shortest undirected path from u to v, where length of a path is measured by the number of edges in the path.

For any integer h, the h-*fuzz* $\mathbf{G}^{(h)}$ of a graph $\mathbf{G}$ is the graph obtained from $\mathbf{G}$ by adding an edge (u, v) whenever the graphical distance between the nodes u and v lies between 2 and h (inclusive) [5]. When $\mathbf{G}$ is directed, the directions of the new edges are arbitrary.

We say that a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ can be *embedded* in another graph $\bar{\mathbf{G}} = (\bar{\mathbf{V}}, \bar{\mathbf{E}})$, if there is an injective map $\eta : \mathbf{V} \to \bar{\mathbf{V}}$ such that $(\eta(u), \eta(v)) \in \bar{\mathbf{E}}$ for every $(u, v) \in \mathbf{E}$. The map η is called the *embedding*.

A *d-D lattice in* $\mathbb{R}^d$ is a graph that has a vertex at every point in $\mathbb{R}^d$ with integer coordinates and an edge between every two nodes with an Euclidean distance of 1 between them. We denote the d-D lattice in $\mathbb{R}_d$ by $\mathbf{Z}_d$.

For a directed graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ with n nodes and m edges, the *Incidence Matrix* $A \in \mathbb{R}^{n \times m}$ is a matrix of 1, -1 and 0's defined as $A = [a_{ij}]$, where $a_{ij} = 1$, -1 or 0, if edge e_j is incident on node i and directed away from it, is incident on node i and directed toward it, or is not incident on node i, respectively. In this paper we consider only those graphs that are weakly connected. Hence, the rank of A is $n - 1$[26]. A sub-matrix of A that has full row rank is called a *basis*

---

^{3}The notation (u, v) to denote an edge incident on the nodes u and v lacks precision since there may be multiple edges that are incident on u and v. However, it has the advantage of clarifying which two nodes the edge is incident on. Therefore we stick to this notation in the paper.

incidence matrix A_b. In principle, we can remove any one row of A to get A_b [26], but in this paper we will remove the row corresponding to the reference node when constructing A for a measurement graph. Later we will also consider electrical networks. The basis incidence matrix of an electrical network will be obtained by removing from A the row corresponding to the node where current is extracted. From now on, we will call A_b *the* basis incidence matrix with the understanding that A_b is obtained from A following the convention outlined above. For the measurement graph with node variables $x_i \in \mathbb{R}^k$, we define a *Generalized Incidence Matrix* $\mathcal{A}$ and the *Generalized Basis Incidence Matrix* $\mathcal{A}$ as

$$\mathcal{A} = A \otimes I_k, \tag{3}$$

$$\mathcal{A}_b = A_b \otimes I_k. \tag{4}$$

An "expanded" form the graph Laplacian will play a key role in our discussions. Recall that the *Laplacian L* of a directed graph is defined as:

$$L = AA^T. \tag{5}$$

It should be emphasized that the Laplacian defined above is invariant to the directions assigned to the edges [27]. Therefore the Laplacian of a directed graph, as defined above, can also be thought of as the Laplacian associated with the corresponding undirected graph. There are definitions of Laplacian for directed graphs [28, 29] that are not invariant to the directions of the edges. However, our purpose is best served by the definition (5).

Consider a function $W : \mathbf{E} \to \mathbb{R}^{k \times k}$ that assigns a symmetric, positive definite matrix weight $W(e) > 0$ for every $e \in \mathbf{E}$. For the graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ with $|\mathbf{E}| = m$ and the weight function W, we can further define the weighing matrix $\mathcal{W}$ as $\mathcal{W} := \mathrm{diag}(W(e_1), W(e_2), \dots, W(e_m))$. The matrix $\mathcal{W} \in \mathbb{R}^{km \times km}$ is block diagonal and positive definite. We define the *Generalized Weighted Laplacian Matrix* for a graph $\mathbf{G}$ with a generalized incidence matrix $\mathcal{A}$ as

$$\mathscr{L}_{(W)} = \mathcal{A}\mathcal{W}\mathcal{A}^T. \tag{6}$$

The *Generalized Weighted Grounded Laplacian* is defined as

$$\mathcal{L}_{(W)} = \mathcal{A}_b \mathcal{W} \mathcal{A}_b^T. \tag{7}$$

As before, the weighted generalized Laplacian is also invariant to the directions assigned to the edges.

III. Optimal Estimation from Relative Measurements

By stacking together all the measurements into a single vector $\mathbf{z}$, all node variables into one vector $\mathbf{X}$, and all the measurement errors into a vector $\boldsymbol{\epsilon}$, we can express all the measurement equations (1) in the following form:

$$\mathbf{z} = \mathcal{A}^T \mathbf{X} + \boldsymbol{\epsilon}, \tag{8}$$

where

$$\mathcal{A} := A \otimes I_k \quad \in \mathbb{R}^{kn \times km}, \tag{9}$$

A is the *incidence matrix of* $\mathbf{G}$, I_k is a $k \times k$ identity matrix and "$\otimes$" denotes the Kronecker product. The Kronecker product $A \otimes I_k$ is a much bigger matrix than A obtained by replacing every element of $A = [a_{u,e}]$ by the matrix $a_{u,e}I_k$. The incidence matrix A is an $n \times m$ matrix with one row per node and one column per edge defined by $A := [a_{ue}]$, where a_{ue} is nonzero if and only if the edge $e \in \mathbf{E}$ is incident on the node $u \in \mathbf{V}$ and when nonzero, $a_{ue} = -1$ if the edge e is directed towards u and $a_{ue} = 1$ otherwise. The matrix $\mathcal{A}$ that appears in (8) is an "expanded" version of the incidence matrix A that we call the *generalized incidence matrix of* $\mathbf{G}$. By partitioning $\mathbf{X}$ into a vector $\mathbf{x}$ containing all the unknown node variables and another vector $\mathbf{x}_r$ containing all the known reference node variables, we can re-write (8) as

$$\mathbf{z} = \mathcal{A}_r^T \mathbf{x}_r + \mathcal{A}_b^T \mathbf{x} + \boldsymbol{\epsilon},$$

$$or, \tag{10}$$

$$\mathbf{z} - \mathcal{A}_r^T \mathbf{x}_r = \mathcal{A}_b^T \mathbf{x} + \boldsymbol{\epsilon}, \tag{11}$$

where $\mathcal{A}_r$ contains the rows of $\mathcal{A}$ corresponding to the reference nodes and $\mathcal{A}_b$ contains the rows of $\mathcal{A}$ corresponding to the unknown node variables. See Figure 4 for an example of a measurement graph and the associated matrices described above.

We seek an estimate of the vector $\mathbf{x}$ of the unknown node variables as a linear combination of the available measurements $\mathbf{z}$ that yields an unbiased estimate and achieves minimum estimator error variance. The solution to this problem is the classical Best Linear Unbiased Estimator(BLUE) [30]. For the measurement model (8), where $\mathscr{P} = \mathrm{E}[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T]$ is the

covariance matrix of the measurement error vector, the BLUE estimate of $\mathbf{x}$ is given by the solution to the following equation

$$\mathcal{L}\hat{\mathbf{x}}^* := \mathbf{b} \tag{12}$$

where

$$\mathcal{L} := \mathcal{A}_b \mathscr{P}^{-1} \mathcal{A}_b^T, \tag{13}$$
$$\mathbf{b} := \mathcal{A}_b \mathscr{P}^{-1}(\mathbf{z} - \mathcal{A}_r^T \mathbf{x}_r).$$

Among all linear estimators of $\mathbf{x}$, the BLU estimate has the smallest variance of the estimation errors $\mathbf{x} - \hat{\mathbf{x}}$. Figure 4 shows a simple measurement graph along with the corresponding measurement equations (8), (11) and estimate (12) defined above.

The covariance matrix of the error in the measurement along an edge $e \in \mathbf{E}$ can be associated with the edge e. Let $P(e) := \mathrm{E}[\epsilon_e \epsilon_e^T]$. Thus, we can associate with the measurement graph $\mathbf{G}$ an *edge-covariance function* $P : \mathbf{E} \to \mathbb{R}^{k \times k}$ such that $P_e = P_e^T > 0$ (i.e., positive definite) for all $e \in \mathbf{E}$. In this paper we assume that the measurement errors on two distinct edges are uncorrelated. Under this assumption, the covariance matrix $\mathscr{P}$ of the measurement error vector $\boldsymbol{\epsilon}$ is a positive definite block diagonal matrix which has the edge error covariances along the diagonal:

$$\mathscr{P} = \mathrm{diag}\left(P_1, P_2, \ldots, P_m\right) \in \mathbb{R}^{km \times km} > 0.$$

A measurement Graph $\mathbf{G}$ and its incidence matrix A (row and column indices of A correspond to node and edge indices, respectively).

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & 0 & 1 \end{bmatrix}$$

Matrix form (8) of the measurement equations (1):

$$\underbrace{\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \end{bmatrix}}_{\mathbf{z}} = \underbrace{\begin{bmatrix} I & -I & 0 & 0 \\ I & -I & 0 & 0 \\ 0 & I & 0 & -I \\ 0 & I & -I & 0 \\ 0 & 0 & -I & I \end{bmatrix}}_{\mathcal{A}^T} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{X}} + \underbrace{\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix}}_{\boldsymbol{\epsilon}}$$

Measurement model (11) when node 1 is the reference with $x_1 = 0$:

$$\underbrace{\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \end{bmatrix}}_{\mathbf{z}} = \underbrace{\begin{bmatrix} I \\ I \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\mathcal{A}_r^T} \underbrace{x_1}_{\mathbf{x}_r} + \underbrace{\begin{bmatrix} -I & 0 & 0 \\ -I & 0 & 0 \\ I & 0 & -I \\ I & -I & 0 \\ 0 & -I & I \end{bmatrix}}_{\mathcal{A}_b^T} \underbrace{\begin{bmatrix} x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{X}} + \boldsymbol{\epsilon}$$

Optimal estimates (12) when all measurement covariance matrices are equal to the identity matrix:

$$\begin{bmatrix} \hat{x}_2^* \\ \hat{x}_3^* \\ \hat{x}_4^* \end{bmatrix} = \underbrace{\begin{bmatrix} 4I & -I & -I \\ -I & 2I & -I \\ -I & -I & 2I \end{bmatrix}^{-1}}_{\mathcal{L}} \underbrace{\begin{bmatrix} -I & -I & 0 & 0 & 0 \\ 0 & 0 & 0 & -I & -I \\ 0 & 0 & -I & 0 & I \end{bmatrix}}_{\mathcal{A}_b \mathscr{P}^{-1}} \underbrace{\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \end{bmatrix}}_{\mathbf{z} - \mathcal{A}_r^T \mathbf{x}_r}$$

Covariance matrices of the overall estimation error and of the individual node-variable errors:

$$\boldsymbol{\Sigma} = \underbrace{\frac{1}{6}\begin{bmatrix} 3I & 3I & 3I \\ 3I & 7I & 5I \\ 3I & 5I & 7I \end{bmatrix}}_{\mathcal{L}^{-1}}, \qquad \Sigma_2 = \frac{1}{2}I, \qquad \Sigma_3 = \frac{7}{6}I, \qquad \Sigma_4 = \frac{7}{6}I.$$
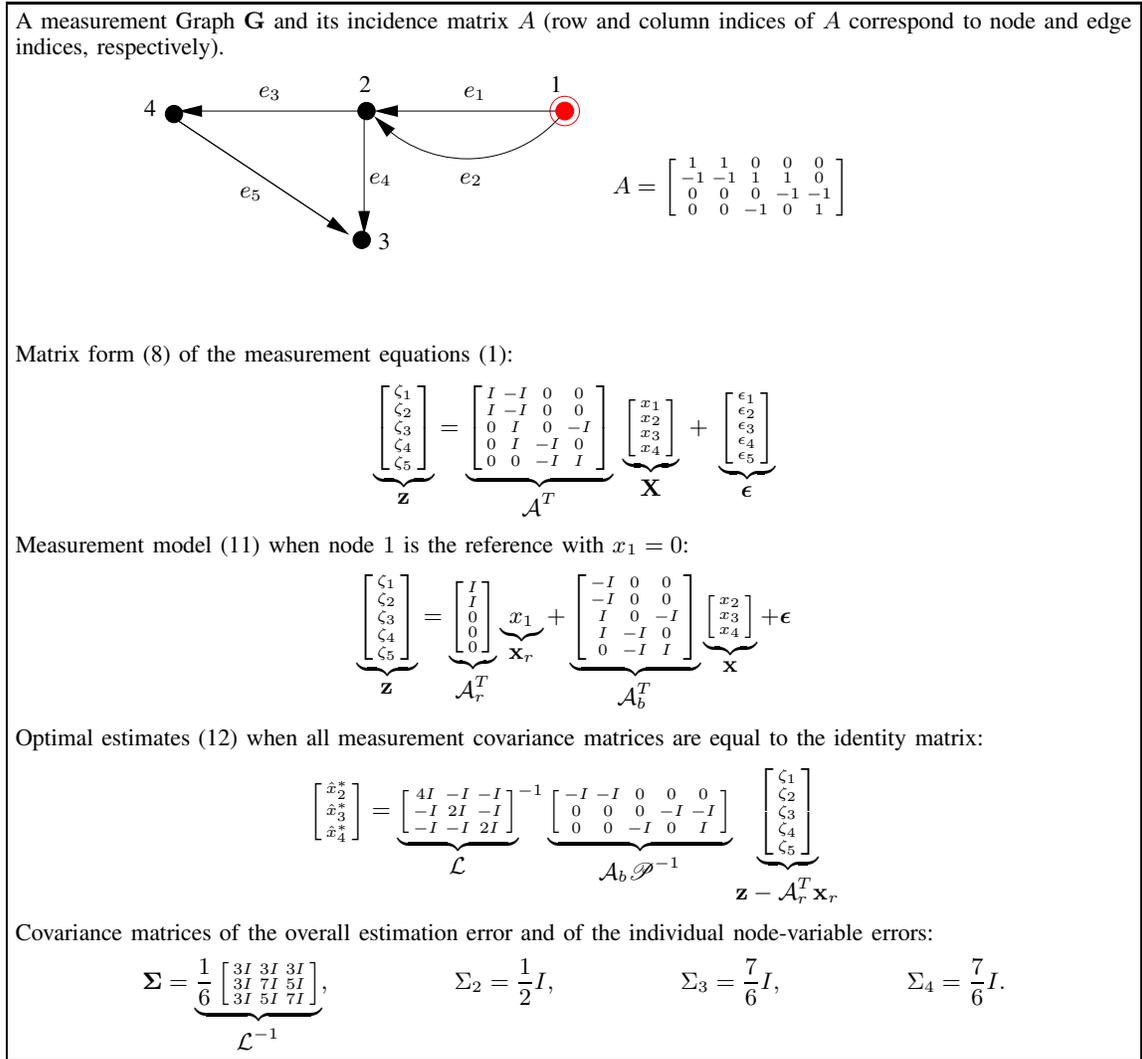
Fig. 4. A measurement graph $\mathbf{G}$ with 4 nodes and 5 edges. Node 1 is the reference. The associated optimal estimates etc. are also shown.

The optimal estimate, and its covariance, depends on the graph (though $\mathcal{A}$) and on the measurement error covariances (through $\mathscr{P}$). Clearly estimating node variables is possible only up to a additive constant if there are no reference nodes

in the graph. The next result (Theorem 1) establishes necessary conditions on the graph $\mathbf{G}$ so that the optimal estimate of node variables is unique and shows how the covariance of the estimation error $\mathbf{x} - \hat{\mathbf{x}}$ relates to the matrices associated with the graph $\mathbf{G}$.

Before proceeding, we need to introduce some terminology. A directed graph is called *weakly connected* if it is possible to go from any node of the graph to any other by traversing the edges, not necessarily respecting the edge directions. If a graph is not weakly connected, it consists of more than one weakly connected components. A *weakly connected component* of a directed graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ is a *maximally weakly connected subgraph* of $\mathbf{G}$. That is, if $\mathbf{G}_s = (\mathbf{V}_s, \mathbf{E}_s)$ is a subgraph of $\mathbf{G}$ (i.e., $\mathbf{V}_s \subset \mathbf{V}, \mathbf{E}_s \subset \mathbf{E}$) such that $\mathbf{G}_s$ is weakly connected and any "larger" subgraph $\bar{\mathbf{G}}_s = (\bar{\mathbf{V}}_s, \bar{\mathbf{E}}_s)$ such that $\bar{\mathbf{V}}_s \supset \mathbf{V}_s$ and $\bar{\mathbf{E}}_s \supset \mathbf{E}_s$ is not weakly connected, then $\mathbf{G}_s$ is a connected component of $\mathbf{G}$. If a graph consists of only one node, it is connected.

Theorem 1. *The matrix $\mathcal{L}$ defined in* (13) *is invertible if and only if every weakly connected component of the graph has at least one reference node. When $\mathcal{L}$ is non-singular, the estimation error covariance matrix $\mathbf{\Sigma} := \mathrm{E}[(\mathbf{x} - \hat{\mathbf{x}}^*)(\mathbf{x} - \hat{\mathbf{x}}^*)^T]$ is given by*

$$\mathbf{\Sigma} = \mathcal{L}^{-1}. \quad \square$$

Proof: [Proof of Theorem 1] We will first consider the case when $\mathbf{G}$ has only one connected component and prove that $\mathcal{L}$ is invertible if and only if the graph has at least one reference node. When $\mathbf{G}$ is weakly connected, the rank of its incidence matrix $\mathrm{rank}(A) = n - 1$, where n is the number of nodes [26]. If $\mathbf{G}$ has no reference nodes, then $A = A_b$, which makes A_b, and thereby $\mathcal{A}_b$, rank deficient. Then $\mathcal{L} = \mathcal{A}_b \mathscr{P}^{-1} \mathcal{A}_b$ is singular. On the other hand, any sub matrix obtained from A by removing one or more rows has full row rank. To prove it, take a submatrix $\bar{A}$ by removing the row corresponding to an arbitrary node u. Since u has at least one edge that connects it to some other node in the graph, say v, the coefficient α_v in the linear combination of the rows of $\bar{A}$ must be zero for the linear combination to be zero. But if α_v is 0, then α_w must also be zero where w is a node that shares an edge with v, and so on, which shows that all the coefficients must be 0 for the linear combination of the rows of $\bar{A}$ to be zero. Hence $\bar{A}$ is full row rank. Any smaller submatrix must obviously be full row rank. Now if the weakly connected graph $\mathbf{G}$ has at least one reference node, A_b, is full row rank by the previous argument, and so is $\mathcal{A}_b$. To prove that $\mathcal{L}$ is non-singular, assume that $\exists x \neq 0$ s.t. $x^T(\mathcal{A}_b \mathscr{P}^{-1} \mathcal{A}_b^T)x = 0$. Since $\mathscr{P}$ is symmetric positive definite, this implies $\mathscr{P}^{-1/2} \mathcal{A}_b^T x = 0$, where $\mathscr{P}^{-1/2}$ is the unique positive definite square root of $\mathscr{P}^{-1}$. Therefore $\mathcal{A}_b^T x = 0$, which is a contradiction. This proves that when $\mathbf{G}$ is weakly connected, $\mathcal{L}$ is invertible if and only if there is at least one reference node.

To examine the situation when $\mathbf{G}$ has more than one weakly connected components, assume w.l.o.g. that it has two components $\mathbf{G}_1 = (\mathbf{V}_1, \mathbf{E}_1)$ and $\mathbf{G}_2 = (\mathbf{V}_2, \mathbf{E}_2)$. Since the two components cannot have an edge or a node in common, the generalized incidence matrix $\mathcal{A}$ of $\mathbf{G}$ can be written as

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}_1 & 0 \\ 0 & \mathcal{A}_2 \end{bmatrix},$$

where $\mathcal{A}_i$ is the generalized incidence matrix of the component $\mathbf{G}_i$. Similarly,

$$\mathcal{A}_b = \begin{bmatrix} \mathcal{A}_{1,b} & 0 \\ 0 & \mathcal{A}_{2,b} \end{bmatrix},$$

where $\mathcal{A}_{i,b}$ correspond to the component $\mathbf{G}_i$. As a result the matrix $\mathcal{L} = \mathcal{A}_b \mathscr{P}^{-1} \mathcal{A}_b^T$ for $\mathbf{G}$ can be written as

$$\mathcal{L} = \begin{bmatrix} \mathcal{A}_{1,b} \mathscr{P}_1^{-1} \mathcal{A}_{1,b}^T & 0 \\ 0 & \mathcal{A}_{2,b} \mathscr{P}_2^{-1} \mathcal{A}_{2,b}^T \end{bmatrix},$$

where $\mathscr{P}_i^{-1}$ contains all the edge-covariance matrices belonging to the edges in $\mathbf{G}_i$. If one of the components, say $\mathbf{G}_1$ does not have a reference node, then $\mathcal{A}_{1,b} = \mathcal{A}_1$ and so $\mathcal{A}_{1,b} \mathscr{P}_1^{-1} \mathcal{A}_{1,b}^T$ is singular, which makes $\mathcal{L}$ singular. If both components have at least one reference node each, each of the diagonal blocks of $\mathcal{L}$ is non-singular, which makes $\mathcal{L}$ invertible. This proves the theorem. ∎

Remark 1. We will shortly see that the elements of $\mathcal{L}$ are independent of the way edge directions are chosen, and so $\mathbf{\Sigma}$ is independent of the edge directions too. The measurement graph is directed due to the need to distinguish between a measurement of $x_u - x_v$ vs. that of $x_v - x_u$, but the variance of the estimation error does not depend on the edge directions chosen.

A. Grounded Laplacian and BLUE

The Laplacian matrix of a graph is usually defined as $L := AA^T$. The matrix $\mathcal{L}$ has a structure similar to the graph Laplacian. To explore this connection, we first define a *matrix-weighted graph* or a *network* as a pair $(\mathbf{G}, W)$ where $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ is a directed graph and $W : \mathbf{E} \to \mathbb{R}^{k \times k}$ is function that assigns symmetric positive definite edge weights $W_e = W_e^T > 0$ to every edge $e \in \mathbf{E}$. For a network $(\mathbf{G}, W)$ with n nodes we define the *Generalized or Matrix- Weighted Graph Laplacian* as

$$\mathcal{L} := \mathcal{A}\mathcal{W}\mathcal{A} \in \mathbb{R}^{kn \times kn}, \tag{14}$$

where $\mathcal{A}$ is the generalized incidence matrix of $\mathbf{G}$, $\mathcal{W}$ is a block matrix with edge weights on its diagonal: $\mathcal{W} = \mathrm{diag}(W_1, \ldots, W_m)$. Expanding it we get

$$\begin{aligned}
\mathcal{L} &= \mathcal{A}\mathcal{W}\mathcal{A}^T \\
&= \begin{bmatrix} \mathcal{A}_b \\ \mathcal{A}_r \end{bmatrix} \mathcal{W} \begin{bmatrix} \mathcal{A}_b^T & \mathcal{A}_r^T \end{bmatrix} \\
&= \left[\begin{array}{c|c} \mathcal{A}_b \mathcal{W} \mathcal{A}_b^T & \mathcal{A}_b \mathcal{W} \mathcal{A}_r^T \\ \hline \mathcal{A}_r \mathcal{W} \mathcal{A}_b^T & \mathcal{A}_r \mathcal{W} \mathcal{A}_r^T \end{array} \right].
\end{aligned}$$

For a measurement graph if we assign edge weights as the inverses of measurement error covariances, i.e., $W_e = P_e^{-1}$ for all $e \in \mathbf{E}$, then $\mathcal{W} = \mathcal{P}$ and $\mathcal{L} = \mathcal{A}_b \mathcal{W} \mathcal{A}_b^T$. So $\mathcal{L}$ is a principal submatrix of the generalized Laplacian. We call $\mathcal{L}$ *the Dirichlet or the Grounded matrix-weighted Laplacian.*

Principal submatrices of the usual graph Laplacian $L = AA^T$ are called Dirichlet Laplacians since they appear in the numerical solution of PDEs with Dirichlet boundary conditions. They also appear in electrical network analysis, hence the moniker "grounded". In fact, we will shortly see that $\mathcal{L}$ plays a key role in a abstract, generalized electrical network with matrix valued currents and voltages.

Note that the matrix- weighted Laplacian defined above in (14) is independent of the edge directions chosen. The matrix $\mathcal{L}$ shares this property, too, and hence the covariance of the optimal estimate Σ does not depend on the edge directions (cf. remark 1).

We list below a few properties of the Laplacian and the incidence matrix that will be used in establishing certain results later in the paper.

Proposition 1. *Let $\mathbf{G}$ be a measurement graph, A its incidence matrix and A_b be the basis incidence matrix constructed by removing the rows corresponding to the reference nodes from A.*
1) *If $\mathbf{G}$ is weakly connected and has n nodes, then the rank of A is $n - 1$, and $\mathbf{1}^T A = 0$, where $\mathbf{1} \in \mathbb{R}^n$ is a vector of all 1's.*
2) *The basis incidence matrix A_b has full row rank if and only if every weakly connected component of the graph has at least one reference node.*
3) *$\mathcal{L}$ has at least (exactly) k zero eigenvalues (if and only if $\mathbf{G}$ is weakly connected),*
4) *and $\mathcal{L} \times (\mathbf{1} \otimes I_k) = 0$.* □

The last two statements are direct consequences of the first two, whose proofs are contained in the proof of Theorem 1.

IV. Electrical Analogy

The main results of the paper are based on the fact that the error covariance matrices of the BLU estimator are numerically equal to the effective resistances in an appropriately defined *generalized* resistive electrical network where currents, potentials and resistances are matrices. Resistances in such a network are always square positive definite matrices in $\mathbb{R}^{k \times k}$ and are called *generalized resistances*. Currents and potentials are matrices in $\mathbb{R}^{k \times k'}$, where $k' \leq k$, and are called *generalized currents* and *generalized potentials*. Generalized currents and potentials (also called voltages) can be non-square matrices, but they must of the same dimension.

In particular, a *generalized electrical network* consists of a connected graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ together with a *generalized resistance* function $R : \mathbf{E} \to \mathbb{R}^{k \times k}$ such that $R(e) = R(e)^T > 0, \forall e \in \mathbf{E}$. For such a network, a *generalized flow from a node $p \in \mathbf{V}$ to another node $q \in \mathbf{V}$ with matrix-intensity $\mathbf{j} \in \mathbb{R}^{k \times k'}$* is a function $j : \mathbf{E} \to \mathbb{R}^{k \times k'}$ with the property that

$$\sum_{\substack{(u,v) \in \mathbf{E} \\ u = \bar{u}}} j_{u,v} - \sum_{\substack{(v,u) \in \mathbf{E} \\ u = \bar{u}}} j_{v,u} = \begin{cases} \mathbf{j} & \bar{u} = p \\ -\mathbf{j} & \bar{u} = q \\ \mathbf{0} & \text{otherwise} \end{cases} \tag{15}$$

$\forall \bar{u} \in V$. In other words, the flow "out of" a node is 0 except for the nodes p and q, and the flow out of p is the same as the flow into q, which is equal to the flow intensity $\mathbf{j}$. Note that (15) is the "generalized" version of Kirchoff's current law. The notation

$$\sum_{\substack{(u,v) \in \mathbf{E} \\ u = \bar{u}}}$$

means the summation is taken over all edges in $\mathbf{E}$ that are incident on, and directed away from, the node $\bar{u}$.

A *generalized current from* $p \in \mathbf{V}$ *to* $q \in \mathbf{V}$ *with matrix-intensity* $\mathbf{i} \in \mathbb{R}^{k \times k'}$ is a flow $i : \mathbf{E} \to \mathbb{R}^{k \times k'}$ for which there exists a function $V : \mathbf{V} \to \mathbb{R}^{k \times k'}$ such that

$$R_{u,v} i_{u,v} = V_u - V_v, \quad \forall (u,v) \in \mathbf{E}. \tag{16}$$

The function V is called a *generalized node potential associated with the current* i. Equation (16) can be thought of the combined outcome of generalized versions of Kirchoff's voltage law and Ohm's law. In "regular" electrical networks, Kirchoff's voltage law states that the net potential drop in a circuit is 0, which shows there must be a potential function defined on nodes. Combined with Ohm's law which relates potential drop along an edge to it resistance and the current flowing through it, we get a relationship of the form (16) in "regular" networks. In generalized networks we simply impose the same relationship, while keeping in mind that the operations are now between matrices.

An electrical network may have currents injected and extracted at more than one node. A *divergence* for a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ is a function $\omega : \mathbf{V} \to \mathbb{R}^{k \times k'}$ so that $\sum_{u \in \mathbf{V}} \omega_u = 0$. One can view a divergence as an assignment of flow sources at nodes of a graph so that total flow into the graph is equal to the total flow out of it. A *flow j with divergence ω in a graph* $\mathbf{G}$ is a function $j : \mathbf{E} \to \mathbb{R}^{k \times k'}$ so that

$$\sum_{\substack{(u,v) \in \mathbf{E} \\ u = \bar{u}}} j_{u,v} - \sum_{\substack{(v,u) \in \mathbf{E} \\ u = \bar{u}}} j_{v,u} = \omega_{\bar{u}}, \quad \bar{u} \in \mathbf{V}. \tag{17}$$

One can think of the above as the counterpart of Kirchoff's current law (15) when current is injected and extracted at more than one node. We can represent it in a more compact form by using the incidence matrix of $\mathbf{G}$. For a flow j with divergence ω, stacking all the edge flows j_e, $e \in \mathbf{E}$ into one tall matrix $\mathcal{J} \in \mathbb{R}^{mk \times k'}$ and stacking all the node divergences ω_u, $u \in \mathbf{V}$ into one tall matrix $\boldsymbol{\omega} \in \mathbb{R}^{nk \times k'}$, we can express (17) as

$$\mathcal{A} \mathcal{J} = \boldsymbol{\omega}. \tag{18}$$

This set of equations represent a set of nkk' linear constraints on the flow j. However, not all these constraints are independent since $\mathcal{A}$ does not have full row rank (Prop. 1). Let $o \in \mathbf{V}$ and let A_b be the submatrix of the incidence matrix $\mathcal{A}$ of $\mathbf{G}$ obtained by removing a row corresponding to o from $\mathcal{A}$. Since $\mathbf{G}$ is weakly connected, $\mathcal{A}_b = A_b \otimes I_k$ has full row rank (see Prop. 1), and so the maximum number of independent constraints is $(n-1)kk'$, and these constraints are given in a compact form by

$$\mathcal{A}_b \mathcal{J} = \boldsymbol{\omega}_b, \tag{19}$$

where $\boldsymbol{\omega}_b \in \mathbb{R}^{k(n-1) \times k'}$ is the tall vector obtained by stacking all but one node-divergences, that node being node $o \in \mathbf{V}$.

For $k = 1$, generalized electrical networks are the usual electrical networks with scalar currents, potentials, and resistors. Given the source currents in a regular electrical network, Kirchoff's current and voltage laws and Ohm's law can be used to compute all the edge currents and edge potentials in the network. However, there is an alternative way of solving for the currents, which has to do with the characterization of current as the flow with minimum energy dissipation. The energy dissipation in the network is simply the sum of the energy dissipated by each resistor r_e in the network: $\sum_e i_e^2 r_e$. That current minimizes the energy dissipation among all flows in a regular electrical network is known as Thomson's minimum energy principle [5]. We will show that Thomson's principle also holds for generalized networks, and use it to compute the generalized currents. This method also yields the node voltages in a surprisingly elegant way.

To proceed it is convenient to introduce the following definition: The *energy dissipated by a generalized flow j from* $p \in \mathbf{V}$ *to* $q \in \mathbf{V}$ is defined to be

$$\sum_{(u,v) \in \mathbf{E}} Tr\left(j_{u,v}^T R_{u,v} j_{u,v}\right).$$

The following technical lemma helps us prove several important results.

Lemma 1. *Consider a generalized electric network $G = (\mathbf{V}, \mathbf{E}, R)$. Let i and j be a current and a flow, respectively, both from $p \in \mathbf{V}$ to $q \in \mathbf{V}$ with intensities $\mathbf{i} \in \mathbb{R}^{k \times k'}$ and $\mathbf{j} \in \mathbb{R}^{k \times k'}$. Then*

$$\sum_{(u,v) \in \mathbf{E}} i_{u,v}^T R_{u,v} j_{u,v} = (V_p - V_q)^T \mathbf{j},$$

where V is a potential associated with i. □

Proof: [Proof of Lemma 1] From (16), we conclude that

$$\sum_{(u,v) \in \mathbf{E}} i_{u,v}^T R_{u,v} j_{u,v} = \sum_{(u,v) \in \mathbf{E}} (V_u - V_v)^T j_{u,v}$$

$$= \sum_{\bar{u} \in \mathbf{V}} V_{\bar{u}}^T \sum_{\substack{(u,v) \in \mathbf{E} \\ u = \bar{u}}} j_{u,v} - \sum_{\bar{v} \in \mathbf{V}} V_{\bar{v}}^T \sum_{\substack{(u,v) \in \mathbf{E} \\ v = \bar{v}}} j_{u,v}.$$

Using (15) to expand the first term on the right hand side, we then conclude that

$$\sum_{(u,v) \in \mathbf{E}} i_{u,v}^T R_{u,v} j_{u,v} = (V_p - V_q)^T \mathbf{j} + \sum_{\bar{u} \in \mathbf{V}} V_{\bar{u}}^T \sum_{\substack{(v,u) \in \mathbf{E} \\ u = \bar{u}}} j_{v,u} - \sum_{\bar{v} \in \mathbf{V}} V_{\bar{v}}^T \sum_{\substack{(u,v) \in \mathbf{E} \\ v = \bar{v}}} j_{u,v}.$$

The result follows from the fact that the two summations are equal and therefore cancel each other.

Now we are ready to state Thomson's minimum energy principle for generalized networks.

Theorem 2 (Generalized Thompson's Minimum Energy Principle). *Among all flows with the same matrix-intensity between two nodes in a generalized electrical network, current minimizes the energy dissipation. Specifically, Let i and j be a current and a flow, respectively, both from $p \in \mathbf{V}$ to $q \in \mathbf{V}$ with the same matrix-intensity. Then*

$$\sum_{(u,v) \in \mathbf{E}} Tr\left(i_{u,v}^T R_{u,v} i_{u,v}\right) \leq \sum_{(u,v) \in \mathbf{E}} Tr\left(j_{u,v}^T R_{u,v} j_{u,v}\right). \tag{20}$$

Conversely, a flow i that satisfies (20) for all flows j between the same pair of nodes with the same intensity must be a current. □

Proof: [Proof of Theorem 2] Defining $d := j - i$, this function is a flow from $p \in \mathbf{V}$ to $q \in \mathbf{V}$ with zero intensity. Moreover,

$$\sum_{(u,v) \in \mathbf{E}} Tr(j_{u,v}^T R_{u,v} j_{u,v}) = \sum_{(u,v) \in \mathbf{E}} Tr(i_{u,v}^T R_{u,v} i_{u,v}) + \sum_{(u,v) \in \mathbf{E}} Tr(d_{u,v}^T R_{u,v} d_{u,v}) + 2 \sum_{(u,v) \in \mathbf{E}} Tr(i_{u,v}^T R_{u,v} d_{u,v}). \tag{21}$$

Since d has zero intensity and $i_{u,v}^T R_{u,v} = (V_u - V_v)^T$, we conclude from lemma 1 that the last term of (21) is zero. The inequality (20) then follows because the term before the last is nonnegative.

To prove the converse result, we use the fact that if i satisfies (20) for all flows j, then it minimizes the quadratic cost

$$\sum_{(u,v) \in \mathbf{E}} Tr\left(j_{u,v}^T R_{u,v} j_{u,v}\right)$$

subject to the linear constraint specified by generalized Kirchoff's current laws. Assume temporarily that $k' = 1$, i.e., the generalized edge-flows $j_e, \forall e \in \mathbf{E}$ and the flow intensity $\mathbf{j}$ are vectors in $\mathbb{R}^k$. Denote by $\mathcal{I} \in \mathbb{R}^{km}$ the tall matrix obtained by stacking all the edge currents $i_e, e \in \mathbf{E}$. Then, $\mathcal{I}$ minimizes the energy consumption $\mathcal{E}(\mathcal{J}) := \mathcal{J}^T \mathcal{R} \mathcal{J}$ subject to the constraint $h(\mathcal{J}) = 0$ where $h(\mathcal{J}) := \mathcal{A}_b \mathcal{J} - \boldsymbol{\omega}$. Note that we only use the linearly independent set of constraints given by (19). Since $\mathcal{A}_b$ has full row rank, $h'(\mathcal{J}) = \mathcal{A}_b^T$ has linearly independent columns and hence the constraint is regular. From the Lagrange Multiplier Theorem [31, Chapter 3] we conclude that there must exist Lagrange Multiplier vector [4] $\lambda \in \mathbb{R}^{(n-1)k}$ such that

$$\mathcal{E}'(\mathcal{J}) + h'(\mathcal{I})\lambda = 0,$$

from which we get

$$2\mathcal{R}\mathcal{I} + \mathcal{A}_b^T \lambda = 0.$$

This is equivalent to

$$2R_e i_e + \lambda_u - \lambda_v = 0 \quad e = (u,v), u \neq q, v \neq q,$$
$$2R_e i_e + \lambda_u = 0 \quad e = (u,q), u \neq q,$$
$$2R_e i_e - \lambda_v = 0 \quad e = (q,v), v \neq q.$$

---

[4] Draft note: the previous proof was incorrect. Regularity was overlooked, but more importantly, the claim "there must exist Lagrange multipliers $\lambda_u, u \in \mathbf{V}$" was incorrect. Only $|\mathbf{V}| - 1$ multipliers can be determined, and the reason again has to do with regularity.

Thus (16) holds with the potential $V_u = -\lambda_u/2$ for all $u \in \mathbf{V} \setminus \{q\}$ and $V_q = 0$. The same can be shown to hold when currents are $k \times k'$ matrices instead of k-vectors, by first converting the matrix valued quantities into vectors by stacking the columns to apply the Lagrange multiplier theorem. This proves that i is indeed a current. ∎

Theorem 3 (Existence, Uniqueness and Linearity). *Assume that* $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ *is weakly connected and that* $R : \mathbf{E} \to \mathbb{R}^{k \times k}$ *is a generalized resistance function. For a given matrix-intensity* $\mathbf{i} \in \mathbb{R}^{k \times k'}$, *there is a unique current* $i : \mathbf{E} \to \mathbb{R}^{k \times k'}$ *from* $p \in \mathbf{V}$ *to* $q \in \mathbf{V}$ *and the corresponding potential* $V : \mathbf{V} \to \mathbb{R}^{k \times k'}$ *is also unique when one enforces the normalization constraint* $V_q = 0$. *Moreover,* i *and* V *are linear functions of the intensity* $\mathbf{i}$. □

Proof: We start by showing that a flow j from $p \in \mathbf{V}$ to $q \in \mathbf{V}$ with vector-intensity $\mathbf{i}$ always exists. To construct such a flow, take a path $\mathcal{P}$ from p to q, and define j as follows

$$j_{u,v} = \begin{cases} \mathbf{i} & \text{orientations of } (u,v) \text{ and } \mathcal{P} \text{ coincide} \\ -\mathbf{i} & \text{orientations of } (u,v) \text{ and } \mathcal{P} \text{ are opposite} \\ 0 & \text{otherwise,} \end{cases}$$

where the "orientation" of an edge $e = (u, v)$ is said to coincide with that of $\mathcal{P}$ if e is traversed from u to v, and is said to be opposite if e is traversed from v to u in $\mathcal{P}$ (see Section II). By construction, such a flow satisfies generalized Kirchoff's current law (15). Having established that this constraint is feasible, there must exist[5] a flow minimizes the quadratic cost

$$\sum_{(u,v) \in \mathbf{E}} Tr\left(j_{u,v}^T R_{u,v} j_{u,v}\right)$$

subject to the linear constraint (15). The quadratic cost is the energy dissipated by the flow j. Since current minimizes energy dissipation among all flows with the same intensity (Theorem 2), the flow that minimizes the quadratic cost is a current from $p \in \mathbf{V}$ to $q \in \mathbf{V}$ with intensity $\mathbf{i}$.

To prove uniqueness, let i and $\tilde{i}$ be two currents from $p \in \mathbf{V}$ to $q \in \mathbf{V}$ with intensity $\mathbf{i}$. Defining $d := i - \tilde{i}$, we see this function is a current from $p \in \mathbf{V}$ to $q \in \mathbf{V}$ with zero intensity. Therefore, we conclude from lemma 1 that

$$\sum_{(u,v) \in \mathbf{E}} d_{u,v}^T R_{u,v} d_{u,v} = 0 \qquad \Rightarrow d_{u,v} = 0 \quad \forall (u,v) \in \mathbf{E},$$

since all the $R_{u,v} > 0$. We conclude that $d = 0$ and therefore $i = \tilde{i}$.

Suppose now that V and $\tilde{V}$ are two potentials associated with i, normalized by $V_q = \tilde{V}_q = 0$. Because of (16), we conclude that

$$V_u - V_v = \tilde{V}_u - \tilde{V}_v \quad \Leftrightarrow \quad D_u = D_v, \quad \forall (u,v) \in \mathbf{E}.$$

where $D = V - \tilde{V}$. Since $\mathbf{G}$ is connected, we conclude that D must be constant in $\mathbf{V}$. Since $D_q = 0$, we conclude that D is identically zero and therefore $V = \tilde{V}$.

The linearity from $\mathbf{i}$ to i results from the fact that if i is a current with intensity $\mathbf{i}$ and potential V, and $\bar{i}$ is a current with intensity $\bar{\mathbf{i}}$ and potential $\bar{V}$, then $\alpha i + \beta \bar{i}$ is also a current with vector-intensity $\alpha \mathbf{i} + \beta \bar{\mathbf{i}}$ and potential $\alpha V + \beta \bar{V}$, for any real scalars α, β. ∎

A. Generalized Effective Resistance and BLUE covariance

Consider a generalized electric network $(\mathbf{G}, R^{-1})$ with graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ and a positive definite generalized resistance function $R : \mathbf{E} \times \mathbb{R}^{k \times k}$. Let i be a current of matrix-intensity $\mathbf{i} \in \mathbb{R}^{k \times k'}$ from $p \in \mathbf{V}$ to $q \in \mathbf{V}$. Since the resulting generalized potential V is a linear function of the matrix-intensity $\mathbf{i}$, for each $v \in \mathbf{V}$ there must exist a matrix $R_{p,v;q}^{\text{cross}}$ such that

$$V_v - V_q = R_{p,v;q}^{\text{cross}} \mathbf{i}, \quad \forall \mathbf{i} \in \mathbb{R}^{k \times k'}. \tag{22}$$

We call such matrix the *generalized cross resistance from p to v with respect to q*. When $v = p$, we call it the *generalized effective resistance between p and q*, i.e.,

$$V_p - V_q = R_{p,q}^{\text{eff}} \mathbf{i}, \quad \forall \mathbf{i} \in \mathbb{R}^k, \tag{23}$$

From Theorem 3 and lemma 1, we get the following corollary:

---

[5] *Draft note: where does this existence claim come from? If we can't quote a theorem about convex functional or something, we have to say stuff about Hilbert space and minimum norm problem to justify existence.*
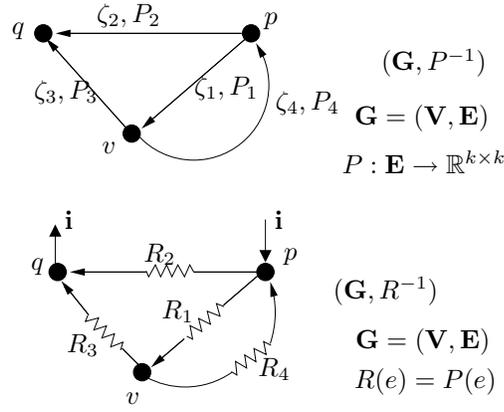
Fig. 5. A measurement graph $(\mathbf{G}, P^{-1})$ and its analogous electrical network $(\mathbf{G}, R^{-1})$. A generalized current is injected at p and extracted at q in the electrical network $(\mathbf{G}, R^{-1})$.

Corollary 1.

$$\sum_{(u,v)\in\mathbf{E}} i_{u,v}^T R_{u,v} i_{u,v} = \mathbf{i}^T R_{p,q}^{\text{eff}} \mathbf{i}, \quad \forall \mathbf{i} \in \mathbb{R}^{k \times k'}, \tag{24}$$

where i is the current in the network with matrix-intensity $\mathbf{i}$ from p to q. □

When the matrix-intensity of the injected current is the identity matrix, then the right hand side of (24) is simply the effective resistance. This is an analog of the interpretation of the effective resistance in regular electrical networks as the total energy dissipation when 1 Ampere current flows through the network[6].

It should be emphasized that the effective resistances $R_{p,q}^{\text{eff}}$ for all pairs of nodes $p, q \in \mathbf{V}$ as well as the edge resistances $R_e, e \in \mathbf{E}$ are always symmetric positive definite $k \times k$ matrices, whereas the currents and the voltages can be $k \times k'$ matrices for any $1 \le k' \le k$. Moreover, the effective resistance between two nodes only depends on the graph $\mathbf{G}$ and the edge resistances R_e. Whether we calculate it by injecting matrix-valued currents into the network or vector-valued currents, its value will be the same.

For the purpose of brevity, from now on we refer to generalized resistances simply as resistances when there is no room for confusion.

From the measurement graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ with an edge covariance function $P : \mathbf{E} \to \mathbb{R}^{k \times k}$, we form the *analogous* generalized electrical network $(\mathbf{G}, R^{-1})$ on the same graph $\mathbf{G}$ by assigning a generalized electrical resistance $R(e)$ to every edge $e \in \mathbf{E}$ equal to the covariance matrix $P(e)$ of the measurement error vector associated with that edge, i.e., $R(e) = P(e)$. Figure 5 shows an example of a measurement graph and the corresponding generalized electrical network.

Recall that $\Sigma_{u,o}$ is the error covariance of the BLU estimate of x_u when o is the reference node, i.e., $\Sigma_{u,o} = \text{E}[(x_u - \hat{x}_u^*)(x_u - \hat{x}_u^*)^T]$. We denote by $\Sigma_{u,v;o}$ the covariance between the estimation errors of x_u and x_v, i.e., $\Sigma_{u,v;o} = \text{E}[(x_u - \hat{x}_u^*)(x_v - \hat{x}_v^*)^T]$.

The following is the main result of this section, which we call by the name *Electrical Analogy*.

Theorem 4 (Electrical Analogy). *Consider a weakly connected measurement graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ with a single reference node $o \in \mathbf{V}$, consisting of $|\mathbf{V}| = n$ nodes and $|\mathbf{E}| = m$ edges. Construct the analogous generalized electrical network $(\mathbf{G}, R^{-1})$ by assigning edge-resistances $R(e) = P(e)$, where $P(e) \in \mathbb{R}^{k \times k}$ is the covariance matrix of the error in the measurement ζ_e, i.e., $P(e) = \text{E}[\epsilon_e \epsilon_e^T]$. Then,*

1) *The covariance matrix of the estimation error of x_u is numerically equal to $R_{u,o}^{\text{eff}}$, that is,*

$$\Sigma_{u,o} = R_{u,o}^{\text{eff}}.$$

2) *The covariance matrix between the estimation errors of x_u and x_v is numerically equal to $R_{u,v;o}^{\text{cross}}$:*

$$\Sigma_{u,v;o} = R_{u,v;o}^{\text{cross}}.$$

---

[6] *Draft note: We can't go further w/o defining $\sum_{(u,v)\in\mathbf{E}} i_{u,v}^T R_{u,v} i_{u,v}$ as some sort of pseudo-energy. If we did, we can make the analogy more exact. Without that, is this at all helpful?*

By indexing the nodes such that the reference $o \in \mathbf{V}$ is the n-th node in a network of n nodes, the analogy between terms of the BLU estimator covariance matrix with effective and cross-resistances can be expressed in an compact form in the following way:

$$\Sigma = \mathbf{R}^{\text{eff}}, \tag{25}$$

where $\mathbf{R}^{\text{eff}} \in \mathbb{R}^{k(n-1) \times k(n-1)}$ is defined as

$$\mathbf{R}^{\text{eff}} = \begin{bmatrix} R_{1,n}^{\text{eff}} & R_{1,2;n}^{\text{cross}} & .. & \\ R_{2,1;n}^{\text{cross}} & R_{2,n}^{\text{eff}} & .. & \\ & & \ddots & R_{n-2,n-1;n}^{\text{cross}} \\ .. & & R_{n-1,n-2;n}^{\text{cross}} & R_{n-1,n}^{\text{eff}} \end{bmatrix}$$

3) *Let the optimal estimate of a node variable x_u be given by*

$$\hat{x}_u^* = \sum_{e \in \mathbf{E}} c_e \zeta_e, \tag{26}$$

where $\{c_e, e \in \mathbf{E}\}$ define the BLU estimator of x_u. Let $i_e, e \in \mathbf{E}$ be the edge currents when a current of matrix-intensity I_k is injected at u and extracted at o, where $I_k \in \mathbb{R}^{k \times k}$ is the identity matrix. Then

$$c_e = i_e, \quad \forall e \in \mathbf{E}. \qquad \qquad \square$$

A proof of the first and the third statements of the theorem above, for the special case $k = 1$, was provided by Karp *et. al.* [8]. They considered the problem of estimating time differences among clocks in a network of sensors. Theorem 4 provides an electrical analogy also for the covariances, and more importantly, generalizes this result to estimation of vector valued variables.

We have noted earlier that $\mathcal{L}$ is independent of edge directions though $\mathbf{G}$ is directed, and that the BLUE covariances are independent of the edge directions, too. Similarly, though the graph of an electrical network is directed, the effective resistances does not depend on the edge directions. However, the currents will depend on the edge directions, and so will the BLU *estimator* of a node variable x_u.

Proof: [Proof of Theorem 4] We put a current source between u and o external to the network so that a current of matrix-intensity $\mathbf{i} \in \mathbb{R}^{k \times k}$ flows from u to o through the network, and then compute the edge currents and node potentials. By stacking together all the generalized edge currents in a tall matrix $\mathscr{I} \in \mathbb{R}^{km \times k}$, all the generalized node potentials in a tall matrix $\mathcal{V} \in \mathbb{R}^{kn \times k}$ and all the generalized edge-resistances diagonally in a square matrix $\mathscr{R} = \text{diag}\,(R_1, \ldots, R_m) \in \mathbb{R}^{km \times km}$, we can write the generalized Kirchoff's and Ohm's Laws described in section IV as follows:

1) Generalized Kirchoff's current law:

$$\mathcal{A}\mathscr{I} = e_{uo} \otimes I_k, \tag{27}$$

where $e_{uo} \in \mathbb{R}^n$, is a vector of all zeros except at the u-th position, where it has a 1, and at the o-th position, where it has a -1.

2) Generalized Kirchoff's voltage law and Ohm's law:

$$\mathscr{I} = \mathscr{R}^{-1}\mathcal{A}^T\mathcal{V}, \tag{28}$$

Pre-multiplying both side of (28) by $\mathcal{A}$ and using (27), we get

$$\mathcal{A}\mathscr{R}^{-1}\mathcal{A}^T\mathcal{V} = e_{uo} \otimes I_k. \tag{29}$$

Since $\mathscr{P} = \mathscr{R}$, we have $\mathcal{L} = \mathcal{A}\mathscr{P}^{-1}\mathcal{A}^T = \mathcal{A}\mathscr{R}^{-1}\mathcal{A}^T$. Had $\mathcal{L}$ been invertible, we could have solved for the generalized node potentials from this equations. However, since it is singular (cf. Proposition 1), we impose the normalization constraint that the node potential at the reference node is 0. Then $\mathcal{V}$ can be written as $[\mathcal{V}_b^T, 0]^T$ where $\mathcal{V}_b \in \mathbb{R}^{k(n-1) \times k}$ is the tall matrix of the node potentials of all the nodes except the reference. We expand the product $\mathcal{A}\mathscr{R}^{-1}\mathcal{A}^T$ using $\mathcal{A} = [\mathcal{A}_b^T, \mathcal{A}_r^T]^T$, where $\mathcal{A}_r$ consists of the rows of $\mathcal{A}$ corresponding to the single reference node $o \in \mathbf{V}$ and $\mathcal{A}_b$ of all the rest. Indexing the rows and columns of the matrices in question appropriately, (29) can now be written as

$$\left[\begin{array}{c|c} \mathcal{A}_b\mathscr{R}^{-1}\mathcal{A}_b^T & \mathcal{A}_b\mathscr{R}^{-1}\mathcal{A}_r^T \\ \hline \mathcal{A}_r\mathscr{R}^{-1}\mathcal{A}_b^T & \mathcal{A}_r\mathscr{R}^{-1}\mathcal{A}_r^T \end{array} \right] \begin{bmatrix} \mathcal{V}_b \\ 0 \end{bmatrix} = \begin{bmatrix} e_u \otimes I_k \\ \hline -I_k \end{bmatrix}$$

where $e_u \in \mathbb{R}^{n-1}$ is vector with all zeros except at the uth position, where it has a 1. The resulting equations are

$$\mathcal{L}\mathcal{V}_b = e_u \otimes I_k,$$
$$\mathcal{A}_r\mathscr{R}^{-1}\mathcal{A}_b^T\mathcal{V}_b = -I_k$$

Since $\mathbf{G}$ is weakly connected, by theorem 1, $\mathcal{L}$ is invertible. The node potentials can now be determined from the first equation as

$$\mathcal{V}_b = \mathcal{L}^{-1}(e_u \otimes I_k), \tag{30}$$

which satisfies the second equation, as can be easily verified. Therefore, by theorem 3, eq. (30) is the unique solution of node potentials (with the normalizing constraint $V_o = 0$). Since $\Sigma = \mathcal{L}^{-1}$ (Theorem 1), we have

$$
\begin{aligned}
\Sigma_{u,v;o} &= \mathrm{E}[(x_u - \hat{x}_u^*)(x_v - \hat{x}_v^*)^T] \\
&= \mathrm{E}[(x_v - \hat{x}_v^*)(x_u - \hat{x}_u^*)^T] \\
&= (e_v \otimes I_k)^T \mathcal{L}^{-1}(e_u \otimes I_k) \\
&= (e_v \otimes I_k)^T \mathcal{V}_b \\
&= R_{u,v;o}^{\mathrm{cross}}
\end{aligned}
$$

where the last equality follows from the definition of cross resistance given by (22) and from (30). The first two statement of the theorem follow immediately.

We construct $n - 1$ such node potential solutions, each time injecting "unit" current at a node u and extracting at o, for every $u \in \mathbf{V} \setminus \{o\}$. Denote each node voltage solution "vector" by $\mathcal{V}_b^u$. Stacking these solutions together horizontally we get

$$
\begin{aligned}
\left[\mathcal{V}_b^1|\ldots|\mathcal{V}_b^{n-1}\right] &= \mathcal{L}^{-1}\left[e_1|\ldots|e_{n-1}\right] \otimes I_k \\
&= \mathcal{L}^{-1}.
\end{aligned}
$$

On the other hand, the definitions of effective and cross resistances (23),(22) yield

$$\left[\mathcal{V}_b^1|\ldots|\mathcal{V}_b^{n-1}\right]^T = \mathbf{R}^{\mathrm{eff}},$$

which proves (25).

To prove the last part of the theorem about the currents defining the optimal estimator, we write using (28) and (30)

$$
\begin{aligned}
\mathscr{I} &= \mathscr{R}^{-1}[\mathcal{A}_b^T \ \mathcal{A}_r^T]\begin{bmatrix} \mathcal{L}^{-1}(e_u \otimes I_k) \\ 0 \end{bmatrix} \\
&= \mathscr{R}^{-1}\mathcal{A}_b^T\mathcal{L}^{-1}(e_u \otimes I_k)
\end{aligned} \tag{31}
$$

Let $\hat{x}_u^* = \mathbf{c}^T\mathbf{z}$, i.e., $\mathbf{c}$ is a tall matrix obtained by stacking the coefficient matrices c_e in (26). From the BLU estimate (12), and since $\mathscr{P} = \mathscr{R}$, we see that $\hat{x}_u^* = (e_u \otimes I_k)^T\mathcal{L}^{-1}\mathcal{A}_b\mathscr{R}^{-1}\mathbf{z}$. Therefore,

$$\mathbf{c} = \mathscr{R}^{-1}\mathcal{A}_b^T\mathcal{L}^{-1}(e_u \otimes I_k). \tag{32}$$

comparing (32) with (31), the last statement of the theorem follows. ∎

The optimal estimates themselves have an interesting electrical analog in terms of generalized potentials, which is presented in the next theorem.

Theorem 5 (Further Electrical Analogy). *Consider a measurement graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ with a single reference node $o \in \mathbf{V}$ and edge-covariances $P(e), e \in \mathbf{E}$, and its corresponding generalized electrical network $(\mathbf{G}, P^{-1})$. Consider the divergence*

$$\boldsymbol{\omega} = \mathcal{A}\mathscr{P}^{-1}\mathbf{z} \in \mathbb{R}^{kn}, \tag{33}$$

for the graph $\mathbf{G}$, where $\mathbf{z}$ is the tall vector obtained by stacking together all the measurements. If the generalized potential at the node o is fixed as the known node variable x_o as a normalizing constraint, i.e., $V_o = x_o$, the resulting potential of a node u in the generalized electrical network $(\mathbf{G}, R)$ is the optimal estimate $\hat{x}_u^*$ in the measurement graph $\mathbf{G}$:

$$\hat{x}_u^* = V_u, \quad \forall u \in \mathbf{V}. \qquad \square$$

It should be emphasized that the currents and the potentials are in this case k-vectors, not matrices.

Proof: [Proof of Theorem 5] First define $\boldsymbol{\omega}_b = \mathcal{A}_b\mathscr{P}^{-1}\mathbf{z}$. We solve for the node potentials with normalizing constraint $V_o = x_o$. The compact version of Kirchoff's current law (18) gives us $\mathcal{A}\mathcal{I} = \boldsymbol{\omega}$, and proceeding again as in the proof of theorem 4, we get

$$
\left[\begin{array}{c|c} \mathcal{A}_b\mathscr{R}^{-1}\mathcal{A}_b^T & \mathcal{A}_b\mathscr{R}^{-1}\mathcal{A}_r^T \\ \hline \mathcal{A}_r\mathscr{R}^{-1}\mathcal{A}_b^T & \mathcal{A}_r\mathscr{R}^{-1}\mathcal{A}_r^T \end{array}\right]\begin{bmatrix} \mathcal{V}_b \\ 0 \end{bmatrix} = \begin{bmatrix} \mathcal{A}_b\mathscr{P}^{-1}\mathbf{z} \\ \mathcal{A}_r\mathscr{P}^{-1}\mathbf{z} \end{bmatrix}.
$$

since $\mathscr{P} = \mathscr{R}$, the solution for the node potentials is

$$\mathcal{V}_b = \mathcal{L}^{-1} \mathcal{A}_b \mathscr{P}^{-1} \mathbf{z}.$$

The right hand side is the optimal estimate of the node variables $\hat{\mathbf{x}}^*$ (from (12)). This concludes the proof. ■

The following two propositions are straightforward extensions of the series and parallel resistor results for regular electrical networks to the generalized case.

Proposition 2 (Generalized Resistors in Series). *The effective resistance between two nodes in a tree is the sum of edge resistances along the unique shortest path between them.*

Proposition 3 (Generalized Resistors in Parallel). *Consider an electrical network $(\mathbf{G}, R^{-1})$ where the graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ may have multiple edges between a pair of nodes. Construct a graph $\mathbf{G}' = (\mathbf{V}, \mathbf{E}')$ with the same nodes as $\mathbf{V}$ such that every pair of nodes that has an edge in $\mathbf{E}$ also has an edge in $\mathbf{E}'$, but without multiple edges between the same pair of nodes. If the edge resistances in $\mathbf{E}'$ are assigned as*

$$R_{u,v}'^{-1} = \sum_{\substack{e \sim u,v \\ e \in \mathbf{E}}} R_e^{-1}, \quad \forall (u,v) \in \mathbf{E}',$$

where $e \sim u, v$ denotes all edges that are incident on the two nodes u and v, the effective resistance between any pair of nodes in $(\mathbf{G}, R^{-1})$ is the same as that between them in $(\mathbf{G}, R'^{-1})$.

In other words, parallel edges between two nodes u and v in a graph can replaced by a single edge with a resistance that is given by the "parallel resistance formula" $\left(\sum_{e \sim u,v} R_e^{-1} \right)^{-1}$ without changing the effective resistances. The proof of this lemma follows simply upon constructing the Dirichlet Laplacians $\mathcal{L}_{(R^{-1})}$ and $\mathcal{L}_{(R'^{-1})}$ of the two networks $(\mathbf{G}, R^{-1})$ and $(\mathbf{G}', R'^{-1})$ and using the result $\mathcal{L}^{-1} = \mathbf{R}^{\text{eff}}$ from (25).

If the effective resistance between two nodes in a "regular" electrical network with 1-Ohm resistors is known, then the generalized effective resistance in the same graph with equal matrix resistances can be determined quickly with the help of the following result.

Lemma 2. *Let $r_{u,v}^{\text{eff}}$ denote the* scalar *effective resistance between two nodes u and v in an a electrical network $(\mathbf{G}, 1)$ that has 1-Ohm resistors on every edge of the graph $\mathbf{G}$. Let $(\mathbf{G}, R_o^{-1})$ be a generalized electrical network constructed from the same graph $\mathbf{G}$ by assigning equal generalized resistance $R_o \in \mathbb{R}^{k \times k}$ to every edge of $\mathbf{G}$. Then, the generalized effective resistance $R_{u,v}^{\text{eff}}$ is related to $r_{u,v}^{\text{eff}}$ by*

$$R_{u,v}^{\text{eff}} = r_{u,v}^{\text{eff}} R_o. \qquad \square$$

Proof: Since all the generalized edge resistances are equal to R_o, we have $\mathscr{R} = \text{diag}(R_o, \dots, R_o) = I_m \times R_o$, and so the Dirichlet Laplacian $\mathcal{L}$ for the network $(\mathbf{G}, R_o)$ satisfies

$$
\begin{aligned}
\mathcal{L}^{-1} &= \left((A_b \otimes I)(I_m \otimes R_o)^{-1}(A_b \otimes I)^T \right)^{-1} \\
&= \left((A_b A_b^T) \otimes R_o^{-1} \right)^{-1} \\
&= (A_b A_b^T)^{-1} \otimes R_o
\end{aligned}
$$

Now $L_b = A_b A_b^T$ is the Dirichlet Laplacian of the weighted graph $(\mathbf{G}, 1)$ with every edge-weight equal to 1. From Theorem 4 for the case $k = 1$, we know that $(L_b^{-1})_{u,u} = r_{u,v}^{\text{eff}}$, from which and the last equation above, the result follows. ■

It is known that the scalar effective resistance is a distance metric, since it is symmetric $r_{u,v}^{\text{eff}} = r_{v,u}^{\text{eff}}$, positive definite $r_{u,v}^{\text{eff}} \geq 0$ and $r_{u,v}^{\text{eff}} = 0$ only when $u = v$, and obeys the triangle inequality [32]. For this reason effective resistance is sometimes called *resistance distance*. It turns out that the generalized effective resistance also obeys the triangle inequality. This is stated formally in the next theorem[7].

Theorem 6 (Triangle Inequality for Effective Resistances). *Let p, r and q be three nodes in an electrical network with graph $G = (\mathbf{V}, \mathbf{E})$ with generalized resistance function $R : \mathbf{E} \to \mathbb{R}^{k \times k}$. Let G be weakly connected. Then,*

$$R_{p,q}^{\text{eff}} \leq R_{p,r}^{\text{eff}} + R_{r,q}^{\text{eff}}, \tag{34}$$

and the equality holds if every undirected path from p to q passes through r[8]. $\square$

Proof: see Appendix I.

---

[7] *Draft note: The symmetry and positive-definiteness properties are not proved since they have not been used visibly. But perhaps we should?*

[8] Conjecture: this is an if and only if condition, but so far I have been able to prove only the "if" part.

V. EFFECTIVE RESISTANCE AND GRAPH EMBEDDING

The effective resistance in a "regular" electrical network obeys Rayleigh's monotonicity law, which states that *if the resistances of a circuit are increased, the effective resistance between any two points can only increase, and if they are decreased, it can only decrease.* [5]. We will show that this monotonicity property is true even for generalized electrical networks. The monotonicity of effective resistance can be used to bound the effective resistance in a complicated graph by that in a simpler graph, one in which the growth of the effective resistance with distance is known. To exploit the full power of this approach, we need to not only reduce or increase resistances, but also to remove edges altogether or introduce new edges. For a "regular" electrical network, removing a resistor corresponds to making the resistance of that edge to infinity. In generalized networks however one needs to be more careful since a matrix can become unbounded in many ways. This can be accommodated through "graph embedding". We now proceed to make these ideas precise.

A. Graph Embedding

Given two graphs $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ and $\bar{\mathbf{G}} = (\bar{\mathbf{V}}, \bar{\mathbf{E}})$, we say that $\mathbf{G}$ *can be embedded in* $\bar{\mathbf{G}}$ if there exists an injective map $\eta : \mathbf{V} \to \bar{\mathbf{V}}$ so that $(\eta(\bar{u}), \eta(\bar{v})) \in \mathbf{E}$ whenever either $(u, v) \in \mathbf{E}$ or $(v, u) \in \mathbf{E}$. When use $\mathbf{G} \subset \bar{\mathbf{G}}$ to denote that $\mathbf{G}$ can be embedded in $\bar{\mathbf{G}}$. When $\mathbf{G} \subset \bar{\mathbf{G}}$, we also say that $\bar{\mathbf{G}}$ *embeds* $\mathbf{G}$. Figure 6 shows an example to make the idea clear. Essentially, the undirected version of $\mathbf{G}$ should be a "subgraph" of the undirected version of $\bar{\mathbf{G}}$ for $\mathbf{G}$ to be embedded in $\bar{\mathbf{G}}$. Note that edge directions play no role in embedding. The reason is that we use embedding to investigate effective resistance, which does not depend on the edge directions.
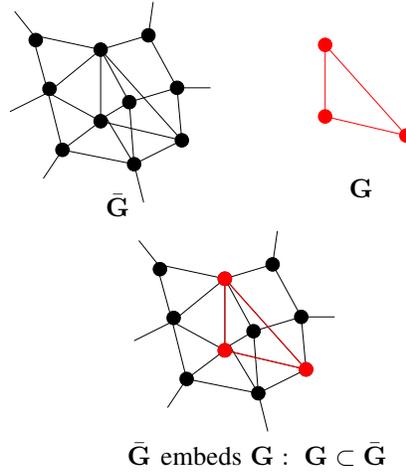

Fig. 6. A graph embedding example. The graph $\bar{\mathbf{G}}$ embeds $\mathbf{G}$, or equivalently, $\mathbf{G}$ can be embedded in $\bar{\mathbf{G}}$, or simply $\mathbf{G} \subset \bar{\mathbf{G}}$.

Theorem 7 (Generalized Rayleigh's Monotonicity Law). *Consider two generalized electrical networks* $(\mathbf{G}, R)$ *and* $(\mathbf{G}, \bar{R})$ *so that* $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ *can be embedded in* $\bar{\mathbf{G}} = (\bar{\mathbf{V}}, \bar{\mathbf{E}})$ *with an embedding* η *and* $R(e) \geq \bar{R}(\eta(e))$, $\forall e \in \mathbf{E}$, *where* $\eta(e)$ *denotes the edge in* $\bar{\mathbf{E}}$ *corresponding to* $e \in \mathbf{E}$. *Then the following monotonicity property holds for the generalized effective resistances in the two networks.*

$$R_{p,q}^{\text{eff}} \geq R_{\eta(p),\eta(q)}^{\text{eff}}, \quad \forall p, q \in \mathbf{V}.$$

Proof: Throughout the proof we will use $\bar{u} \in \bar{\mathbf{V}}$ to denote $\eta(u)$, the node in $\bar{\mathbf{G}}$ that corresponds to the node u in $\mathbf{G}$. Let i be the current in the electrical network $(\mathbf{G}, R^{-1})$, of intensity $\mathbf{i}$ from $p \in \mathbf{V}$ to $q \in \mathbf{V}$. Similarly, let $\bar{i}$ be the current for the electrical network $(\bar{\mathbf{G}}, \bar{R}^{-1})$ of the same intensity $\mathbf{i}$ from $\bar{p} \in \bar{\mathbf{V}}$ to $\bar{q} \in \bar{\mathbf{V}}$. Define a new flow $\hat{i}$ in the graph $\bar{\mathbf{G}}$ as

$$\hat{i}(\bar{e}) = \begin{cases} i(e) & e \in \mathbf{E} \\ 0 & \text{otherwise,} \end{cases}$$
$$\bar{e} \in \bar{\mathbf{E}}$$

where $e \in \mathbf{E}$ is the edge corresponding to $\bar{e} \in \bar{\mathbf{E}}$. Since $\hat{i}$ is a flow in the electrical network $(\bar{\mathbf{G}}, \bar{R}^{-1})$, we conclude from Thomson's minimum energy principle (Theorem 2) that

$$\sum_{(\bar{u},\bar{v}) \in \bar{\mathbf{E}}} Tr(\bar{i}_{\bar{u},\bar{v}}^T \bar{R}_{\bar{u},\bar{v}} \bar{i}_{\bar{u},\bar{v}}) \leq \sum_{(\bar{u},\bar{v}) \in \bar{\mathbf{E}}} Tr(\hat{i}_{\bar{u},\bar{v}}^T \bar{R}_{\bar{u},\bar{v}} \hat{i}_{\bar{u},\bar{v}}).$$

But since $\hat{i}$ is equal to i on $\mathbf{E}$ and zero outside this set, we further conclude that

$$\sum_{(u,v)\in\bar{\mathbf{E}}} Tr\left(\bar{i}_{u,v}^T \bar{R}_{u,v}\bar{i}_{u,v}\right) \leq \sum_{(u,v)\in\mathbf{E}} Tr\left(i_{u,v}^T \bar{R}_{u,v}i_{u,v}\right)$$
$$\leq \sum_{(u,v)\in\mathbf{E}} Tr\left(i_{u,v}^T R_{u,v}i_{u,v}\right).$$

From corollary 1, we conclude

$$Tr\left(\mathbf{i}^T \bar{R}_{\bar{u},\bar{v}}^{\mathrm{eff}}\mathbf{i}\right) \leq Tr\left(\mathbf{i}^T R_{u,v}^{\mathrm{eff}}\mathbf{i}\right).$$

Since this inequality holds for every intensity $\mathbf{i} \in \mathbb{R}^{k\times k'}$, the result follows[9].

B. Lattices, Fuzzes and Their Effective Resistances

Lattice graphs (cf. Figure 8) are familiar examples of regular degree graphs, that is, graphs in which every node has the same degree, and perhaps require no introduction. The effective resistance in lattices (with scalar resistors on every edge) have been studied in the literature, and we will show that similar results can be obtained for *generalized* lattice networks. Effective resistance in lattices, and in a class of graphs derived from them called lattice fuzzes, will be especially useful in studying the scaling laws of effective resistance in large graphs.

For any integer h, the *h- fuzz* $\mathbf{G}^{(h)}$ *of a graph* $\mathbf{G}$ is the graph obtained from $\mathbf{G}$ by adding an edge (u, v) whenever the graphical distance between the nodes u and v is less than or equal to h in $\mathbf{G}$ [5]. Recall that the *graphical distance* between two nodes u and v in a graph $\mathbf{G}$, denoted by $d_{\mathbf{G}}(u,v)$, is the minimum number of edges one has to traverse in going from one node to the other. In this definition, we allow edges to be traversed in any direction and therefore $d_{\mathbf{G}}(u,v) = d_{\mathbf{G}}(v,u)$. The directions of the "new" edges in $\mathbf{G}^{(h)}$ are arbitrary (see remark 1 in Section III). Figure 7 shows a graph and its 2-fuzz.
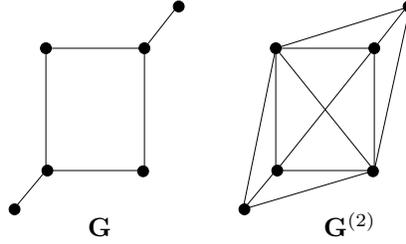

PSfrag replacements

$\mathbf{G}$ $\mathbf{G}^{(2)}$

Fig. 7.   A graph $\mathbf{G}$ and its 2-fuzz $\mathbf{G}^{(2)}$.

An h-fuzz will clearly have lower effective resistance than the original graph because of Rayleigh's Generalized Monotonicity Law, but it is lower only by a constant factor. The following result is a straightforward extension to the generalized case of a result about scalar effective resistance established by Doyle [33], that the effective resistance in a graph and its h-fuzz differs only by a constant factor.

Lemma 3. *Let* (G, R_o^{-1}) *be a generalized electrical network where* $G = (\mathbf{V}, \mathbf{E})$ *is bounded degree*[10] *weakly connected graph and each edge of G has equal generalized resistance* $R_o \in \mathbb{R}^{k\times k}$. *Let* $(G^{(h)}, R_o^{-1})$ *be the electrical network constructed by assigning a generalized resistance R_o to every edge of $G^{(h)}$, the h-fuzz of G. For every pair of nodes u and v in* $\mathbf{V}$,

$$\alpha R_{u,v}^{\mathrm{eff}}(\mathbf{G}) \leq R_{u,v}^{\mathrm{eff}}(\mathbf{G}^{(h)}) \leq R_{u,v}^{\mathrm{eff}}(\mathbf{G}),$$

where $R_{u,v}^{\mathrm{eff}}(\cdot)$ is the effective resistance in the electrical network $(\cdot, R^{-1})$ and $\alpha \in (0,1]$ is a positive constant that does not depend on u and v. $\square$

The arguments used in the proof by Doyle [33] can be applied to the generalized case to get the desired result. The proof is included in Appendix I.

A *d-D lattice in* $\mathbb{R}^d$ is a graph that has a vertex at every point in $\mathbb{R}^d$ with integer coordinates and an edge between every two vertices with an Euclidean distance of 1 between them. Edge directions are arbitrary. We denote the d-D lattice in $\mathbb{R}_d$ by $\mathbf{Z}_d$. Figure 8 shows 1-, 2-, and 3-dimensional lattices. Lattices have infinitely many nodes and edges, and are therefore examples of *infinite graphs*. In practice, they serve as proxies for very large graphs.

---

[9]*Draft note: it seems from the proof that the stronger result $R_{\bar{u},\bar{v}}^{\mathrm{eff}} \prec R_{u,v}^{\mathrm{eff}}$ is also true. needs further checking.*

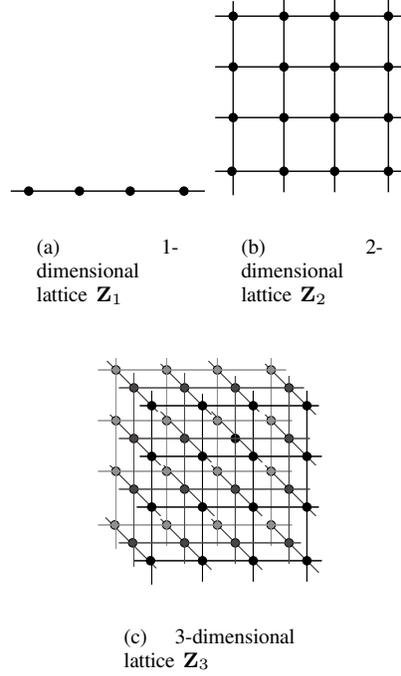[10]The maximum number of edges incident on any node of the graph is bounded.

Fig. 8. Lattices

The following lemma establishes the effective resistance of a d-D lattice and its h-fuzz.

Lemma 4. *Consider the electrical network $(\mathbf{Z}_d^{(h)}, R_o^{-1})$ with a constant generalized resistance $R_o \in \mathbb{R}^{k \times k}$ on every edge of the h-fuzz of the d-D lattice, where h is a positive integer. We denote by $d_{\mathbf{Z}_d}(u, v)$ the graphical distance between two nodes u and v in the d-D lattice $\mathbf{Z}_d$. The effective resistance $R_{u,v}^{\mathrm{eff}}$ between two nodes u and v in the electrical network $(\mathbf{Z}_d^{(h)}, R_o^{-1})$ satisfies*

1) $R_{u,v}^{\mathrm{eff}}(\mathbf{Z}_1^{(h)}) = \Theta\big(d_{\mathbf{Z}_1}(u, v)\big)$
2) $R_{u,v}^{\mathrm{eff}}(\mathbf{Z}_2^{(h)}) = \Theta\big(\log d_{\mathbf{Z}_2}(u, v)\big)$,
3) $R_{u,v}^{\mathrm{eff}}(\mathbf{Z}_3^{(h)}) = \Theta\big(1\big)$. $\hfill\square$

The usual asymptotic notation $\Theta(\cdot)$ is used with matrix valued functions in the following way. First of all, for two matrices $A, B \in \mathbb{R}^{k \times k}$, $A \leq (<)B$ means $A - B$ is negative semidefinite (definite). For two functions $g : \mathbb{R} \to \mathbb{R}^{k \times k}$ and $f : \mathbb{R} \to \mathbb{R}$, the notation $g(x) = \Theta(f(x))$ means there exists a constant x_o and two positive definite matrices A and B such that $Af(x) \leq g(x) \leq Bf(x)$ for all $x > x_o$.

The proof of Lemma 4 follows by generalizing into the matrix-resistance case known results about effective resistance in "regular" lattice electrical networks with 1-Ohm resistors. In the sequel we use the notation $g(x) = \Theta(f(x))$ both in the traditional sense and in the matrix-valued sense described above, dimension of $g(x)$ determining in what sense it is used.

Proof: [Proof of Lemma 4] Consider the "regular" electrical network $(\mathbf{Z}_d, 1)$ formed by assigning a 1-Ohm resistance to every edge of the d-D lattice $\mathbf{Z}_d$. The effective resistance between two nodes in the one-dimensional lattice network $(\mathbf{Z}_1, 1)$ is given by $r_{u,v}^{\mathrm{eff}} = d_{\mathbf{Z}_1}(u, v)$, which follows from series resistance formula. From this, Lemma 2 and Lemma 3, the first statement follows.

It was shown in [34] that for the regular electrical network $(\mathbf{Z}_2, 1)$, the effective resistances obeys $r_{u,v}^{\mathrm{eff}} = \Theta\left(\log d_{\mathbf{Z}_2}(u, v)\right)$. The second statement follows from the application of Lemma 2 and Lemma 3 to this result.

For $d = 3$, it is known that for the regular electrical network $(\mathbf{Z}_3, 1)$, the effective resistances obeys $r_{u,v}^{\mathrm{eff}} = \Theta(1)$ [35]. Again, the third statement follows from the application of Lemma 2 and Lemma 3 to this result. $\hfill\blacksquare$

The fact that in a 1-dimensional lattice the effective resistance grows linearly with the distance between nodes can be trivially deduced from the well known formula for the effective resistance of a series of resistors (which generalizes to generalized electrical networks). In two-dimensional lattices the effective resistance only grows with the logarithm of the graphical distance and therefore the effective resistance grows very slowly with the distance between nodes. Far more

surprising is the fact that in three-dimensional latices the effective resistance is actually bounded by a constant even when the distance is arbitrarily large.

VI. Scaling of BLUE Covariance with Distance: Dense and Sparse Graphs

In this section we show how to combine the tools developed so far to determine the scaling laws of the estimation error variances for general classes of measurement graphs and investigate what graph properties determine this scaling. Roughly speaking, our approach is the following: we determine what structural properties must a graph satisfy so that it can either embed, or be embedded in a lattice or its h-fuzz. These conditions come from looking at different "drawings" of the graph. When a graph can be embedded in a lattice, Rayleigh's Generalized Monotonicity Law gives us a lower bound on the generalized effective resistance of the graph in terms of the effective resistance in the lattice, which we already know as a function of distance from the Lattice Effective Resistance Lemma 4. We still need to relate the distance in the graph to the distance in the corresponding lattice to get a complete answer to the scaling law, which are again obtain by considering the "drawing" of the graph. When a graph embeds a lattice, we get an upper bound.

Before we go into describing these concepts precisely, we might ask ourselves if there are no simple indicators of the relationship between graph structure and estimator accuracy that might be used to answer the question of variance scaling without going into the somewhat complex route we have outlined above. In fact, in the sensor networks literature, it is recognized that higher "density" of nodes/edges usually lead to better estimation accuracy. Usually, the average number of nodes in an unit of area or the average degree of a node (that is, its number of neighbors) are used to quantify the notion of denseness [3, 4]. We have already seen that when one graph can be embedded in another, the one with the higher number of edges have a lower effective resistance, and consequently, lower estimator variance. One could therefore expect that a higher density of edges and nodes in the measurement graph should lead to better estimates. However, "naive" measures of node and edge density turn out to be very misleading predictors for how the *estimation error variance scales with distance*. We now present a few examples to motivate the search for deeper graph-structural properties that determine how variance scales with distance.

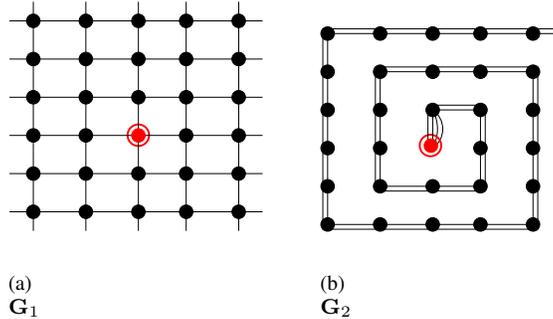### A. Counterexamples to conventional wisdom

Fig. 9. Two measurement graphs drawn in 2-dimensional with the same number of nodes and edges per unit area and with the same node degree of 4 (except for a single node of $\mathbf{G}_2$ that has a higher degree of 6).

Consider the two graphs $\mathbf{G}_1$ and $\mathbf{G}_2$ in Figure 9 and imagine that these two graphs represent the measurement graphs of two sensor networks deployed in 2-d Euclidean space, and that we drew the graphs so that the position of the vertices in the paper matches precisely the physical locations of the sensors in the plane where they were deploys (perhaps up to some scaling). For simplicity we will assume that all measurements affected by noise with the same covariance matrix P. For both graphs the number of nodes per unit area is exactly the same and the degree of every node in both the graphs is four (except for a single node of $\mathbf{G}_2$ that has a higher degree of 6). So "naive" notions of denseness would classify the graphs as similar.

The graph $\mathbf{G}_1$ is a 2-dimensional lattice so from the Electrical Analogy Theorem 4 and the Lattice Effective Resistance Lemma 4, we conclude that the estimation error variance grows asymptotically with the logarithm of the graphical distance from the node to the reference. The exact same statement is true if we replace "graphical distance" by "Euclidean distance" because this drawing of the graph has the property that the graphical distance is always between the Euclidean distance and $\sqrt{2}$ times the Euclidean distance.
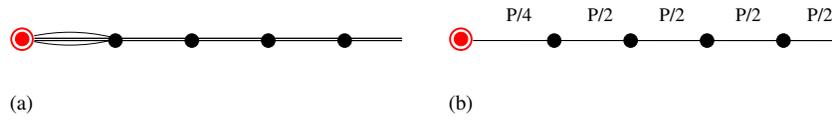
Fig. 10. (a)A 1-dimensional drawing of the graph G_2 in Figure 9, and (b)another equivalent drawing, with single edges between nodes.

To study the estimation error variance of $\mathbf{G}_2$, we redraw this graph as in Figure 10(a). It is now easy to recognize that $\mathbf{G}_2$ is also equivalent to an 1-dimensional lattice shown in Figure 10(b), with the noise covariance matrix equal to $P/2$ at every edge, except the first one that for which the covariance matrix is $P/4$. This equivalence follows from the fat that the parallel of n equal matrix resistors R has an effective resistance of $R/4$, just as it is so in case of regular resistors. This means that for $\mathbf{G}_2$ the estimation error variance grows asymptotically with the logarithm of the graphical distance from the node to the reference. Moreover, for the drawing of $\mathbf{G}_2$ in Figure 10(a), the exact same statement is true if we replace "graphical distance" by "Euclidean distance." However, for the drawing of $\mathbf{G}_2$ in Figure 9, the Euclidean distance provides little information about the effective resistance because nodes separated by small Euclidean distances can be separated by very large graphical distances (as one moves further away from the center node). Even though $\mathbf{G}_1$ and $\mathbf{G}_2$ are very similar by all conventional measures of denseness, the effective resistance scales very differently with distance in these two graphs. The difference between the graphs became apparent when we *drew* them appropriately.

Two conclusions can be drawn from this example:
1) Two graphs with the same node degree and node/edge densities can exhibit fundamentally different scaling laws of error variance with distance.
2) Some drawings of the graph appear to be more adequate than others to determine error variance scaling laws.

(a) A triangular 2-dimensional lattice   (b) A 3-fuzz of a 1-dimensional lattice
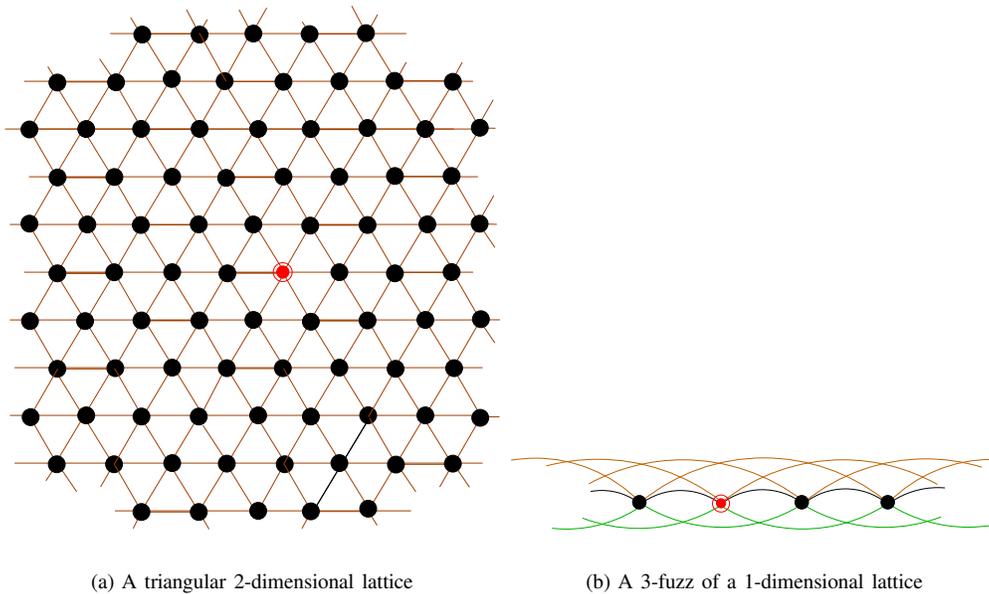
Fig. 11. Two different measurement graphs — a triangular 2-dimensional lattice and a 3-fuzz of a 1-dimensional lattice. Both graphs have the same node degree for every node but very different variance growth rates with distance.

The two graphs in Figure 11 – a triangular lattice and a 3-fuzz of a 1-dimensional lattice – could have been used in place of the ones in Figure 9 to make these points. Actually, things are perhaps even more interesting for these graphs: The effective resistance in the 3-fuzz of the 1-dimensional lattice grows linearly with distance, whereas in the triangular lattice it grows only with the logarithm of distance, in spite of both graphs having the *same node degree of* 6. At this point the reader will have to take our word for the stated growth of the effective resistance in triangular lattices, but the tools needed to prove this fact are forthcoming.

B. Drawings of graphs

In graph theory, the branch of mathematics dealing with the study of graphs, a graph is treated purely as a collection of nodes connected by edges, without any regard to the geometry determined by the nodes' locations. However, in sensor

network problems there is an underlying geometry for the measurement graph because generally this graph is tightly related to the physical locations of the sensor nodes. For example, a pair of nodes from a sensor network will typically have an edge if the two nodes are within some "sensing/communication range" of each other. In general, this range may be defined in a very complex fashion (not just determined by the Euclidean distance), but still the geometric configuration of nodes in Euclidean space will play a key role in determining the measurement graph.

A

B
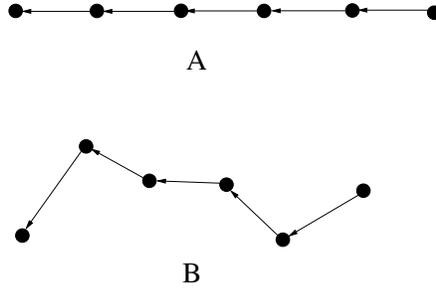
Fig. 12.   Two different drawings of the same graph.

Graph drawings are used to capture the geometry of graphs in Euclidean space. Figure 12 shows two different "pictures" of the same graph. From a graph theoretic point of view, the two graphs are the same because they have the same nodes and edges. However, the two graphs were "drawn" differently. A drawing of a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ is simply a mapping of its nodes to points in some Euclidean space, which can formally be described by a function $f : \mathbf{V} \to \mathbb{R}^d$, $d \geq 1$. An edge (u, v) is shown in the drawing by a line segment from u to v[11]. A drawing is also sometimes called a *representation* of a graph [27].

For a particular drawing f of a graph, we can define an Euclidean distance between nodes, which is simply the distance in that Euclidean space between the drawings of the nodes. In particular, given two nodes $u, v \in \mathbf{V}$ the *Euclidean distance between u and v induced by the drawing $f : \mathbf{V} \to \mathbb{R}^d$* is defined by

$$d_f(u, v) := \|f(v) - f(u)\|,$$

where $\| \cdot \|$ denoted the usual Euclidean norm in d-space. Note that Euclidean distances depend on the drawing and can be completely different from graphical distances. The two drawings of the same graph found in Figure 9(b) and Figure 10(a) should make this point quite clear. It is important to emphasize that the definition of drawing does not require edges to not intersect and therefore every graph has a drawing in any Euclidean space. In fact, it has infinitely many drawings.

For a sensor network, there is a *natural drawing* of its measurement graph that is obtained by associating each node to its position in 1-, 2- or 3-dimensional Euclidean space. In reality, all sensor networks are situated in 3-dimensional space. However, sometimes it maybe more natural to draw them on a 2-dimensional Euclidean space if one dimension (for example, height) does not vary much from node to node, or is somehow irrelevant. In yet another situation, such as the one shown in Figure 1, one could draw the graph in 1-dimension since the nodes essential form a chain even though nodes are situated in 3-dimensional space. For *natural drawings, the Euclidean distance induced by the drawing is, in general, a much more meaningful notion of distance than the graphical distance.* In this article we will see that the Euclidean distance induced by appropriate drawings provide the right measure of distance to determine scaling laws of error variances.

C. Measures of denseness/sparseness

For a particular drawing f and induced Euclidean distance d_f of a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, four parameters can be used to characterize denseness/sparseness. The term *minimum node distance* denotes the minimum Euclidean distance between the drawing of two nodes:

$$s := \inf_{\substack{u,v \in \mathbf{V} \\ v \neq u}} d_f(u, v).$$

The term *maximum connected range* denotes the Euclidean length of the drawing of the longest edge:

$$r := \sup_{(u,v) \in \mathbf{E}} d_f(u, v).$$

---

[11] *Draft note: Should we say "line segment with an arrow from u to v"?*
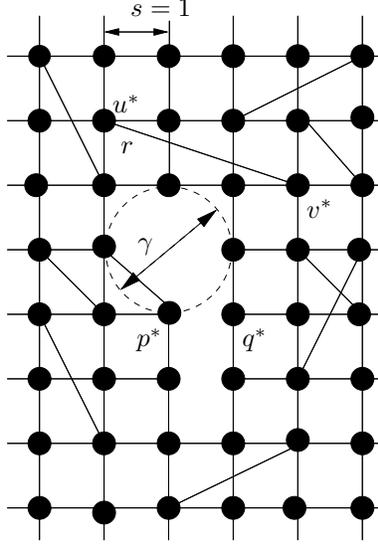
Fig. 13. Drawing of a graph and the corresponding denseness/sparseness parameters: $s = 1$, $r = d_f(u^*, v^*) = \sqrt{10}$, $\gamma = 2$, and $\rho = d_f(p^*, q^*)/d_{\mathbf{G}}(p^*, q^*) = 1/5$.

The term *maximum uncovered diameter* denotes the diameter of the largest open ball that can be placed in $\mathbb{R}^d$ with no drawing of a node inside it:

$$\gamma := \sup \left\{ \delta : \exists B_\delta \text{ such that } f(u) \notin B_\delta, \ \forall u \in \mathbf{V} \right\},$$

where the existential quantification spans over the balls B_δ in $\mathbb{R}^d$ with diameter δ. Finally, the term *asymptotic distance scaling* denotes the largest asymptotic ratio between the graphical and the Euclidean distance between two nodes: as

$$\rho := \lim_{n \to \infty} \inf \left\{ \frac{d_f(u, v)}{d_{\mathbf{G}}(u, v)} : u, v \in \mathbf{V} \text{ and } d_{\mathbf{G}}(u, v) \geq n \right\}.$$

Essentially ρ provides a lower bound for the ratio between the Euclidean and the graphical distance for nodes that are very far apart. Figure 13 shows the drawing of a graph and the four corresponding parameters s, r, γ, and ρ.

1) Dense graphs: The drawing of a graph for which the maximum uncovered diameter is finite ($\gamma < \infty$) and the asymptotic distance scaling is positive ($\rho > 0$) is called a *dense drawing*. We say that a $\mathbf{G}$ *is dense in* $\mathbb{R}^d$ if there exists a dense drawing of the graph in $\mathbb{R}^d$. Intuitively, these drawing are "dense" in the sense that the nodes can cover $\mathbb{R}^d$ without leaving large holes between them and still having sufficiently many edges so that a small Euclidean distance between two nodes in the drawing guarantees a small graphical distance between them. In particular, for dense drawings there are always finite constants α, β for which

$$d_{\mathbf{G}}(u, v) \leq \alpha \, d_f(u, v) + \beta, \quad \forall u, v \in \mathbf{V}. \tag{35}$$

This fact is proved in Proposition 4. Using the natural drawing of a d-dimensional lattice, one concludes that this graph is dense in $\mathbb{R}^d$. One can also show that a d-dimensional lattice can never be dense in $\mathbb{R}^{\bar{d}}$ with $\bar{d} > d$. This means, for example, that any drawing of a 2-dimensional lattice in the 3-dimensional Euclidean space will never be dense.

Proposition 4 (ρ vs. linear growth). *The following two statements are equivalent:*

1) *The asymptotic distance scaling*

$$\rho := \lim_{n \to \infty} \inf \left\{ \frac{d_f(u, v)}{d_{\mathbf{G}}(u, v)} : u, v \in \mathbf{V} \text{ and } d_{\mathbf{G}}(u, v) \geq n \right\}.$$

is strictly positive.

2) *There exist finite constants α, β for which*

$$d_{\mathbf{G}}(u, v) \leq \alpha d_f(u, v) + \beta, \quad \forall u, v \in \mathbf{V}.$$

Proof: See appendix II.

2) Sparse graphs: Graph drawings for which the minimum node distance is positive ($s > 0$) and the maximum connected range is finite ($r < \infty$) are called *civilized drawings*. This definition is essentially a refinement of the one given in [5], with the quantities r and s made to assume precise values. Intuitively, these drawings are "sparse" in the sense that one can keep the edges with finite lengths, without cramping all nodes on top of each other. We say that a graph **G** *is sparse in* $\mathbb{R}^d$ if *it can be drawn in a civilized manner in* d-*dimensional Euclidean space*. For example, we can conclude from the natural drawing of a d-dimensional lattice that this graph is sparse in $\mathbb{R}^d$. In fact, any h-fuzz of a d-dimensional lattice is still sparse in $\mathbb{R}^d$. However, a d-dimensional lattice can never be drawn in a civilized way in $\mathbb{R}^{\bar{d}}$ with $\bar{d} < d$. This means, for example, that any drawing of a 3-dimensional lattice in the 2-dimensional Euclidean space will never be a civilized drawing. A 3-dimensional lattice is therefore not sparse in $\mathbb{R}^2$.

The notions of graph "sparseness" and "denseness" are mostly interesting for infinite graph because every finite graph is sparse in all Euclidean spaces $\mathbb{R}^d$, $\forall d \geq 1$ and no finite graph can ever be dense in any Euclidean space $\mathbb{R}^d$, $\forall d \geq 1$. This is because any drawing of a finite graph that does not place nodes on top of each other will necessarily have a positive minimum node distance and a finite maximum connected range (from which sparseness follows) and it is not possible to achieve a finite maximum uncovered diameter with a finite number of nodes (from which lack of denseness follows). However, in practice infinite graphs serve as proxies for very large graphs that, from the perspective of most nodes, "appear to extend in all directions as far as the eye can see." So conclusions drawn for sparse/dense infinite graphs hold for large graphs, at least far from the graph boundaries.

3) Sparseness, denseness, and embeddings: The notions of sparseness and denseness introduced above are useful because they provide a complete characterization for the classes of graphs that can embed or be embedded in lattices. This is presented in the next two theorems. Once we know whether a graph can embed or be embedded in a lattice (or its fuzz), we can use the Lattice Effective Resistance Lemma 4 to determine an upper or lower bound on the scaling laws for the effective resistance in the graph.

Theorem 8 (Lattice Embedding). *Let* $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ *be a graph that does not have multiple edges between nodes.* **G** *is sparse in* $\mathbb{R}^d$ *if and only if* **G** *can be embedded in an* h-*fuzz of a* d-*dimensional lattice. Formally,*

$$\mathbf{G} \text{ is sparse in } \mathbb{R}^d \quad \Leftrightarrow \quad \exists h < \infty : \mathbf{G} \subset \mathbf{Z}_d^{(h)}$$

Moreover, if $f : \mathbf{V} \to \mathbb{R}^d$ *is a civilized drawing of* **G** *in* $\mathbb{R}^d$ *then,* $\forall u, v \in \mathbf{V}$,

$$d_{\mathbf{Z}_d}(u_z, v_z) \geq \sqrt{d}\left(\frac{1}{s} d_f(u, v) - 2\right), \tag{36}$$

where u_z *is the node in* $\mathbf{Z}_d^{(h)}$ *that corresponds to* $u \in \mathbf{V}$ *when* $\mathbf{G} \subset \mathbf{Z}_d^{(h)}$, $d_{\mathbf{Z}_d}$ *denotes the graphical distance in* $\mathbf{Z}_d$ *and* d_f *denotes the Euclidean distance in* $\mathbb{R}^d$ *induced by* f.

This first statement is essentially taken from [5], where it was proved that if a graph without multiple edges between node pairs can be drawn in a civilized manner in $\mathbb{R}^d$, then it can be embedded in a h-fuzz of a d-lattice, where h depends only on s and r. A careful examination of the proof reveals that it is not only sufficient but also a necessary condition for embedding in lattice fuzzes.

Proof: [Proof of Theorem 8] We will denote by $g : \mathbf{V}_{Z_d} \to \mathbb{R}^d$ the natural drawing of the lattice $\mathbf{Z}_d$.

$\Rightarrow$ Since **G** is sparse in $\mathbb{R}^d$, there is a drawing function $f : \mathbf{V} \to \mathbb{R}^d$ that produces a civilized drawing of **G** with minimum node distance $s > 0$ and maximum connected range $r < \infty$. Consider a new drawing $f' : \mathbf{V} \to \mathbb{R}^d$ of **G** defined as

$$f'(u) = \frac{\sqrt{d}}{s} f(u), \forall u \in \mathbf{V}. \tag{37}$$

The minimum node distance and the maximum connected range in this drawing are

$$s' = \sqrt{d}, \quad r' = \frac{\sqrt{d}}{s} r.$$

Superimpose the two drawings $g(\mathbf{Z}_d)$ and $f'(\mathbf{G})$ (cf. figure 14). In every lattice cell[12] in $\mathbb{R}^d$, there is at most one node of **G**, for if there are two, then in an open ball of diameter $\sqrt{d}$ in $\mathbb{R}^d$, there are two points $f'(u)$ and $f'(v)$ where $u, v \in \mathbf{V}$, which violates the condition that

$$s' = \inf_{\substack{u \neq v \\ u, v \in \mathbf{V}}} \|f'(u) - f'(v)\| = \sqrt{d}.$$

---

[12]A lattice cell is taken as a unit semi open hypercube in $\mathbb{R}^d$, which is a subset of $\mathbb{R}^d$ of the form $[a_1, a_1 + 1) \times [a_2, a_2 + 1) \cdots \times [a_d, a_d + 1)$ for some $a_1, a_2, \ldots, a_d \in \mathbb{R}$.
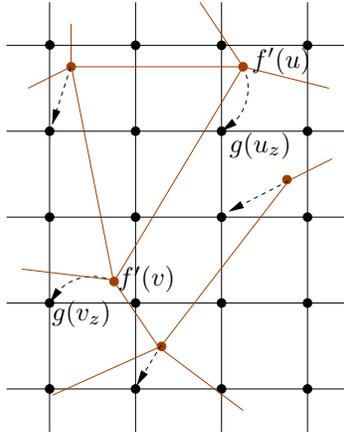
Fig. 14. A graph that can be drawn in a civilized manner in $\mathbb{R}^2$ can be embedded in a fuzz of the 2-D lattice.

Define a mapping $\eta : \mathbf{V} \to \mathbf{V}_{\mathbf{Z}_d}$ by associating every node $u \in \mathbf{V}$ to the lattice vertex with the most negative coordinate in the lattice cell where $f'(u)$ lies (cf. Figure 14). Since a lattice cell contains the f' drawing at most one node, η is injective. Let $u_z := \eta(u)$ and $v_z := \eta(v)$. Since $f'(u)$ and $g(u_z)$ lie in the same lattice cell,

$$\|f'(u) - g(u_z)\| \leq \sqrt{d}, \quad \forall u \in \mathbf{V}. \tag{38}$$

So,

$$
\begin{aligned}
d_{\mathbf{Z}_d}(u_z, v_z) &= \|g(u_z) - g(v_z)\|_1 \\
&\leq \sqrt{d}\|g(u_z) - g(v_z)\| \\
&\leq \sqrt{d}\Big(\|g(u_z) - f'(u)\| + \|f'(u) - f'(v)\| + \|f'(v) - g(v_z)\|\Big) \\
&\leq \sqrt{d}\Big(2\sqrt{d} + d_{f'}(u, v)\Big)
\end{aligned}
$$

If $(u, v) \in \mathbf{E}$, $d_{f'}(u, v) \leq r' = \frac{\sqrt{d}}{s}r$, where r' is the maximum connected range in the f' drawing of $\mathbf{G}$, so we get

$$d_{\mathbf{Z}_d}(u_z, v_z) \leq 2d + \frac{d}{s}r < \infty$$

Defining $h := \lceil 2d + \frac{d}{s}r \rceil$, we see that there is an edge between u_z and v_z, i.e., between $\eta(u)$ and $\eta(v)$, in $\mathbf{Z}_d^{(h)}$. This proves that η is an embedding, so $\mathbf{G} \subset \mathbf{Z}_d^{(h)}$.

$\Leftarrow$ Since $\mathbf{G} \subset \mathbf{Z}_d^{(h)}$ for some $h < \infty$, there is an embedding $\eta : \mathbf{V} \to \mathbf{V}_{\mathbf{Z}_d}$ of $\mathbf{G}$ in $\mathbf{Z}_d$. Consider a drawing f of $\mathbf{G}$ defined as

$$f(u) = g(\eta(u)), \qquad \forall u \in \mathbf{V}.$$

We immediately get $s \geq 1 > 0$ for this drawing. If $(u, v) \in \mathbf{E}$,

$$
\begin{aligned}
\|f(u) - f(v)\| &= \|g(\eta(u)) - g(\eta(v))\| \\
&\leq \|g(\eta(u)) - g(\eta(v))\|_1 \\
&= d_{\mathbf{Z}_d}(\eta(u), \eta(v)) \leq h,
\end{aligned}
$$

where the last inequality follows because $\mathbf{G} \subset \mathbf{Z}_d^{(h)}$. Therefore the maximum connected range r in the drawing f of $\mathbf{G}$ satisfies $r \leq h < \infty$. The drawing of $\mathbf{G}$ specified by f is therefore a civilized drawing in $\mathbb{R}^d$, from which it follows that $\mathbf{G}$ is sparse in $\mathbb{R}^d$.

To prove the relation (36), we go back to the drawing f' defined in (37) based on the civilized drawing f of $\mathbf{G}$. Since $\mathbf{G}$ is sparse in $\mathbb{R}^d$, $\mathbf{G} \subset \mathbf{Z}_d^{(h)}$ for some positive integer h. Let $\eta : \mathbf{V} \to \mathbf{V}_{\mathbf{Z}_d}$ be the embedding of $\mathbf{G}$ into $\mathbf{Z}_d^{(h)}$. Denoting $u_z := \eta(u)$, for any $u, v \in \mathbf{V}$ we get

$$
\begin{aligned}
d_{\mathbf{Z}_d}(u_z, v_z) &= \|g(u_z) - g(v_z)\|_1 \\
&\geq \|g(u_z) - g(v_z)\|.
\end{aligned}
$$

Since $\|f'(u) - f'(v)\| \leq \|f'(u) - g(u_z)\| + \|g(u_z) - g(v_z)\| + \|g(v_z) - f'(v)\|$, we get from the above that

$$d_{\mathbf{Z}_d}(u_z, v_z) \geq \|f'(u) - f'(v)\| - \|f'(u) - g(u_z)\| - \|g(v_z) - f'(v)\|$$
$$\geq d_{f'}(u, v) - 2\sqrt{d}.$$

where we have used that fact that both $f'(u)$ and $g(u_z)$ lie in the same lattice cell for every $u \in \mathbf{V}$. Since $d_{f'}(\cdot) = d_f(\cdot)\frac{\sqrt{d}}{s}$, the result follows. ∎

The next theorem shows that fuzzes of dense graphs can embed lattices. A sufficient condition for a graph to embed a lattice-fuzz was previously published in [6]; here we tighten those results and derive necessary and sufficient conditions.

Theorem 9 (Lattice Embedding). $\mathbf{G}$ *is dense in* $\mathbb{R}^d$ *if and only if (i) the* d-*dimensional lattice can be embedded in an* h-*fuzz of* $\mathbf{G}$ *for some positive integer* h *and (ii) every node of* $\mathbf{G}$ *that is not mapped to a node of* $\mathbf{Z}_d$ *is at a uniformly bounded graphical distance from a node that is mapped to* $\mathbf{Z}_d$. *Formally,*

$$\mathbf{G} \text{ is dense in } \mathbb{R}^d \quad \Leftrightarrow \quad \exists h, c < \infty :$$
$$(i)\mathbf{G}^{(h)} \supset \mathbf{Z}_d \quad \&$$
$$(ii)\forall u \in \mathbf{V} \; \exists \bar{u} \in \mathbf{V}_{\text{lat}}(\mathbf{G}) : d_{\mathbf{G}}(u, \bar{u}) \leq c,$$

where $\mathbf{V}_{\text{lat}}(\mathbf{G})$ *denotes the set of nodes of* $\mathbf{G}$ *that are mapped to nodes of* $\mathbf{Z}_d$.

Moreover, if $f : \mathbf{V} \to \mathbb{R}^d$ *is a dense drawing of* $\mathbf{G}$ *in* $\mathbb{R}^d$, *then* $\forall u, v \in \mathbf{V}$, *there exists* $u_z, v_z \in \mathbf{V}_{\mathbf{Z}_d}$, *a positive constant* c_1, *and an embedding* $\eta : \mathbf{V}_{\mathbf{Z}_d} \to \mathbf{V}$ *of* $\mathbf{Z}_d$ *into* $\mathbf{G}^{(h)}$ *such that*

$$d_{\mathbf{G}}(u, \eta(u_z)) \leq c_1, \tag{39}$$
$$d_{\mathbf{G}}(v, \eta(v_z)) \leq c_1, \tag{40}$$
$$d_{\mathbf{Z}_d}(u_z, v_z) \leq 4d + \frac{\sqrt{d}}{\gamma} d_f(u, v), \tag{41}$$

where γ *is the maximum uncovered diameter of the* f-*drawing of* $\mathbf{G}$.

Proof: [Proof of Theorem 9] As in the previous theorem, $g : \mathbf{V}_{\mathbf{Z}_d} \to \mathbb{R}^d$ denotes the natural drawing of the lattice $\mathbf{Z}_d$ in $\mathbb{R}^d$.

$\Rightarrow$ Since $\mathbf{G}$ is dense in $\mathbb{R}^d$, there is a drawing function $f : \mathbf{V} \to \mathbb{R}^d$ so that the f-drawing of $\mathbf{G}$ has a $\gamma < \infty$ and $\rho > 0$. Define a new drawing $f' : \mathbf{V} \to \mathbb{R}^d$ as

$$f'(u) = \frac{1}{\gamma} f(u), \quad \forall u \in \mathbf{V},$$

so that the maximum uncovered diameter γ' of the f' drawing of $\mathbf{G}$ is 1. Note that f' is still a dense drawing of $\mathbf{G}$. Now we superimpose the natural g-drawing of $\mathbf{Z}_d$ on the f'-drawing of $\mathbf{G}$, and draw open balls of diameter one $B(g(u_z), \frac{1}{2})$ centered at the natural drawing $g(u_z)$ of every lattice node. Figure 15 shows an example in $\mathbb{R}^2$. Since $\gamma' = 1$, it follows from the definition of denseness that in every one of those balls, there is at least one node $u \in \mathbf{V}$. To construct the
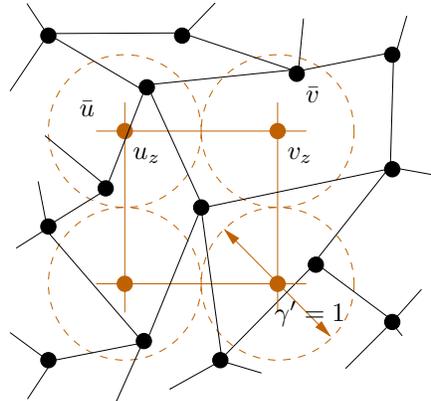

Fig. 15. Superimposing a 2-dimensional lattice (brown) on a 2-dimensional dense graph (black).

embedding, we associate each node of the lattice to a node of $\mathbf{G}$ whose drawing appears inside the ball centered around the lattice node. This defines a injective function $\eta : \mathbf{V}_{\mathbf{Z}_d} \to \mathbf{V}$. Consider two adjacent nodes of the lattice $u_z, v_z \in \mathbf{V}_{\mathbf{Z}_d}$,

i.e., there is an edge between u_z and v_z in $\mathbf{Z}_d$. Let $\bar{u} := \eta(u_z), \bar{v} := \eta(v_z)$. Since $f'(\bar{u})$ and $f'(\bar{v})$ belong to adjacent balls of unit diameter (cf. Figure 15),

$$d_{f'}(\bar{u}, \bar{v}) = \|f'(\bar{u}) - f'(\bar{v})\| \leq 2.$$

From Proposition 4 and f' being a dense drawing in $\mathbb{R}^d$, it follows that $d_{\mathbf{G}}(\bar{u}, \bar{v}) \leq 2\alpha + \beta$, for some positive constants α and β. Define $h := \lceil 2\alpha + \beta \rceil$. Then $\bar{u}$ and $\bar{v}$ will have an edge between them in the h-fuzz $\mathbf{G}^{(h)}$. So $\mathbf{G}^{(h)} \supset \mathbf{Z}_d$, and we have (i).

By definition, $\mathbf{V}_{\mathrm{lat}}(\mathbf{G}) = \eta(\mathbf{V}_{\mathbf{Z}_d})$. If $u \in \eta(\mathbf{V}_{\mathbf{Z}_d})$, then (ii) is trivially true (choose $\bar{u} := u$), so only nodes in $\mathbf{V} \setminus \eta(\mathbf{V}_{\mathbf{Z}_d})$ are interesting. For every $u \in \mathbf{V}$, find $u_z \in \mathbf{V}_{\mathbf{Z}_d}$ as

$$u_z = \arg \min_{\substack{p \in B(u_z', 1/2) \\ u_z' \in \mathbf{V}_{\mathbf{Z}_d}}} \|f'(u) - p\|$$

There are only 2^d balls one needs to check to determine the minimum, so u_z exists, though it may not be unique. If there are multiple minima, pick any one. This procedure defines an onto map $\xi : \mathbf{V} \to \mathbf{V}_{\mathbf{Z}_d}$. Let $\eta : \mathbf{V}_{\mathbf{Z}_d} \to \mathbf{V}$ be the embedding of $\mathbf{Z}_d$ into $\mathbf{G}^{(h)}$ as described earlier in this proof. Define $\psi : \mathbf{V} \to \mathbf{V}$ as $:= (\eta \circ \xi)$. Recall that since f' is a dense drawing with $\gamma' = 1$, every ball in $\mathbb{R}^d$ of unit diameter contains the drawing $f'(u)$ for at least one node $u \in \mathbf{V}$. Now

$$d_{f'}(u, \psi(u)) \leq \|f'(u) - g(u_z)\| + \|g(u_z) - f'(\psi(u))\|$$
$$< \sqrt{d} + \frac{1}{2} \leq \frac{3}{2}\sqrt{d}, \tag{42}$$

where we have used the fact that $f'(u)$ either lies in the ball centered at $g(u_z)$ or in the gaps between that ball and the neighboring balls, and $f'(\psi(u)) \in B(g(u_z), \frac{1}{2})$. From Proposition 4 and the denseness of the f-drawing of $\mathbf{G}$, we get

$$d_{\mathbf{G}}(u, \psi(u)) \leq \alpha d_f(u, \psi(u)) + \beta \tag{43}$$
$$= \alpha\gamma d_{f'}(u, \psi(u)) + \beta \tag{44}$$
$$< \frac{3}{2}\alpha\gamma\sqrt{d} + \beta. \tag{45}$$

Define $c := \lceil \frac{3}{2}\alpha\gamma\sqrt{d} + \beta \rceil$. Then for every $u \in \mathbf{V}$, there exists a $\bar{u} := \psi(u) \in \eta(\mathbf{V}_{\mathbf{Z}_d}) \subset \mathbf{V}$ such that $d_{\mathbf{G}}(u, \bar{u}) < c$, which is the desired condition (ii).

$\Leftarrow$ We have to prove that if (i) and (ii) are satisfied, then $\mathbf{G}$ is dense in $\mathbb{R}^d$. Since $\mathbf{G}^{(h)} \supset \mathbf{Z}$, there is an injective map $\eta : \mathbf{V}_{\mathbf{Z}_d} \to \mathbf{V}$ such that $\mathbf{V} \supset \eta(\mathbf{V}_{\mathbf{Z}_d})$. We will construct a drawing f of $\mathbf{G}$ in $\mathbb{R}^d$ with the following procedure and then prove that it is a dense drawing. Pick a node u in $\mathbf{V}$ that has not been drawn yet. By (ii), there exists a positive constant c and two nodes $\bar{u} \in \mathbf{V}$ and $u_z \in \mathbf{V}_{\mathbf{Z}_d}$ such that $\bar{u} = \eta(u_z)$ and $d_{\mathbf{G}}(u, \bar{u}) < c$. Note that $\bar{u}$ could be u itself. If $\bar{u}$ has not yet been drawn, then draw it the lattice location of its corresponding lattice nodes, i.e.,

$$f(\bar{u}) = g(u_z).$$

Now draw u by choosing a random location inside an open ball of diameter 1 with the center at $f(\bar{u})$. To show that a drawing obtained this way is dense, first note that the largest uncovered diameter $\gamma < 2$ since a subset of the nodes of $\mathbf{V}$ occupy the lattice node positions. Pick any two nodes $u, v \in \mathbf{V}$. Again, from (ii), we know that there exists $\bar{u}, \bar{v} \in \eta(\mathbf{V}_{\mathbf{Z}_d}) \subset \mathbf{V}$ such that $d_{\mathbf{G}}(u, \bar{u}) \leq c$ and $d_{\mathbf{G}}(v, \bar{v}) \leq c$ for some positive constant c.

$$d_{\mathbf{G}}(u, v) \leq d_{\mathbf{G}}(u, \bar{u}) + d_{\mathbf{G}}(\bar{u}, \bar{v}) + d_{\mathbf{G}}(\bar{v}, v)$$
$$\leq 2c + h\, d_{\mathbf{G}^{(h)}}(\bar{u}, \bar{v})$$

Denoting $\bar{u} := \eta(u_z)$ and $\bar{v} := \eta(v_z)$,

$$d_{\mathbf{G}^{(h)}}(\bar{u}, \bar{v}) \leq d_{\mathbf{Z}_d}(\eta^{-1}(\bar{u}), \eta^{-1}(\bar{v}))$$
$$= \|g(u_z) - g(v_z)\|_1$$
$$\leq \sqrt{d}\|g(u_z) - g(v_z)\|$$
$$= \sqrt{d}\|f(\bar{u}) - f(\bar{v})\|$$
$$= \sqrt{d}\, d_f(\bar{u}, \bar{v}).$$

So we have $d_{\mathbf{G}}(u, v) \leq 2c + h\sqrt{d}\, d_f(\bar{u}, \bar{v})$. From Proposition 4, we see that the asymptotic distance scaling $\rho > 0$ for the f-drawing of $\mathbf{G}$, which establishes that $\mathbf{G}$ is dense in $\mathbb{R}^d$.

To prove the relationship (41) for any dense drawing f, consider again the scaled drawing f' defined as $f' = f/\gamma$, so that the maximum uncovered diameter of f' is 1. Since $\mathbf{G}$ is dense in $\mathbb{R}^d$, $\mathbf{Z}_d$ can be embedded in $\mathbf{G}^{(h)}$ with an embedding $\eta : \mathbf{V}_{\mathbf{Z}_d} \to \mathbf{V}$. We choose the embedding η as described in the first part of the proof, where every lattice node was mapped to a node of the graph whose f' drawing appeared inside a unit ball in $\mathbb{R}^d$ centered around the lattice node. For every $u \in \mathbf{V}$, call $u_z := \xi(u)$ and $\bar{u} := \psi(u)$, where $\xi : \mathbf{V} \to \mathbf{V}_{\mathbf{Z}_d}$ and $\psi : \mathbf{V} \to \mathbf{V}$ were defined earlier in this proof. Now,

$$
\begin{aligned}
d_{\mathbf{G}}(u, \bar{u}) &\leq \alpha d_f(u, \bar{u}) + \beta \\
&= \alpha\gamma d_{f'}(u, \bar{u}) + \beta \\
&\leq \alpha\gamma(\frac{3}{2}\sqrt{d}) + \beta,
\end{aligned}
$$

where we have used (42). Defining $c_1 := \alpha\gamma(\frac{3}{2}\sqrt{d}) + \beta$, we get the first part.
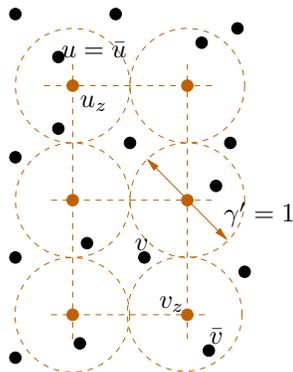

Fig. 16.   Natural drawing of the 2-D lattice (brown) superimposed on the f' drawing of $\mathbf{G}$. Edges are not shown to prevent clutter. In this example, $u = \bar{u}$ but $v \neq \bar{v}$.

Now consider two arbitrary nodes $u, v \in \mathbf{V}$ and let $u_z := \xi(u)$, $v_z := \xi(v)$(cf. Figure 16). Now,

$$
\begin{aligned}
d_{\mathbf{Z}_d}(u_z, v_z) &= \|g(u_z) - g(v_z)\|_1 \\
&\leq \sqrt{d}\|g(u_z) - g(v_z)\|,
\end{aligned}
$$

and

$$
\|g(u_z) - g(v_z)\| \leq \|g(u_z) - f'(\bar{u})\| + \|f'(\bar{u}) - f'(u)\| + \|f'(u) - f'(v)\| + \|f'(v) - f'(\bar{v})\| + \|f'(\bar{v}) - g(v_z)\|.
$$

We know that $\|g(u_z) - f'(\bar{u})\| \leq \frac{1}{2} \leq \frac{\sqrt{d}}{2}$ since $f'(\bar{u}) \in B(g(u_z), \frac{1}{2})$, and $\|f'(u) - f'(\bar{u})\| < \frac{3}{2}\sqrt{d}$ from (42). Using these in the above, we get

$$
\|g(u_z) - g(v_z)\| \leq 4\sqrt{d} + d_{f'}(u, v),
$$

$$
\Rightarrow d_{\mathbf{Z}_d}(u_z, v_z) \leq 4d + \frac{\sqrt{d}}{\gamma}d_f(u, v)
$$

which is the desired result.   ■

D. Covariance vs. Distance

It turns out that if a graph is dense/sparse in $\mathbb{R}^d$, we can derive upper/lower bounds on the scaling of BLU estimator covariance of a node u as a function of the Euclidean distance $d_f(u, o)$ from the reference node, where f is any drawing function that establishes the graph's denseness/sparseness in $\mathbb{R}^d$. The two theorems presented here are the main results of the paper; they utilize all the tools developed and results established so far in the paper. The results of these two theorems are also summarized in Table I for easy reference.

First some notation. The usual asymptotic notations $\Theta(\cdot), \Omega(\cdot)$ and $\mathcal{O}(\cdot)$ are used with matrix valued functions in the following way. First of all, for two matrices $A, B \in \mathbb{R}^{k \times k}$, $A \leq (<)B$ means $A - B$ is negative semidefinite (definite). For two functions $g : \mathbb{R} \to \mathbb{R}^{k \times k}$ and $f : \mathbb{R} \to \mathbb{R}$, the notation $g(x) = \mathcal{O}(f(x))$ means that there exists a constant x_o and

a positive definite matrix A such that $g(x) \leq Af(x)$ for all $x > x_o$. Similarly, the notation $g(x) = \Omega(f(x))$ means there exists a constant x_o and a positive definite matrix B such that $g(x) \geq Bf(x)$ for all $x > x_o$. Finally, $g(y) = \Theta(p(y))$ means that $g(y) = \Omega(p(y))$ and $g(y) = \mathcal{O}(p(y))$.

The first theorem establishes the upper bounds on the BLU estimator error covariance $\Sigma_{u,o}$ of a node variable x_u, $u \in \mathbf{V}$ for measurement graphs that are dense in $\mathbb{R}^d$.

Theorem 10. *Let* $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ *be a bounded degree measurement graph with a single reference node* $o \in \mathbf{V}$ *and with bounded edge covariances, i.e., there exists symmetric positive definite matrices* P_m *and* P_M *such that*

$$P_m \leq P(e) \leq P_M; \quad \forall e \in \mathbf{E}.$$

Then the following statements hold:
1) *If* G *is dense in* $\mathbb{R}^1$ *and* $f : \mathbf{V} \to \mathbb{R}^1$ *is any dense drawing of* $\mathbf{G}$*, then* $\Sigma_{u,o} = \mathcal{O}(d_f(u, o))$.
2) *If* G *is dense in* $\mathbb{R}^2$ *and* $f : \mathbf{V} \to \mathbb{R}^2$ *is any dense drawing of* $\mathbf{G}$*, then* $\Sigma_{u,o} = \mathcal{O}(\log d_f(u, o))$.
3) *If* G *is dense in* $\mathbb{R}^3$ *and* $f : \mathbf{V} \to \mathbb{R}^3$ *is any dense drawing of* $\mathbf{G}$*, then* $\Sigma_{u,o} = \mathcal{O}(1)$. $\qquad\square$

Proof: [Proof of Theorem 10] Throughout this proof, we will use $R^{\text{eff}}_{u,v}(\mathcal{G})$, for any graph $\mathcal{G}$, to denote the effective resistance between nodes u and v in the electrical network $(\mathcal{G}, P_M^{-1})$ with every edge of $\mathcal{G}$ having a generalized resistance of P_M. Consider the generalized electrical network $(\mathbf{G}, P_M^{-1})$ formed by assigning a constant generalized resistances of P_M to every edge of the measurement graph $\mathbf{G}$. From the Electrical analogy and Monotonicity Law (theorems 4 and 7), we get

$$\Sigma_{u,o} \leq R^{\text{eff}}_{u,o}(\mathbf{G}).$$

Since $\mathbf{G}$ is dense in $\mathbb{R}^2$, there exists a positive integer h such that the d-D lattice $\mathbf{Z}_d$ can be embedded in the h-fuzz of $\mathbf{G}$. Moreover, Theorem 9 tells us that $\exists u_z, o_z \in \mathbf{V}_{\mathbf{Z}_d}, c_1 > 0$ and an embedding $\eta : \mathbf{V}_{\mathbf{Z}_d} \to \mathbf{V}$ of $\mathbf{Z}_d$ into $\mathbf{G}^{(h)}$ such that

$$d_{\mathbf{G}}(u, \eta(u_z)) \leq c_1, \ d_{\mathbf{G}}(o, \eta(o_z)) \leq c_1$$

$$d_{\mathbf{Z}_d}(u_z, o_z) < 4d + \frac{\sqrt{d}}{\gamma} d_f(u, o), \tag{46}$$

where γ is the maximum uncovered diameter of the f-drawing of $\mathbf{G}$. Consider the electrical network $(\mathbf{G}^{(h)}, P_M^{-1})$ formed by assigning every edge of $G^{(h)}$ a resistance of P_M. From the Triangle Inequality for effective resistances (Theorem 6),

$$R^{\text{eff}}_{u,o}(\mathbf{G}^{(h)}) \leq R^{\text{eff}}_{u,\bar{u}}(\mathbf{G}^{(h)}) + R^{\text{eff}}_{\bar{u},\bar{o}}(\mathbf{G}^{(h)}) + R^{\text{eff}}_{\bar{o},o}(\mathbf{G}^{(h)}).$$

For any two nodes $u, v \in \mathbf{V}$, triangle inequality gives us $R^{\text{eff}}_{u,v}(\mathbf{G}^{(h)}) \leq d_{\mathbf{G}^{(h)}}(u, v)P_M \leq \frac{2}{h}d_{\mathbf{G}}(u, v)P_M$. Therefore

$$R^{\text{eff}}_{u,o}(\mathbf{G}^{(h)}) \leq \frac{2c_1}{h}P_M + R^{\text{eff}}_{\bar{u},\bar{o}}(\mathbf{G}^{(h)}). \tag{47}$$

Define $\bar{u} := \eta(u_z)$ for every $u_z \in \mathbf{Z}_d$. Since $\mathbf{G}^{(h)} \supset \mathbf{Z}_d$, from Monotonicity (Theorem 7), we get

$$R^{\text{eff}}_{\bar{u},\bar{o}}(\mathbf{G}^{(h)}) \leq R^{\text{eff}}_{u_z,o_z}(\mathbf{Z}_d).$$

When $\mathbf{G}$ is dense in, say, 2-Dimension, we have from Lemma 4 that

$$R^{\text{eff}}_{u_z,o_z}(\mathbf{Z}_2) = \Theta\left(\log d_{\mathbf{Z}_2}(u_z, o_z)\right).$$

Combining this with (46) and (47), we get

$$R^{\text{eff}}_{u,o}(\mathbf{G}^{(h)}) = \mathcal{O}\left(\log d_f(u, o)\right).$$

Since $\mathbf{G}$ is a bounded degree graph, from Lemma 3 we know that the effective resistance in $\mathbf{G}$ and its h-fuzz is of the same order:

$$R^{\text{eff}}_{u,o}(\mathbf{G}) = \Theta\left(R^{\text{eff}}_{u,o}(\mathbf{G}^{(h)})\right),$$

which finally gives us the desired result

$$R^{\text{eff}}_{u,o}(\mathbf{G}) = \mathcal{O}\left(\log d_f(u, o)\right).$$

The statements for $\mathbb{R}^1$ and $\mathbb{R}^3$ can be proved similarly. $\qquad\blacksquare$

The next theorem establishes lower bounds on the BLU estimator error covariance $\Sigma_{u,o}$ of a node variable x_u, $u \in \mathbf{V}$ for measurement graphs that are sparse in $\mathbb{R}^d$.

Theorem 11. *Let* $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ *be a bounded degree measurement graph with a single reference node* $o \in \mathbf{V}$ *and with bounded edge covariances, i.e., there exists symmetric positive definite matrices* P_m *and* P_M *such that*

$$P_m \leq P(e) \leq P_M; \quad \forall e \in \mathbf{E}.$$

Then the following statements hold:

1) *If G is sparse in $\mathbb{R}^1$ and $f : \mathbf{V} \to \mathbb{R}$ is any civilized drawing of G, then $\Sigma_{u,o} = \Omega\left(d_f(u,o)\right)$.*
2) *If G is sparse in $\mathbb{R}^2$ and $f : \mathbf{V} \to \mathbb{R}^2$ is any civilized drawing of G, then $\Sigma_{u,o} = \Omega\left(\log d_f(u,o)\right)$.*
3) *If G is sparse in $\mathbb{R}^3$ and $f : \mathbf{V} \to \mathbb{R}^3$ is any civilized drawing of G, then $\Sigma_{u,o} = \Omega(1)$.*

Proof: Throughout this proof, we will use $R_{u,v}^{\text{eff}}(\mathcal{G})$, for any graph $\mathcal{G}$, to denote the effective resistance between nodes u and v in the electrical network $(\mathcal{G}, P_m^{-1})$ with every edge of $\mathcal{G}$ having a generalized resistance of P_m. Consider the generalized electrical network $(\mathbf{G}, P_m^{-1})$ formed by assigning a constant generalized resistances of P_m to every edge of the measurement graph $\mathbf{G}$. From the Electrical analogy and Monotonicity Law (theorems 4 and 7), we get

$$\Sigma_{u,o} \geq R_{u,o}^{\text{eff}}(\mathbf{G}). \tag{48}$$

If $\mathbf{G}$ has multiple edges between some node pairs, let ℓ be the maximum number of parallel edges between any two node pairs. This upper bound exists since $\mathbf{G}$ is of bounded degree. We construct a graph $\mathbf{G}'$ from $\mathbf{G}$ by introducing multiple edges between nodes pairs until every pair of adjacent nodes in $\mathbf{G}'$ has exactly ℓ edges between them. Construct an electrical network $(\mathbf{G}', P_m)$ by assigning a generalized resistance P_m to every edge of $\mathbf{G}'$. From Rayleigh's Monotonicity law

$$R_{u,o}^{\text{eff}}(\mathbf{G}) \geq R_{u,o}^{\text{eff}}(\mathbf{G}').$$

We construct $\mathbf{G}''$ by removing $\ell - 1$ edges between every pair of nodes having an edge between them in $\mathbf{G}'$, so that there are no parallel edges in $\mathbf{G}''$. Construct $(\mathbf{G}'', P_m)$ by assigning a resistance of P_m to every edge of $\mathbf{G}''$. From Proposition 3 it can be deduced that

$$R_{u,o}^{\text{eff}}(\mathbf{G}') = \frac{1}{\ell} R_{u,o}^{\text{eff}}(\mathbf{G}''). \tag{49}$$

If there are no multiple edges in $\mathbf{G}$, then $\ell = 1$, $\mathbf{G} = \mathbf{G}' = \mathbf{G}''$, and the above equation holds trivially.

Since $\mathbf{G}$ is sparse in $\mathbb{R}^d$, $\mathbf{G}'$ and $\mathbf{G}''$ are sparse, too, and so $\mathbf{G}''$ can be drawn in a civilized manner in $\mathbb{R}^d$ with some drawing function f. Moreover, since $\mathbf{G}''$ has no multiple edges between node pairs, it follows from Theorem 8 that there exists a positive integer h, such that $\mathbf{G}''$ can be embedded in a h-fuzz of the d-D lattice. Consider the generalized electrical network $(\mathbf{Z}_d^{(h)}, P_m^{-1})$ formed by assigning a generalized resistance of P_m to every edge of $\mathbf{Z}_d^{(h)}$. From Rayleigh's Generalized Monotonicity Theorem 7, we get

$$R_{u,o}^{\text{eff}}(\mathbf{G}'') \geq R_{\bar{u},\bar{o}}^{\text{eff}}(\mathbf{Z}_d^{(h)}),$$

where $\bar{u}, \bar{o}$ refer to the nodes in $\mathbf{Z}_d^{(h)}$ that correspond to the nodes u, o in $\mathbf{G}$. In 2-dimension, for example, we have from Lemma 4 and Theorem 8

$$R_{u,o}^{\text{eff}}(\mathbf{G}'') = \Omega\left(\log d_{\mathbf{Z}_2}(\bar{u}, \bar{o})\right),$$
$$= \Omega\left(\log d_f(u, o)\right).$$

combining this with (48) and (49), we get the desired result in $\mathbb{R}^2$. The same can be done in $\mathbb{R}^1$ and $\mathbb{R}^3$. ∎

Graphs that are both sparse and dense in some Euclidean space $\mathbb{R}^d$ are a particularly nice class of graphs since we can provide exact scaling laws for the variance. We call a graph *coarsely equivalent to* $\mathbb{R}^d$ if it both sparse and dense in $\mathbb{R}^d$. The effective resistance in such a graph will grow with distance in the same rate as it will grow in the d-D lattice.

E. Are sensor networks Dense/Sparse?

A natural question to ask is - is it common for sensor networks to be sparse and/or dense in some Euclidean space $\mathbb{R}^d$? The answer happens to be "very much so" and this is often seen by considering the natural drawing of the network. Recall that a natural drawing of a sensor network is obtained by associating each node to its physical position in $1, 2$ or 3-dimensional Euclidean space. All natural drawings of sensor networks are likely to be sparse in 3-dimensional space, since the only requirements for sparseness are that nodes not lie on top of each other and edges be of finite length. When a sensor network is deployed in a 2-dimensional domain or when the third physical dimension is irrelevant, again the natural drawing is likely to be sparse in 2-dimensional space for the same reasons. It is slightly harder for a graph to satisfy the denseness requirements. Formally, a graph has to be infinite to be dense. However, what matters in practice are the properties of the graph "not too close to the boundary". Thus, a large-scale sensor network satisfies the denseness requirements, as long as there are no big holes between nodes and sufficiently many interconnections between them. A commonly encountered deployment model of sensor networks consists of nodes that are Poisson distributed random points in 2-dimensional space with an edge between every pair of nodes that they are within a certain range. Graphs produced by this model likely to be dense in 2-dimensional for a sufficiently large range (when compared to the intensity of the Poisson process). Underwater sensors filling a 3-dimensional volume using a similar model are likely to be dense in 3-dimensional space.

VII. Conclusion

Our results make precise the intuition that a "denser" graph will have a better estimate than a "sparse" graph, and shows (i) why naive measures of denseness such as node/edge density, node degree etc. does not work, (ii) what is the right way to define "denseness", and (iii) how does the variance of the BLU estimate scales with distance for large graphs that satisfy appropriate denseness properties. We have shown that the growth of variance with distance depends on how the graph *looks* when drawn, and therefore a proper measure of denseness of a graph lies in the properties of the drawing. For two classes of graphs - sparse and dense in $\mathbb{R}^d$, $1 \leq d \leq 3$, we established asymptotic lower and upper bounds on the variance of a node variable x_u's BLU estimate, as a function of the node's Euclidean distance from the reference node in a drawing of the graph. The main results are summarized in table I. Although the drawing of a graph is not unique, in practice this distance-in-the-drawing can be directly related to the distance in physical space where the objects in the network are located. It is reasonable to expect that real sensor networks will satisfy the criteria for being sparse and hence the lower bounds derived here will hold in practice. Denseness properties are slightly harder to satisfy, but large sensor networks can be dense for all practical purposes in an appropriate Euclidean space – when there are no arbitrarily big holes in it and the nodes have sufficiently many interconnections between them so that nodes situated close to each other in the Euclidean space have small graphical distance. In that case, we can get upper bounds on the estimation error of the optimal estimate. In practice, this provides a valuable guidelines for deploying a network so that accurate estimation is possible.

The bounds derived in this paper are for the BLU estimator, which achieves the minimum variance among all linear estimators. Thus these bounds serve as a benchmark against which the performance of other, possibly non-optimal, estimation algorithms can be compared. Due to various practical constraints in sensor networks, there might be a role for estimation algorithms other than the BLU estimator, in order to balance estimation accuracy and energy consumption. We have developed distributed algorithms to compute the BLU estimate in an iterative manner [1, 36]. However, in a poorly connected network, the number of iterations needed to get close to the BLU estimate may be large, which may lead to excessive energy costs. In such a situation, estimation algorithms that minimize energy consumption at the cost of a somewhat less accurate estimate may be warranted. Estimation error variance of such an algorithm can be compared against the variance bounds of the BLU estimator derived in this paper to benchmark their estimation accuracy.

References

[1] P. Barooah and J. P. Hespanha. Distributed optimal estimation from relative measurements. In *3rd International Conference on Intelligent Sensing and Information Processing*. 2005.

[2] P. Barooah, N. M. da Silva, and J. P. Hespanha. Distributed optimal estimation from relative measurements in sensor networks: Applications to localizationa and time synchronization. Tech. rep., University of California, Santa Barbara, 2006. URL `http://www.ccec.ece.ucsb.edu/~pbarooah/publications/TR3.html`.

[3] D. Niculescu and B. Nath. Error characteristics of ad hoc positioning systems (APS). In *MobiHoc '04: Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*. 2004.

[4] A. Savvides, W. Garber, , R. Moses, *et al.*. An analysis of error inducing parameters in multihop sensor node localization. *IEEE Transactions on Mobile Computing*, vol. 4, no. 6: 567–577, 2005.

[5] P. G. Doyle and J. L. Snell. Random walks and electric networks. Math. Assoc. of America, 1984.

[6] P. Barooah and J. P. Hespanha. Estimation from relative measurements: Error bounds from electrical analogy. In *Proceedings of the 2nd International Conference on Intelligent Sensing and Information Processing(ICISIP 05)*. 2005.

[7] P. G. Doyle. Personal communication. 2004.

[8] R. Karp, J. Elson, D. Estrin, and S. Shenker. Optimal and global time synchronization in sensornets. Tech. rep., Center for Embedded Networked Sensing,Univ. of California, Los Angeles, 2003.

[9] J. Caffery and G. Stber. Overview of radiolocation in CDMA cellular systems. *IEEE Communications Mag.*, vol. 36, no. 4: 38–45, 1998.

[10] D. Moore, J. Leonard, D. Rus, and S. Teller. Robust distributed network localization with noisy range measurements. In *Proceedings of the Second ACM Conference on Embedded Networked Sensor Systems (SenSys 04)*. 2004.

[11] N. B. Priyantha, A. K. L. Miu, H. Balakrishnan, and S. J. Teller. The cricket compass for context-aware mobile applications. In *Mobile Computing and Networking*, pp. 1–14. 2001.

[12] Sick laser range finder. URL `http://www.sick.de`.

[13] Crossbow technologies. URL `http://www.xbow.com`.

[14] N. Malhotra, M. Krasniewski, C. Yang, S. Bagchi, *et al.*. Location estimation in ad-hoc networks with directional antenna. In *25th International Conference on Distributed Computing Systems: ICDCS(To appear)*. 2005.

[15] A. Kalis and T. Antanakopolous. Direction finding in IEEE802.11 wireless networks. *IEEE Transactions on Instrumentation and Measurement*, vol. 51, no. 5: 940–948, 2002.

[16] A. Kalis, T. Antanakopolous, and V. Makios. A printed circuit switched array antenna for indoor communications. *IEEE Transactions on Consumer Electronics*, vol. 46: 531–538, 2000.

[17] T. Ueda, S. Tanaka, D. Saha, S. Roy, *et al.*. An efficient mac protocol with direction finding scheme in wireless ad hoc network using directional antenna. In *Radio and Wireless Conference (RaWCon)*. 2003.

[18] N. Patwari, A. O. H. III, and M. Perkins. Relative location estimation in wireless sensor networks. *IEEE Transactions in Signal Processing*, vol. 51, no. 8: 2137–2148, 2003.

[19] A. Savvides, C. Han, and M. B. Strivastava. Dynamic fine-grained localization in ad-hoc networks of sensors. In *Mobile Computing and Networking, MobiCom 2001, Rome, Italy*, pp. 166–179. 2001.

[20] L. Doherty, K. S. J. Pister, and L. E. Ghaoui. Convex position estimation in wireless sensor network. In *IEEE INFOCOM*, pp. 1655–1663. 2001.

[21] D. Niculescu and B. Nath. Ad hoc positioning system (APS) using AOA. In *IEEE INFOCOM*. 2003.

[22] K. Chintalapudi, A. Dhariwal, R. Govindan, and G. Sukhatme. Ad-hoc localization using ranging and sectoring. In *IEEE Infocomm*. 2004.

[23] H. A. Pratik Biswas and Y. Ye. Integration of angle of arrival information for multimodal sensor network localization using semidefinite programming. *EURASIP Journal on Applied Signal Processing*. Submitted.

[24] T. Eren, W. Whiteley, and P. N. Belhumeur. Further results on sensor network localization using rigidity. In *Proceedings of the Second European Workshop on Sensor Networks (EWSN)*, pp. 405–409. 2005.

[25] A. Savvides, W. Garber, S. Adlakha, R. Moses, *et al.*. On the error characteristics of multi-hop node localization in ad-hoc sensor networks. In *Second International Workshop on Information Processing in Sensor Networks (IPSN)*. 2003.

[26] W.-K. Chen. *Applied Graph Theory*. North Holland Publishing Company, 1971.

[27] C. Godsil and G. Royle. *Algebraic Graph Theory*. Graduate Texts in Mathematics. Springer, 2001.

[28] F. Chung. Laplacians and the cheeger inequality for directed graphs. *Annals of Combinatorics*, vol. 9, no. 1: 1–19, 2005.

[29] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, vol. 49: 1520–1533, 2004.

[30] J. M. Mendel. *Lessons in Estimation Theory for Signal Processing, Communications and Control*. Prentice Hall P T R, 1995.

[31] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 1999.

[32] D. Klein and M. Randić. Resistance distance. *Journal of Mathematical Chemistry*, vol. 12: 81–95, 1993.

[33] P. G. Doyle. Application of Rayleigh's short-cut method to Polya's recurrence problem. online, 1998. URL `http://math.dartmouth.edu/~doyle/docs/thesis/thesis/`.

[34] G. Venezian. On the resistance between two points on a grid. *Am. Jour. Phys.*, vol. 62, no. 11: 1000–1004., 1994.

[35] J. Cserti. Application of the lattice green's function for calculating the resistance of and infinite network of resistors. *Am. Jour. Phys.*, vol. 68, no. 10: 896–906., 2000.

[36] P. Barooah, N. M. da Silva, and J. P. Hespanha. Distributed optimal estimation from relative measurements in sensor networks: Applications to localizationa and time synchronization., 2006. Submitted, available at `http://www.ece.ucsb.edu/~pbarooah/publications.html`.

APPENDIX I

Proof: [Proof of Theorem 6] Consider three generalized electrical networks $\mathcal{G}_1, \mathcal{G}_2$ and $\mathcal{G}_3$ with the same graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ and the same edge-resistance function R, but with different points of current injection. In particular, in $\mathcal{G}_1$, a generalized current of matrix-intensity $\mathbf{i} \in \mathbb{R}^{k \times k'}$ is injected at $p \in \mathbf{V}$ and extracted at $q \in \mathbf{V}$, in $\mathcal{G}_2$, a current of the same intensity $\mathbf{i}$ is injected at p and extracted at r, and in $\mathcal{G}_3$ a current of the same intensity $\mathbf{i}$ is injected at r and extracted at q. Because of linearity, the superposition of currents and potentials in networks $\mathcal{G}_2$ and $\mathcal{G}_3$ gives the currents and potentials in $\mathcal{G}_1$. The situation is schematically represented in figure 17. Denoting by $i_j(e)$ the current along the edge e in the electrical network $\mathcal{G}_j$, we have

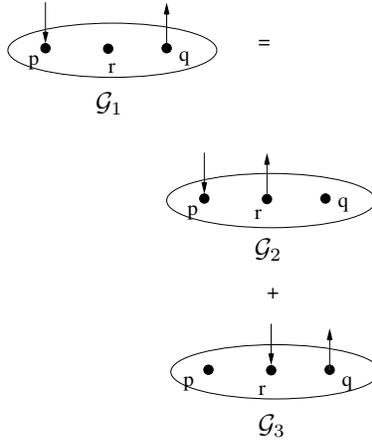$$i_1(e) = i_2(e) + i_3(e), \quad \forall e \in \mathbf{E}.$$

Fig. 17.   A schematic of networks $\mathcal{G}_1$, $\mathcal{G}_2$, $\mathcal{G}_3$.

Therefore

$$\sum_{e \in \mathbf{E}} i_1^T(e)R(e)i_1(e) = \sum_{e \in \mathbf{E}} (i_2^T(e) + i_3^T(e))R(e)(i_2(e) + i_3(e))$$

$$= \sum_{e \in \mathbf{E}} i_2^T(e)R(e)i_2(e) + \sum_{(e) \in \mathbf{E}} i_3^T(e)R(e)i_3(e) + \sum_{e \in \mathbf{E}} i_2^T(e)R(e)i_3(e) + \sum_{e \in \mathbf{E}} i_3^T(e)R(e)i_2. \quad (50)$$

From (24), it follows that

$$\mathbf{i}^T R_{p,q}^{\text{eff}} \mathbf{i} = \mathbf{i}^T R_{p,r}^{\text{eff}} \mathbf{i} + \mathbf{i}^T R_{r,q}^{\text{eff}} \mathbf{i} + \sum_{e \in \mathbf{E}} i_2^T(e)R(e)i_3(e) + \sum_{e \in \mathbf{E}} i_3^T(e)R(e)i_2. \quad (51)$$

Furthermore, consider a network $\mathcal{G}_{3'}$ that is almost the same as $\mathcal{G}_3$, except that the current is injected at q and extracted at r. In all three networks $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_{3'}$, the node r is assigned a generalized potential of 0. From Linearity again, we have

$$i_1(e) = i_2(e) - i_{3'}(e), \quad \forall e \in \mathbf{E},$$
$$V_1(u) = V_2(u) - V_{3'}(u), \quad \forall u \in \mathbf{V}.$$

Let $\mathcal{V}_b^{(j)}$ be the tall vector obtained by stacking the voltages of all nodes $u \in \mathbf{V} \setminus \{r\}$ in the network $\mathcal{G}_j$. Similarly, let $\mathcal{I}^{(j)}$ be the tall vector obtained by stacking the currents on all edges $e \in \mathbf{E}$ in the network $\mathcal{G}_j$. We index the nodes appropriately so that node r corresponds to the last "rows" in the matrix $\mathcal{V}$ as well as in the gen. incidence matrix $\mathcal{A}$ and the Dirichlet Laplacian $\mathcal{L}$ of the weighted graph $(\mathbf{G}, R^{-1})$. We can show, as it was done in the proof of Theorem 4, that

$$\mathcal{I}^{(j)} = \mathscr{R}^{-1}\mathcal{A}^T \begin{bmatrix} \mathcal{V}_b^{(j)} \\ 0 \end{bmatrix}$$

and

$$\mathcal{V}^{(2)} = \mathcal{L}^{-1}(e_p \otimes \mathbf{i}),$$
$$\mathcal{V}^{(3')} = \mathcal{L}^{-1}(e_q \otimes \mathbf{i}),$$

where $e_u \in \mathbb{R}^{n-1}$ is a vector of all zeros except at the uth position, where it has a 1. Now,

$$\sum_{e \in \mathbf{E}} i_2^T(e)R(e)i_3(e) = -\sum_{e \in \mathbf{E}} i_2^T(e)R(e)i_{3'}(e)$$

$$= -\mathcal{I}^{(2)}\mathscr{R}\mathcal{I}^{(3')}$$

$$= -[\mathcal{V}_b^{(2)}|0]\mathcal{A}\mathscr{R}^{-1}\mathscr{R}\mathscr{R}^{-1}\mathcal{A}^T \begin{bmatrix} \mathcal{V}_b^{(3')} \\ 0 \end{bmatrix}$$

$$= -\begin{bmatrix}(e_p \otimes \mathbf{i})^T\mathcal{L}^{-1} & 0\end{bmatrix} \begin{bmatrix} \mathcal{L} & \cdot \\ \cdot & \cdot \end{bmatrix} \times \begin{bmatrix} \mathcal{L}^{-1}(e_q \otimes \mathbf{i}) \\ 0 \end{bmatrix}$$

$$= -[(e_p \otimes \mathbf{i})^T\mathcal{L}^{-1}(e_q \otimes \mathbf{i})]$$

When $\mathbf{i} = I_k$ (the $k \times k$ identity matrix), we get from the above and Theorem 4

$$\sum_{e \in \mathbf{E}} i_2^T(e) R(e) i_3(e) = -R_{p,q;r}^{\mathrm{cross}}$$

and similarly,

$$\sum_{e \in \mathbf{E}} i_3^T(e) R(e) i_2(e) = -R_{q,p;r}^{\mathrm{cross}} = -R_{p,q;r}^{\mathrm{cross}}.$$

Therefore, using $\mathbf{i} = I_k$ in (51), we get

$$R_{p,q}^{\mathrm{eff}} = R_{p,r}^{\mathrm{eff}} + R_{r,q}^{\mathrm{eff}} - 2R_{p,q;r}^{\mathrm{cross}}$$
$$\leq R_{p,r}^{\mathrm{eff}} + R_{r,q}^{\mathrm{eff}},$$

since $R_{p,q;r}^{\mathrm{cross}} = \mathrm{E}[(x_p - \hat{x}_p^*)(x_q - \hat{x}_q^*)^T]$ (Theorem 4) is a symmetric positive semidefinite matrix.

That equality is achieved when every path from p to q passes through r is straightforward, and is left as an exercise for the reader. $\blacksquare$

The next lemma is required to prove Lemma 3. We say that two nodes u and v in a network $(\mathbf{G}, R)$ are *shorted* if there is an edge $e \sim u, v$ between them with a resistance 0, i.e., $R_e = \mathbf{0} \in \mathbb{R}^{k \times k}$. We sometimes say that an *edge* (u,v) *is shorted*, to mean that u and v are shorted. Theorem 7 states that taking edges out of a network, or "cutting" edges increases effective resistance. The following result is the counterpart of that, and states that "shorting" edges decreases effective resistance. These are the exact analogs of Rayleigh's cutting and shorting laws described by Doyle and Snell [5] for regular electrical networks.

Lemma 5 (Shorting Law). *Consider two electrical networks* $(\mathbf{G}, R)$ *and* $(\mathbf{G}, \bar{R})$ *with the same graph* $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ *and edge resistances such that* $R(e) = \bar{R}(e) \in \mathbb{R}^{k \times k}$ *for every edge* $e \in \mathbf{E}$ *except for one edge* $e^* \sim p, q$, *for which* $R(e^*) > 0$ *and* $\bar{R}(e^*) = 0$. *That is,* p *and* q *are shorted in* $(\mathbf{G}, \bar{R})$. *Then,*

$$R_{u,v}^{\mathrm{eff}} \text{ in } (\mathbf{G}, R) \geq R_{u,v}^{\mathrm{eff}} \text{ in } (\mathbf{G}, \bar{R})$$

for every pair of nodes $u, v \in \mathbf{V}$ *such that* $\{u,v\} \neq \{p,q\}$. $\square$

Proof: Let i be the current in $(\mathbf{G}, R)$ with intensity $\mathbf{i} \in \mathbb{R}^{k \times k'}$ from u to v, where u, v are not p, q. Similarly, let $\bar{i}$ be the current in $(\mathbf{G}, \bar{R})$ with the same intensity from u to v. Since $\bar{R}_{e^*} = 0$, we get

$$\sum_{e \in \mathbf{E}} Tr(i_e^T R_e i_e) \geq \sum_{e \in \mathbf{E}} Tr(i_e^T \bar{R}_e i_e)$$
$$\geq \sum_{e \in \mathbf{E}} Tr(\bar{i}_e^T \bar{R}_e \bar{i}_e),$$

where the second inequality follows from i being a flow in $\mathbf{G}$ (Thomson's minimum energy principle). From corollary 1,

$$Tr(\mathbf{i}^T R_{u,v}^{\mathrm{eff}} \mathbf{i}) \geq Tr(\mathbf{i}^T \bar{R}_{u,v}^{\mathrm{eff}} \mathbf{i}).$$

Sine this is true for every $\mathbf{i} \in \mathbb{R}^{k \times k'}$, the result follows. $\blacksquare$

Remark 2. It should be emphasized that so far we have only allowed edge resistance to be positive definite. It turns out that effective resistance between nodes that are not shorted together is well defined even when there are 0 edge-resistances in the network. The problem of determining currents and voltages in the edges of the network is no longer well-posed if we allow shorted edges. However, when one constructs a "short-equivalent" network by replacing shorted node sets by single nodes and all edges between shorted nodes, currents and voltages in this short-equivalent network is well-defined. It should be emphasized that general positive semi-definite edge resistances are still not allowed. $\square$

Now we are ready to prove Lemma 3. The proof follows exactly along the steps of that by [33] for "regular" electrical networks. The only technicality involved here is that of "shorting" for generalized networks, which is taken care of by the previous lemma.

Proof: [Proof of Lemma 3] The upper bound $R^{\mathrm{eff}}(\mathbf{G}^{(h)}) \leq R^{\mathrm{eff}}(\mathbf{G})$ follows from Rayleigh's Monotonicity Law. The edge set $E^{(h)}$ of the h-fuzz $\mathbf{G}^{(h)}$ consists of two disjoint subsets $\mathbf{E}$ and $\mathbf{E}_h$, where $\mathbf{E}_h$ is the set of "new" edges in $\mathbf{G}^{(h)}$ that were not there in $\mathbf{G}$. Thus, $\mathbf{E}^{(h)} = \mathbf{E} \cup \mathbf{E}_h$ and $\mathbf{E} \cap \mathbf{E}_h = \emptyset$. To every edge $e \in \mathbf{E}_h$, there corresponds a path $\mathcal{P}_e$ of length $\ell_e \leq h$ in $\mathbf{G}$. See figure 18(A) and (B) for an example. Replace every edge $e \in \mathbf{E}_h$ by a series of ℓ_e edges and call the resulting graph $\overline{\mathbf{G}^{(h)}} = (\overline{\mathbf{V}}, \overline{\mathbf{E}^{(h)}})$. Each new edge is assigned a resistance of R_o. We also introduce new nodes in doing so. To every one of these new nodes in $\overline{\mathbf{G}^{(h)}}$, there corresponds a node in $\mathbf{G}$ (cf. Figure 18(C)).
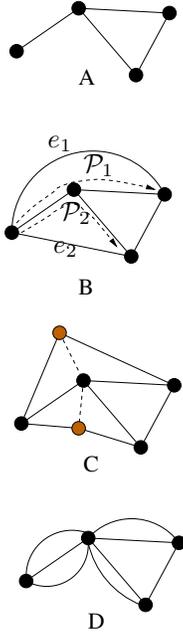
Fig. 18. Fuzzing doesn't change the effective resistance too much.

Let $R_{u,v}^{\text{eff}}(\mathbf{G})$ (for any graph $\mathbf{G}$) refer to the effective resistance between nodes u and v in the network $(\mathbf{G}, R_o)$ that has a generalized resistance of R_o on every edge of $\mathbf{G}$. By Rayleigh's Monotonicity Law, the effective resistance $R_{u,v}^{\text{eff}}(\overline{\mathbf{G}^{(h)}})$ is higher than $R_{u,v}^{\text{eff}}(\mathbf{G}^{(h)})$. However, since we have increased the resistance of any edge by no more than a factor of h, the increase in effective resistance is no more than a factor of h:

$$R_{u,v}^{\text{eff}}(\overline{\mathbf{G}^{(h)}}) \leq h R_{u,v}^{\text{eff}}(\mathbf{G}^{(h)}). \tag{52}$$

Now for every edge e of $\mathbf{E}_h$, look at the corresponding series of resistors in $\overline{\mathbf{E}^{(h)}}$. Its endpoints lie in the original graph $\mathbf{G}$. Take its intermediate vertices and "short" them to vertices of $\overline{\mathbf{G}^{(h)}}$ along the path $\mathcal{P}_e$ with edges with a $\mathbf{0} \in k \times k$ resistance. We do this for every edge $e \in \mathbf{E}_h$ and call the resulting graph $\mathbf{G}'$ (cf. Figure 18(D)). From the shorting Law (Lemma 5), we get

$$R_{u,v}^{\text{eff}}(\mathbf{G}') \leq R_{u,v}^{\text{eff}}(\overline{\mathbf{G}^{(h)}}), \quad \forall u, v \in \mathbf{V}. \tag{53}$$

The graph $\mathbf{G}'$ differs from $\mathbf{G}$ only in having extra parallel edges between nodes (compare Figure 18(A) and (B)). However, the number of edges in $\mathbf{G}'$ that are parallel to an edge e in $\mathbf{G}$ are no more than the number of paths in $\mathbf{G}$ of length $\leq h$ that traverse the edge e. Since $\mathbf{G}$ is a bounded degree graph, there is an upper bound to this number, which we call ℓ_d. The effective resistance in the network $(\mathbf{G}', R_o)$ is higher than that in a network $(\mathbf{G}'', R_o)$ where $\mathbf{G}''$ is a graph that is obtained from $\mathbf{G}$ by replacing every edge in $\mathbf{G}$ by ℓ_d parallel edges. Each edge in $(\mathbf{G}'', R_o)$ has the same resistance R_o. Thus,

$$\begin{aligned} R_{u,v}^{\text{eff}}(\mathbf{G}') &\geq R_{u,v}^{\text{eff}}(\mathbf{G}'') \\ &= R_{u,v}^{\text{eff}} \text{ in } (\mathbf{G}, R_o/\ell_d), \\ &= \frac{1}{\ell_d} R_{u,v}^{\text{eff}}(\mathbf{G}) \end{aligned} \tag{54}$$

where the second equality follows from Proposition 3 and third from the fact that changing the edge-resistances by a constant factor changes the effective resistances by the same factor. Equations (52),(53) and (54) give us

$$R_{u,v}^{\text{eff}}(\mathbf{G}^{(h)}) \geq \frac{1}{h\ell_d} R_{u,v}^{\text{eff}}(\mathbf{G}), \tag{55}$$

which is the desired result, since the constant $\frac{1}{h\ell_d}$ does not depend on u and v.

APPENDIX II

Proof: [Proof of Proposition 4] We prove that 1 implies 2 by contradiction. Assuming that 2 does not hold, we have that

$$\forall \alpha \quad \forall \beta \quad \exists \bar{u}, \bar{v} \in \mathbf{V} \text{ such that } d_{\mathbf{G}}(\bar{u}, \bar{v}) > \alpha d_f(\bar{u}, \bar{v}) + \beta.$$

or equivalently

$$\forall \alpha \quad \forall \beta \quad \exists \bar{u}, \bar{v} \in \mathbf{V}$$

such that

$$\frac{d_f(\bar{u}, \bar{v})}{d_{\mathbf{G}}(\bar{u}, \bar{v})} < \frac{1}{\alpha} - \frac{\beta}{\alpha d_{\mathbf{G}}(\bar{u}, \bar{v})}.$$

This means that for a given α, β, the set

$$\left\{ \frac{d_f(u, v)}{d_{\mathbf{G}}(u, v)} : u, v \in \mathbf{V} \text{ and } d_{\mathbf{G}}(u, v) \geq \beta \right\}$$

contains at least the element

$$\frac{d_f(\bar{u}, \bar{v})}{d_{\mathbf{G}}(\bar{u}, \bar{v})} < \frac{1}{\alpha} - \frac{\beta}{\alpha d_{\mathbf{G}}(\bar{u}, \bar{v})} < \frac{1}{\alpha}$$

and therefore

$$\inf \left\{ \frac{d_f(u, v)}{d_{\mathbf{G}}(u, v)} : u, v \in \mathbf{V} \text{ and } d_{\mathbf{G}}(u, v) \geq \beta \right\} < \frac{1}{\alpha}.$$

Making $\beta \to \infty$ we obtain that

$$\rho = \lim_{\beta \to \infty} \inf \left\{ \frac{d_f(u, v)}{d_{\mathbf{G}}(u, v)} : u, v \in \mathbf{V} \text{ and } d_{\mathbf{G}}(u, v) \geq \beta \right\} < \frac{1}{\alpha}.$$

But since α can be arbitrarily large, the above actually implies that $\rho = 0$, which contradicts 1.

To prove that 2 implies 1, we note that when 2 holds, we conclude that for every pair of nodes $u, v \in \mathbf{V}$, for which $d_{\mathbf{G}}(u, v) \geq n$, we have that

$$\frac{d_f(u, v)}{d_{\mathbf{G}}(u, v)} \geq \frac{1}{\alpha} - \frac{\beta}{d_{\mathbf{G}}(u, v)} \geq \frac{1}{\alpha} - \frac{\beta}{n}, \quad \forall u \neq v \in \mathbf{V}.$$

Therefore,

$$\inf \left\{ \frac{d_f(u, v)}{d_{\mathbf{G}}(u, v)} : u, v \in \mathbf{V} \text{ and } d_{\mathbf{G}}(u, v) \geq n \right\} \geq \frac{1}{\alpha} - \frac{\beta}{n}.$$

As $n \to \infty$, the left-hand side converges to ρ and the right-hand side converges to $\frac{1}{\alpha} > 0$, from which 1 follows. $\qquad \square$