A graph theoretical framework for consistent time differences of arrival

Martin Kreißig, Bin Yang

Kurzdresse: Lehrstuhl für Systemtheorie und Signalverarbeitung, Universität Stuttgart, 70550 und Stuttgart
E-Mail: {martin.kreissig,bin.yang}@lss.uni-stuttgart.de
Web: www.lss.uni-stuttgart.de

Abstract

The estimation of time difference of arrival (TDOA) plays an important role in acoustic source localization, blind source separation, acoustic sound control, and improved speech communication. However, when multiple sources are active in a reverberant environment, the TDOA estimates are known to suffer from the multipath and multiple source ambiguity, and hence are less reliable. In a previous work [1], we introduced the idea of consistent graphs to combat these TDOA ambiguities. In this paper, we go one step further and present a graph theoretical framework to study this consistency problem. It turns out that many concepts from the graph theory are useful for this purpose.

1 Introduction

In a variety of acoustic signal processing applications, one needs to estimate the TDOA between a source and a pair of microphones from the microphone signals [2]. Different families of methods like generalized cross-correlation (GCC), adaptive eigenvalue decomposition, and blind source separation exist for this purpose. A common problem is, however, that these methods often return several TDOA estimates for one pair of microphones. They can stem from the direct or echo path and can originate from different sources. The matching of these ambiguous TDOA estimates among different microphone pairs is thus an important and challenging problem.

In [1, 3], we made a simple observation that the matching TDOA values, i.e. these from the same source and the same propagation paths, satisfy the so called zero cyclic sum condition: the sum of them is zero for any closed path over a number of microphones. These matching TDOA values are said to be consistent. In addition, we proposed the DATEMM algorithm, an ad hoc bottom-up synthesis approach, to find consistent TDOA values from sets of TDOA candidates for different microphone pairs. This algorithm worked pretty well in experiments.

The goal of this work is to provide a theoretically more rigorous, graph oriented framework to analyze the consistency problem. We discover a parallel between TDOA estimation and circuit analysis, discuss general properties of consistent graphs, present some novel approaches for their synthesis, and address open questions.

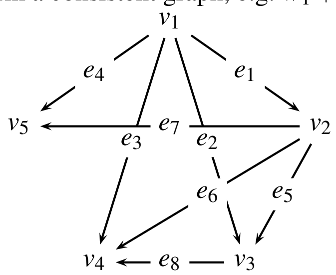
2 Basics of graph theory

Below we briefly summarize some basic definitions from the graph theory we need in this paper [4, 5].

A graph $G(V,E)$ is defined as a set of M vertices (or nodes) $V = \{v_1,\dots,v_M\}$ and a set of N edges (or lines, branches, arcs) $E = \{e_1,\dots,e_N\}$. We consider only simple graphs in which there is at most one edge between two vertices and there are no loops involving only one vertex. Hence, the number of edges is limited by $N \le N_{\max} = \binom{M}{2} = \frac{M(M-1)}{2}$. The graph is complete if $N = N_{\max}$, i.e. each disjoint pair of vertices is connected by an edge. The graph is connected if each pair of disjoint vertices is connected by at least one path. A path is a sequence of neighboured edges, that share the same terminal vertices. A necessary condition for a graph to be connected is $N \ge M-1$. A loop (or cycle) in a graph is a closed path whose start and end vertex coincide.

We consider directed graphs in which each edge has a direction (shown by an arrow) pointing from the start vertex to the end vertex. Similarly, we define a direction for each loop. We also consider weighted graphs in which one weight $w_n \in \mathbb{R}$ is assigned to each edge e_n $(1 \le n \le N)$. The weight is defined such that it changes its sign if we change the direction of the edge. A graph is said to be consistent if the sum of edge weights along any loop, taking the edge direction into account, is zero.

Fig. 1 shows a simple, directed, weighted graph consisting of $M = 5$ vertices and $N = 8$ edges. It is connected, but not complete because of two missing edges (v_3, v_4) and (v_4, v_5). The arrows of the edges show their direction. It is easy to verify that the given edge weights w_e in Fig. 1 form a consistent graph, e.g. $w_1 + w_5 - w_2 = 0$.



$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ w_7 \\ w_8 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 4 \\ 1 \\ -2 \\ 1 \\ -2 \\ 3 \end{bmatrix}$$

Figure 1: Example of a consistent graph containing 5 vertices and 8 edges

Such a graph representation has been successively applied to the analysis of electrical circuits by the Kirchhoff laws. It is also a useful description of TDOA based source localization. Table 1 shows the parallels between these different topics.

graph theory	electrical circuit	TDOA based localization
vertex	node	microphone
vertex value	elec. potential	TOA
edge	branch	pair of microph.
edge weight	voltage, diff. of potentials	TDOA, diff. of TOA
consistent graph	Kirchhoff 2nd law	zero cyclic sum of TDOA

Table 1: Parallels between consistent graphs, analysis of electrical circuits and TDOA based localization

3 Incidence and loop matrix

The topology of a graph can be represented by the incidence matrix $\mathbf{A}$. It is an $N \times M$ matrix containing the values 1, -1, and 0. It shows the relation between the edges and their start and end vertices where vertices are shown in the columns and edges in rows. A start vertex of an edge

ISBN 978-3-8007-3300-2

is marked with 1, the end vertex is marked with -1, and all other matrix elements are zero. The incidence matrix for the graph in Fig. 1 is

$$\mathbf{A} = \left[\begin{array}{l|ccccc} \text{vertices} \rightarrow & v_1 & v_2 & v_3 & v_4 & v_5 \\ \text{edges} \downarrow & & & & & \\ \hline e_1 & 1 & -1 & 0 & 0 & 0 \\ e_2 & 1 & 0 & -1 & 0 & 0 \\ e_3 & 1 & 0 & 0 & -1 & 0 \\ e_4 & 1 & 0 & 0 & 0 & -1 \\ e_5 & 0 & 1 & -1 & 0 & 0 \\ e_6 & 0 & 1 & 0 & -1 & 0 \\ e_7 & 0 & 1 & 0 & 0 & -1 \\ e_8 & 0 & 0 & 1 & -1 & 0 \end{array}\right] \quad (1)$$

Another useful matrix describing the relationship between loops and edges is the loop (or cycle) matrix $\mathbf{B}$. It shows the N edges in rows and all loops in columns. If an edge contributes to a loop, the corresponding element in $\mathbf{B}$ is marked with 1 if the edge and loop direction coincide or -1 otherwise. A zero element in $\mathbf{B}$ indicates that the edge is not contained in the corresponding loop. The loop matrix for the graph in Fig. 1 is

$$\mathbf{B} = [\mathbf{B}_f \ldots], \mathbf{B}_f = \left[\begin{array}{l|cccc} \text{loops} \rightarrow & l_1 & l_2 & l_3 & l_4 \\ \text{edges} \downarrow & & & & \\ \hline e_1 & 1 & 1 & 1 & 0 \\ e_2 & -1 & 0 & 0 & 1 \\ e_3 & 0 & -1 & 0 & -1 \\ e_4 & 0 & 0 & -1 & 0 \\ e_5 & 1 & 0 & 0 & 0 \\ e_6 & 0 & 1 & 0 & 0 \\ e_7 & 0 & 0 & 1 & 0 \\ e_8 & 0 & 0 & 0 & 1 \end{array}\right] \quad (2)$$

where only four (fundamental) loops $l_1, \ldots, l_4$ are shown for simplicity. Fig. 2 shows the loops of Eq. (2).

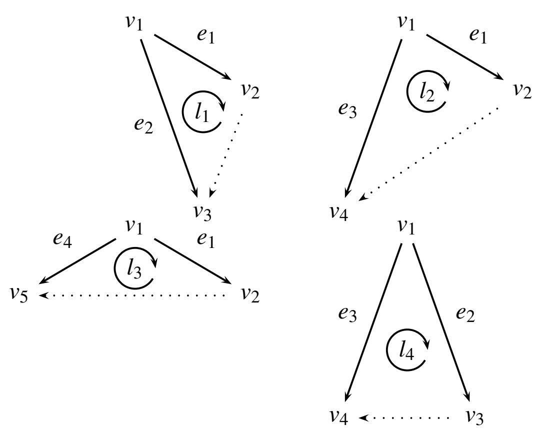


Figure 2: Four fundamental loops $l_1, \ldots, l_4$ of the graph in Fig. 1

Both the incidence and loop matrix are determined by the graph topology only. They are independent of the edge weights and have been shown to be useful in the analysis of electrical networks. For a connected graph (i.e. $N \geq M-1$), the following rank properties of $\mathbf{A}$ and $\mathbf{B}$ are well known [6]:

$$\text{rank}(\mathbf{A}) = M-1, \quad \text{rank}(\mathbf{B}) = N-M+1. \quad (3)$$

In addition, $\mathbf{B}^T\mathbf{A} = \mathbf{0}$. This implies, the column vectors of $\mathbf{A}$ span an $(M-1)$-dimensional subspace in $\mathbb{R}^N$, the column space or range $R(\mathbf{A})$ of $\mathbf{A}$, while the column vectors of $\mathbf{B}$ span the $(N-M+1)$-dimensional orthogonal complement of $R(\mathbf{A})$.

The rank properties of $\mathbf{B}$ implies that only $N-M+1$ columns in $\mathbf{B}$ are linearly independent. Each such column represents a so called fundamental loop. Let $\mathbf{B}_f$ be the $N \times (N-M+1)$ loop matrix of these fundamental loops. Clearly, $R(\mathbf{B}_f) = R(\mathbf{B})$ and $\mathbf{B}_f^T\mathbf{A} = \mathbf{0}$.

These two rank properties have a graphical interpretation. Any set of $M-1$ linearly independent rows of $\mathbf{A}$ represents a subgraph with a minimal set of $M-1$ edges that connect all vertices in G. Such a graph has no loop and is called spanning tree of G. For the graph in Fig. 1, a possible spanning tree is given by the solid lines in Fig. 2, namely e_1, e_2, e_3 and e_4. The remaining $N-M+1$ edges of G compose the complementary tree (cotree), denoted as dotted lines in Fig. 2. As seen in Fig. 2, any edge of the cotree, when added to the spanning tree, closes one fundamental loop. Hence $R(\mathbf{B}) = R(\mathbf{B}_f)$ is spanned by the $N-M+1$ fundamental loops. While the cotree is implicitly given by the spanning tree, the latter is comletely arbitrary.

4 Properties of a consistent graph

Now we consider a graph with the edge weight vector $\mathbf{w} = [w_1, \ldots, w_N]^T$. The consistency of $\mathbf{w}$ implies $\mathbf{B}^T\mathbf{w} = \mathbf{0}$, i.e. the cyclic sum of edge weights along all loops is zero. However, the total number of loops in a graph can be very large. Fortunately, it is only necessary to check the consistency of $\mathbf{w}$ along the $N-M+1$ fundamental loops since all columns of $\mathbf{B}$ can be written as a linear combination of the columns of $\mathbf{B}_f$. A necessary and sufficient condition for $\mathbf{w}$ being consistent is therefore

$$\mathbf{B}_f^T\mathbf{w} = \mathbf{0}. \quad (4)$$

An equivalent consistency condition is that $\mathbf{w}$ is from the column space $R(\mathbf{A})$ of the incidence matrix $\mathbf{A}$.

In the following, we discuss some general properties of a consistent TDOA weight vector $\mathbf{w}$ for a given incidence matrix $\mathbf{A}$ and loop marix $\mathbf{B}_f$, as well as their physical interpretations in the context of TDOA based source localization.

If $\mathbf{w}$ is consistent in the sense of Eq. (4) and $\mathbf{a}_i$ is the i^{th} column vector of the incidence matrix $\mathbf{A}$, then

P1) $\alpha\mathbf{w}$ is also consistent $\forall\alpha \in \mathbb{R}$.

P2) $\mathbf{w} + \alpha\mathbf{a}_i$ is also consistent $\forall\alpha \in \mathbb{R}$.

P3) $\mathbf{w} + \sum_{i=1}^{M} \alpha_i\mathbf{a}_i$ is also consistent $\forall\alpha_i \in \mathbb{R}$.

P4) Each consistent edge weight vector $\mathbf{w}$ can be represented as a unique linear combination of any $M-1$ columns of $\mathbf{A}$, e.g. $\mathbf{w} = \sum_{i=1}^{M-1} \alpha_i \cdot \mathbf{a}_i$.

P5) If $\mathbf{w}_i$ are consistent, then $\sum_i \alpha_i\mathbf{w}_i$ is consistent as well $\forall\alpha_i \in \mathbb{R}$.

The proof of these properties are trivial using Eq. (4) and are hence dropped here. For the physical interpretations, we first consider the initial scenario in Fig. 3 with the source position $\mathbf{p}$, microphone positions $\mathbf{q}_i$, $(1 \leq i \leq M)$, time of arrival (TOA) $t_i = \frac{1}{c}\|\mathbf{q}_i - \mathbf{p}\|$ for microphone i, TDOA $w_{ij} = t_i - t_j$ for the microphone pair (i, j), and sound propagation velocity c. Let $\mathbf{t} = [t_1, \ldots, t_M]^T$ be the vector of all TOA values and $\mathbf{w}$ be the vector of TDOA values for all considered microphone pairs. Clearly, $\mathbf{w} = \mathbf{A}\mathbf{t}$ holds where $\mathbf{A}$ is the incidence matrix.

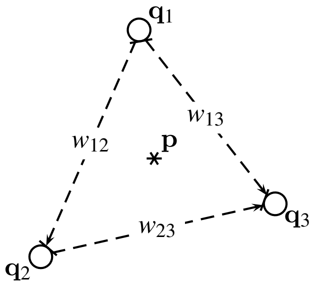


Figure 3: Initial scenario for TDOA based localization

P1) For $\alpha > 0$, we can interpret the property P1 either as a change in the speed of sound from c to c/α due to a different propagation medium or as a scaling of the source and microphone positions by α, see Fig. 4. In both cases, all TOA and TDOA values are scaled by α.

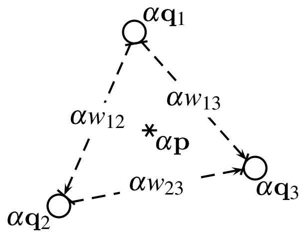


Figure 4: Scaling of the source and microphone positions by $\alpha > 0$

P2) An interpretation of P2 is a change in the distance of the i^{th} microphone to the source, which results in a change of the TOA from the source to microphone i. This happens frequently in reality and is known as the effect of mirrored microphone [1]. Fig. 5 illustrates this phenomenon. The echo path propagation from the source to microphone 2 due to a sound reflection on a wall enhances the TOA value from t_2 to $t_2 + \alpha$ ($\alpha > 0$). The effect is the replacement of microphone 2 by its mirror at position $\mathbf{q}_2'$. The TDOA vector caused by the microphone array at $\mathbf{q}_1, \mathbf{q}_2', \mathbf{q}_3$ becomes then $\mathbf{w} + \alpha\mathbf{a}_2$ where $\mathbf{w}$ is the TDOA vector caused by the original microphone array at $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$.

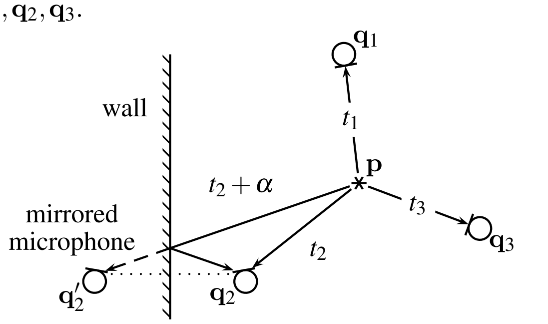


Figure 5: Change of the distance between one microphone and the source as caused by echo path propagation or mirrored microphone

P3) The property P3 is an extension of P2. In this case, the TOA from the source to all microphones has changed from t_i to $t_i + \alpha_i$. The underlying microphone array contains M mirrored microphones.

P4) Since $\mathbf{w} = \mathbf{At} \in R(\mathbf{A})$ and $\text{rank}(\mathbf{A}) = M - 1$, $\mathbf{w}$ can always be written as a linear combination of any $M - 1$ linearly independent columns of $\mathbf{A}$. Due to the structure of $\mathbf{A}$ one can easily verify that $\mathbf{A}[1, \ldots, 1]^T = \mathbf{a}_1 + \ldots + \mathbf{a}_M = 0$ and hence $\mathbf{a}_M = -(\mathbf{a}_1 + \ldots + \mathbf{a}_{M-1})$. This means $\mathbf{w} = \sum_{i=1}^{M} t_i \mathbf{a}_i = \sum_{i=1}^{M-1} t_i \mathbf{a}_i - t_M \sum_{i=1}^{M-1} \mathbf{a}_i = \sum_{i=1}^{M-1} (t_i - t_M) \mathbf{a}_i$. The physical interpretation of P4 is to move all microphones about the same distance $c \cdot t_M$ towards the source, assuming that t_M is the smallest TOA value.

P5) Assume that $\mathbf{w}_s$ is the TDOA vector of a source at position $\mathbf{p}_s$ ($s = 1, 2$) and $\mathbf{t}_s$ is the corresponding TOA vector, i.e. $\mathbf{w}_s = \mathbf{A}\mathbf{t}_s$. Then $\alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 = \mathbf{A}(\alpha_1\mathbf{t}_1 + \alpha_2\mathbf{t}_2)$ can be interpreted as a single source with respect to a new microphone array where the microphone distances to the source are given by $c(\alpha_1\mathbf{t}_1 + \alpha_2\mathbf{t}_2)$. The same apparently applies for $s > 2$.

The properties P1 to P5 show that there exist more consistent graphs than given sources. By using a synthesis algorithm of consistent graphs, we are able to detect and eliminate wrong TDOA vales in W_n which are never consisten in the sense of Eq. (4). But we cannot distinguish between true consistent weight vectors $\mathbf{w}$ that originate from true source and mircrophone positions and those false consistent weight vectors $\mathbf{w}$ which satisfy Eq. (4), but arise due to properties P1 to P5 and thus correspond to a modified microphone array. This problem is already known and a good solution to distinguish the true and false consistent TDOA vectors is the residual TDOA error $\|\mathbf{w} - \tilde{\mathbf{w}}\|$ proposed in[1], where $\tilde{\mathbf{w}}$ is the TDOA vector computed from the estimated source position and exact microphone positions. This measure is small if $\mathbf{w}$ is a true consistent vector, and large if $\mathbf{w}$ is not.

5 Efficient synthesis of consistent graphs

5.1 Problem formulation

The ambiguity problem in TDOA estimation described at the beginning of the paper is only one possible application for the synthesis problem of a consistent graph. There we have a setup of M microphones and obtain N weight sets W_n per microphone pair. Each weight set $W_n = \{w_{n,1}, \ldots, w_{n,k_n}\}$ contains $k_n = |W_n|$ TDOA values for the n^{th} microphone pair estimated by GCC or other methods. This leads to a graph with M vertices and N edges, that is not necessarily connected and may have partially wrong weight sets. The goal is to obtain such graphs with single edge weights $\mathbf{w} \in (W_1 \times \cdots \times W_N)$ that are consistent in the sense of Eq. (4). A brute-force approach would require a check of $\prod_{n=1}^{N} |W_n|$ different possibilities for $\mathbf{w}$ what is computationally too demanding. Hence we introduce here two different efficient synthesis strategies.

5.2 Efficient synthesis strategies

The combinatorial problem of finding weights that form a consistent graph can be solved by a top-down (TD) strategy, that uses sophisticated search procedures on the whole weight vector $\mathbf{w}$, and a bottom-up (BU) strategy, that checks the consistency of small subgraphs and merge them to a full consistent graph. Due to limited space, we will only briefly sketch both strategies. More details about the TD approach can be found in [7].

5.2.1 Top-Down strategy

The TD strategy is mainly based on Eq. (4) that provides an easy test for consistency, once all edges have been assigned a value. This approach contains the main steps in Fig. 6. Here one central step is to find the spanning tree of a connected graph or a connected partition of a given graph. To be more precise, a spanning tree is defined as a

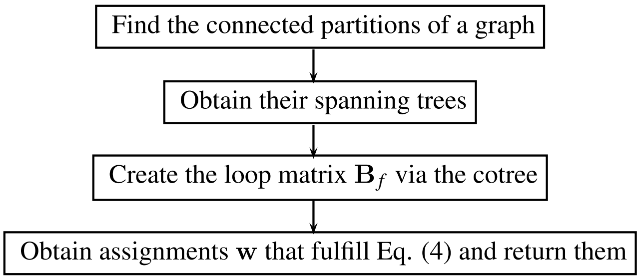


Figure 6: Main steps of the TD approach for synthesing consistent graphs

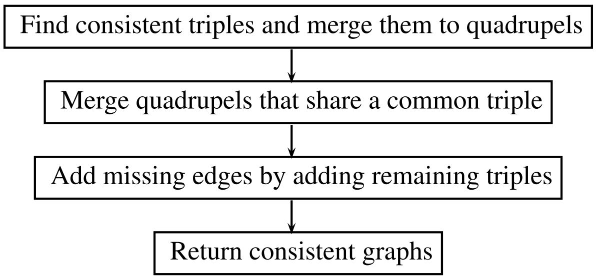


Figure 7: Main steps of the DATEMM algorithm

subgraph $T(V,\tilde{E})$ of $G(V,E)$ with $\tilde{E} \subseteq E$. It reaches all vertices of the graph without closing any loop. In 1971 Tarjan introduced an algorithm called depth-first search (DFS) or backtracking [8] that searches for the tree with least edges per vertex. This algorithm starts with a root vertex and searches for one of its neighbours. Then it starts a search for neighbours of this adjacent vertex and so on. If there is a final vertex with no neighbours but not all vertices have been visited, the algorithm tracks back to the previous vertex and searches there for further unvisited neighbours. This leads to the recursive algorithm 1 below. With the root vertex $v \in V$, it finds a spanning tree T.

Algorithm 1 depth-first search

```
[T, V] = DFS(V \ v, E, [], v)
if V ≠ ∅ then
   the graph is not connected
end if

function [T, V] = DFS(V, E, T, v)
for u ∈ V do
   if {v,u} ∈ E then
      T = T ∪ {v,u}
      [T, V] = DFS(V \ u, E, T, u)
   end if
end for
```

Another algorithm that we want to mention is the breadth-first search (BFS) by Dijkstra [9]. In a BFS we visit all neighbours of the root vertex first and then we search from every neighbour for unvisited vertices. This leads to a wide spanning tree.

As one can see, algorithm 1 already detects if a given graph is connected or not. It is easy to apply it recursively to a given graph to find the spanning trees of its connected partitions. This accomplishes the first two steps in Fig. 6. From then we easily follow the theory explained in section 3 to create the loop matrix $\mathbf{B}_f$.

The final step to find consistent weight vectors $\mathbf{w} \in (W_1 \times \cdots \times W_N)$ is explained in detail in [7].

5.2.2 Bottom-Up strategy

The consistency of subgraphs is a necessary condition for a consistent graph. This property is exploited by the BU strategy. Thus the graph is generated by merging the consistent subgraphs, usually triples, to graphs of higher order until most vertices are included. One example of a BU strategy is the DATEMM algorithm [3], that is summarized in Fig. 7. More details about the BU strategy will be discussed in the future due to limited space here.

5.2.3 Comparison and summary

The BU algorithms obtain partial results directly and can be stopped at every step. Moreover, the synthesis for several sources can be done in parallel. In contrast, the TD strategy can apply sophisticated search procedures on the weight sets. They need a preprocessing step to obtain the spanning tree but thereby one can examine the graph and exclude elements that do not contribute to the synthesis. This includes isolated vertices, single edges and partitions without any loop. Consequently, the combinatorial complexity can be reduced.

One problem for the TD strategy is that all edges need a valid weight for each consistent graph (source), because Eq. (4) will never be true for a non-complete, valid assignment. A more general question is what happens to weight sets W_n that have absolutely no valid weights. The BU algorithms also lack a discussion about the reuse of triples or subgraphs, i.e. if one should reuse triples in the synthesis procedure in order not to lose any consistent graphs. These questions well be answered in future research.

In this paper we have presented the connection between TDOA based source localization and consistent graphs. We showed how to transform the problem of finding consistent TDOA estimates to an abstract graph problem and found further properties that can be interpreted as real world phenomena. Moreover, we sketched two different synthesis approaches to consistent graphs.

Literatur

[1] J. Scheuing and B. Yang, “Disambiguation of TDOA estimation for multiple sources in reverberant environments,” *Transactions on Audio, Speech and Language Processing*, vol. 16, pp. 1479–1489, 2008.

[2] M. Brandstein and D. Ward, Eds., *Microphone Arrays*. Springer-Verlag, 2001.

[3] J. Scheuing and B. Yang, “Disambiguation of TDOA estimates in multi-path multi-source environments (DATEMM),” in *Proc. IEEE Conference on Acoustics, Speech, and Signal Processing*, 2006, pp. 837–840.

[4] J. Bang-Jensen and G. Gutin, *Digraphs - Theory, Algorithms and Applications*. Springer-Verlag, 2006.

[5] R. Diestel, *Graphentheorie*, 3rd ed. Berlin, Heidelberg: Springer-Verlag, 2006.

[6] N. Balabanian and T. A. Bickart, *Electrical Network Theory*. John Wiley & Sons, 1969.

[7] M. Kreißig and B. Yang, “Efficient synthesis of consistent graphs,” in *2010 European Signal Processing Conference*, Aalborg, Denmark, 2010.

[8] R. Tarjan, “Depth-first search and linear graph algorithms,” in *Switching and Automata Theory, 12th Annual Symposium*, vol. 1, 1971, pp. 114–121.

[9] E. W. Dijkstra, “A note on two problems in connexion with graphs,” *Numerische Mathematik*, vol. 1, pp. 269–271, 1959.