SEISMIC FEATURE EXTRACTION USING STEINER TREE METHODS

Ludwig Schmidt, Chinmay Hegde, Piotr Indyk

Massachusetts Insitute of Technology
Cambridge, MA 02139

Ligang Lu, Xingang Chi, Detlef Hohl

Shell International E&P, Inc.
Houston, TX 77082

ABSTRACT

Identifying "interesting" features, such as faults, unconformities, and other events in subsurface images is a challenging task in seismic data processing. Existing state-of-the-art methods usually involve manual intervention in the form of a visual inspection by an expert, but this is time-consuming, expensive, and error-prone. In this paper, we propose an efficient, *automatic* approach for seismic feature extraction. The core idea of our approach involves interpreting a given 2D seismic image as a function defined over the vertices of a specially chosen underlying graph. This enables us to formulate the feature extraction task as an instance of the *Prize-Collecting Steiner Tree* problem encountered in combinatorial optimization. We develop an efficient algorithm to solve this problem, and demonstrate the utility of our method on a number of synthetic and real examples.

***Index Terms*—** Seismic signal processing, Prize Collecting Steiner Tree problem, combinatorial optimization.

1. INTRODUCTION

A key goal in seismic data processing is the detection of geologically interesting *features* in seismic images. Subsurface features, such as faults, unconformities, and other events contain valuable spatio-temporal information which can be of both scientific as well as commercial importance. Therefore, there is a compelling need to accurately and efficiently locate such features in images.

However, the prevailing method for feature extraction in seismic images involves manual intervention in the form of visual interpretation and pixel labeling by an expert. This is time-consuming and expensive. On the other hand, automatic (algorithm-based) approaches for seismic feature extraction suffer from several pitfalls themselves. Subsurface features are of varied shapes and sizes, and parametric modeling approaches often do not work well. Moreover, seismic datasets are typically high-dimensional, and complicated algorithmic approaches are expensive from a computational standpoint. Finally, seismic images contain high amounts of noise, and this further degrades the performance of parametric approaches.

In this paper, we propose a simple and efficient algorithm for seismic image analysis and feature extraction. At a high level, the key idea involves formulating the seismic feature extraction task as a *combinatorial optimization* problem. We interpret a 2D seismic image as a function defined over the vertices of an underlying graph (the choice of this graph is flexible). Formally, an $n \times n$ image can be viewed as a scalar-valued function f defined, for instance, over a *grid* graph $G = (V, E)$ containing n^2 vertices and $\Theta(n^2)$ edges. We focus our attention on features that roughly correspond to nodes v with large function values $f(v)$. Moreover, we assume that the significant coefficients of f form a small number of *connected components* relative to the edge structure of G, since interesting features such as unconformities correspond to pixels that are spatially contiguous in the seismic image. Finally, we assume that there is only a (relatively) small number of such strong, connected features. We make no further assumptions on the shapes and / or sizes of the features.

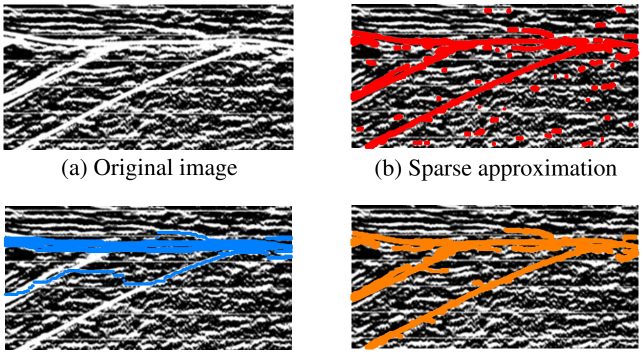
With these assumptions, we formally pose the feature extraction task as the following combinatorial optimization problem: given a function $f : V \to \mathbb{R}_0^+$, find a subset $T \subset V$ such that $|T|$ is small, T constitutes the union of a small number of connected components (features), and $f(V_T) = \sum_{v \in T} f(v)$ is maximized. Observe that if we ignore the connectedness property, this simply reduces to the problem of *sparse approximation* and can be solved by identifying the locations of the largest coefficients of f. The connectedness assumption makes the problem challenging; in fact, the problem of detecting even a single dominant component containing a given number of nodes in an arbitrary graph is known to be NP-hard [1].

Despite this hardness result, efficient algorithms with provable approximation guarantees exist for variants of this problem. In particular, we focus on the *Prize-Collecting Steiner Tree* (PCST) problem [2], which is a generalization of the classical *Steiner tree* problem. At a high level, the goal in the PCST problem is to find a subtree T that balances the cost of connecting some nodes and the penalty paid for omitting the remaining nodes from the solution T. The PCST problem has been well-studied in graph optimization and several algorithms have been proposed, including the seminal primal-dual scheme of Goemans and Williamson [3].

The PCST problem only asks to find a single tree T, and not a set of disjoint connected components (features). Therefore, we consider an *augmented* graph that lets us incorporate the number of features into the problem formulation. We achieve this by connecting all nodes in the image graph to a new *root* node and assigning a special edge weight to the newly created edges. The augmented graph has two nonnegative parameters that control the trade-off between the three competing objectives: (i) the signal energy captured by the features, (ii) the number of discovered features, and (iii) the area covered by the features. By tuning these two parameters, we can arrive at a suitable solution. See Fig. 1 for a representative example.

We conduct experiments to demonstrate that our algorithm can recover several types of interesting features in seismic images. In particular, we are able to identify complicated structures (such as unconformities), which was not possible using previous methods such as [4]. Moreover, our proposed method is computationally efficient. Finally, the method is fairly modular and can be used in conjunction with other pre- and post-processing techniques.

The rest of this paper is organized as follows. Section 2 presents a brief overview of existing methods, including a description of the Prize-Collecting Steiner Tree (PCST) problem and associated algorithms. Section 3 provides details about our proposed approach. Finally, Section 4 describes the results of numerical experiments on synthetic and real test cases, and Section 5 concludes the paper.



(a) Original image

(b) Sparse approximation

(c) EMD [4]

(d) Proposed method

Fig. 1. *Unconformity detection on a section of a real seismic image corresponding to a section of the southeastern Malay basin [5]. (a) The original image contains a single angular unconformity. (b) Conventional sparse approximation identifies the pixels corresponding to the location of the unconformity, but also introduces outliers (false positives). (c) The approach of [4] identifies the unconformity but "short-circuits" different sloping layers. (d) Our proposed Prize-Collecting Steiner Tree (PCST)-based approach provides a more faithful feature identification.*

2. BACKGROUND

2.1. Seismic image analysis

Feature identification in unmigrated trace data, as well as stacked images, is an important task in the seismic data processing pipeline. Typical features of interest include *faults* (shear-like discontinuities caused by the approximately vertical relative movement of rock layers) and *unconformities* (erosional surfaces separating strata of different geological ages, representing discontinous sediment deposition). Both kinds of features contain valuable spatio-temporal information and accurate localization of such features is of considerable interest.

A host of automatic approaches for extraction of both types of features have been proposed in the literature, and we highlight a few representative works (for a more complete list of references, see the recent papers [6, 7]). In the context of fault extraction, these include texture classification approaches [8], coherence-based methods [9], and evolutionary computation-based search heuristics [10]. In the context of unconformity extraction, these include edge-detection based methods [11] and structure-tensor fields [7]. Most of these methods seem to suffer from a low tolerance to noise, high computational complexity, or both.

2.2. Fault Localization using the Earth Mover's Distance

In previous work [4], we proposed an alternate algorithm for automatic fault localization, borrowing ideas from recent advances in signal processing. This algorithm is based on two key assumptions: (i) the seismic image contains s features in each column, i.e., only a fraction of the pixels in each column correspond to dominant reflectors; and (ii) the locations of the features in the columns vary smoothly in terms of the *Earth Mover's Distance* (EMD), except at the location of the faults. Under these assumptions, the method poses the feature extraction task as a combinatorial approximation problem that can be solved using a minimum-cost flow routine over a specially defined graph. Experiments on both (noisy) synthetic images as well as real 2D and 3D images confirmed the robustness of the method.

However, this method suffers from drawbacks. The algorithm assumes a constant number of features per column across the entire image. Consequently, it is unable to capture more complicated patterns, such as unconformities, where the number of features per column varies across different regions. In Section 3 below, we propose a new method for seismic feature extraction that is also graph-based in nature, but that improves over our previous method. Our new method is flexible enough to handle images with a spatially varying number of features per column, and we successfully use this property to demonstrate recovery of unconformities in noisy seismic images.

2.3. The Prize-Collecting Steiner Tree problem

Our method leverages efficient algorithms for the *Prize Collecting Steiner Tree* (PCST) problem. We briefly review the problem and associated algorithms here for completeness and refer to the textbook [12] for a detailed description.

First, we introduce necessary notation. Let $G = (V, E)$ be a given undirected graph. For any function $f : V \to \mathbb{R}$ and node subset $U \subseteq V$, we define $f(U) := \sum_{v \in U} f(v)$. Moreover, we define the complement of U with respect to V as $\overline{V} := V \setminus U$. We use the same conventions for functions defined on the edge set E and for subsets of E.

The goal of the PCST problem is to find a "good" subgraph $T = (V_T, E_T)$ in G, where the objective function is defined as follows. Let $\pi : V \to \mathbb{R}_0^+$ be a nonnegative function representing the *prize* associated with each vertex $v \in V$. The prize of a vertex $\pi(v)$ denotes the cost of excluding v in the solution subgraph T. Furthermore, let $c : E \to \mathbb{R}_0^+$ be a nonnegative function representing the *cost* associated with each edge. The cost of an edge e denotes the cost of adding e to the solution T. Using these definitions, the goal of the PCST problem is to find a connected subgraph T of G that minimizes the prizes of nodes *not* in T and the costs of edges *in* T. This formulation naturally captures the trade-off between connecting a node to the solution T (and hence paying the cost of the respective edge) and omitting a node from T (and hence paying its prize).

Trivially, one can observe that any optimal solution subgraph T_{OPT} is a *tree*; if it were not, then one could simply choose the minimum spanning tree of T_{OPT} as a candidate solution with fewer edges. Therefore, our goal is to find the optimal subtree:

$$T_{OPT} = \operatorname*{arg\,min}_{T \text{ is a tree}} c(T) + \pi(\overline{T}) ,$$

where we use T to refer to either the node set V_T or the edge set E_T, depending on context. Since this problem is a generalization of the classical Steiner tree problem, it is also NP-hard [13]. Nevertheless, there exist several efficient *approximation* algorithms. Many of these algorithms are variants of the seminal work of Goemans and Williamson [3], who developed a primal-dual algorithm that returns a subtree T with the following approximation guarantee:

$$c(T) + 2\pi(\overline{T}) \leq 2c(T_{OPT}) + 2\pi(\overline{T_{OPT}}) .$$

For a graph with n vertices, the algorithm by Goemans and Williamson has a time complexity of $O(n^2 \log n)$ for a special (rooted) variant of the PCST problem, and subsequent improvements [14, 15] achieve this running time for the case of general undirected graphs.

3. PROPOSED ALGORITHM

We now describe our feature extraction algorithm in detail. At a high level, our approach involves converting the feature extraction task to a specific instance of the PCST problem.

First, we formalize the feature extraction task. Suppose we are given an image $I \in \mathbb{R}^{n \times m}$ and an underlying graph structure $G = (V, E)$ on the set of pixel indices, i.e., $V = [n] \times [m]$. In this paper, we consider G to be a grid graph, although the algorithm below works for arbitrary graphs. The objective is to find a set of features $\mathcal{F} = \{F_1, \ldots, F_c\}$ where each feature is a *connected subset* of the image, i.e., $F_i \subseteq [n] \times [m]$ and the pixel indices in F_i form a connected component in G. Our goals are as follows:

- The image coefficients corresponding to features should capture a significant amount of energy in the signal, i.e., we want $\sum_{i=1}^{k} \|I_{F_i}\|_2^2$ to be large. This corresponds to the assumption that the "interesting" features in the image (e.g., faults or unconformities) correspond to pixels with high intensities.
- The total number of features, $|\mathcal{F}|$, should be small. This corresponds to the assumption that there is only a relatively small number of strong features in the image of interest.
- The features should only cover a limited part of the image, i.e., $\sum_{i=1}^{k} |F_i|$ should be small. This corresponds to the assumption that the total area covered by the most important features is typically only a small fraction of the image.

We now formulate the feature extraction task as an instance of the PCST problem. First, we observe that the PCST problem directly allows us to address the first and third objectives: feature energy and feature size. By assigning the corresponding (squared) coefficient intensity value to each node in the graph, the PCST objective function ensures that the solution T contains many of the significant coefficients; otherwise, the solution would incur a large cost in the node-prize term $\pi(\overline{T})$. Moreover, assigning a nonzero cost λ to every edge favors sets of features that cover a small total area since adding a new pixel to the solution incurs cost λ.

Next, we incorporate the second objective (total number of features) in the PCST instance. Unfortunately, this objective is not directly captured by the conventional PCST formulation, which only seeks a *single* subtree T (and not a union of trees). However, we can add the constraint by augmenting the image graph G with a special *root node* which is connected to all other nodes in the image. We assign a special edge cost γ to these edges, which can be different from λ, the cost of "normal" graph edges. See Fig. 2 for an illustration of this construction. We observe that a PCST solution in this augmented graph can give rise to several features which are disconnected in the original graph G (see Fig. 3). By adjusting the weight γ of the edges incident to the root node, we can control the cost of creating a new feature in the PCST solution. Formally, we obtain:

Definition 1 (Feature-detection PCST instance). *Let $I \in \mathbb{R}^{n \times m}$ be an $n \times m$ image and let $G = (V, E)$ be the pixel connectivity structure for the image, i.e., $V = [n] \times [m]$. Then a feature-detection PCST instance is constructed as follows: Let $G' := (V', E')$ be a graph with $V' := V \cup \{r\}$, where r is the root node. Set $E' := E \cup \{(r, v) \,|\, v \in V\}$. For any $v \in V'$ corresponding to an image pixel, set $\pi((i,j)) := I_{i,j}^2$. For the root node, set $\pi(r) := +\infty$. For all original image edges $e \in E$, set the edge cost $c(e) := \lambda$. For all edges incident to the root node, set the edge cost $c((r, v)) := \gamma$ for all $v \in V$.*

Solving the PCST problem on a feature detection instance as defined above gives us a tree T in the graph G'. We can now convert T to a set of features by simply removing the root node r, which partitions the tree T into connected components. We then return each connected component as a "feature" in the image.

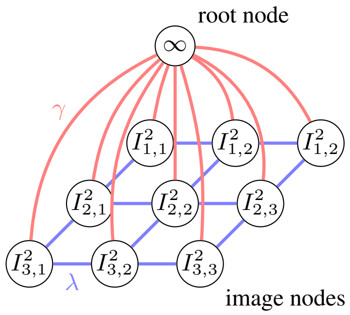


Fig. 2. *The PCST instance for a 3×3-pixel image with intensities $I_{i,j}$ and a simple grid graph as connectivity structure G. The prize of each node is the corresponding intensity squared, i.e., $I_{i,j}^2$. The costs of all grid edges (blue) are λ and the costs of all root edges (red) are γ. Most of the edge labels are omitted for clarity. The root effectively has cost $+\infty$ and hence is always included in the solution.*

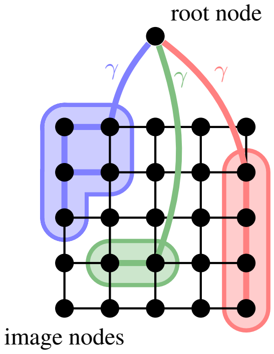


Fig. 3. *A PCST solution in a 5×5-pixel image on a grid graph. The solution decomposes into three features: the red, green, and blue components. Every feature has a single edge to the root node and hence contributes γ to the overall solution cost. The other root edges and most node / edge labels are omitted for clarity.*

Finally, it remains to relate the PCST objective function to the goals in the feature extraction task. If we set $\gamma' := \lambda + \gamma$ in the PCST instance, then we can derive the following equalities:

$$c(T) = \lambda \sum_{F \in \mathcal{F}} |F| + \gamma' |\mathcal{F}|$$

$$\pi(\overline{T}) = \|I\|_2^2 - \sum_{F \in \mathcal{F}} \|I_F\|_2^2 \, .$$

Therefore, minimizing the PCST objective function $c(T) + \pi(\overline{T})$ captures all three goals outlined above: the energy of the image coefficients in the features, the total area covered by the features, and the number of features identified.

To summarize, our algorithm for extracting "interesting" features proceeds as follows: (i) For a given input image I and edge weights λ and γ, construct a feature-detection PCST instance (Definition 1). (ii) Solve the instance using an efficient PCST algorithm. (iii) Repeat steps (i) and (ii) after appropriately tuning λ and γ until a suitable set of image indices is identified as the set of features.

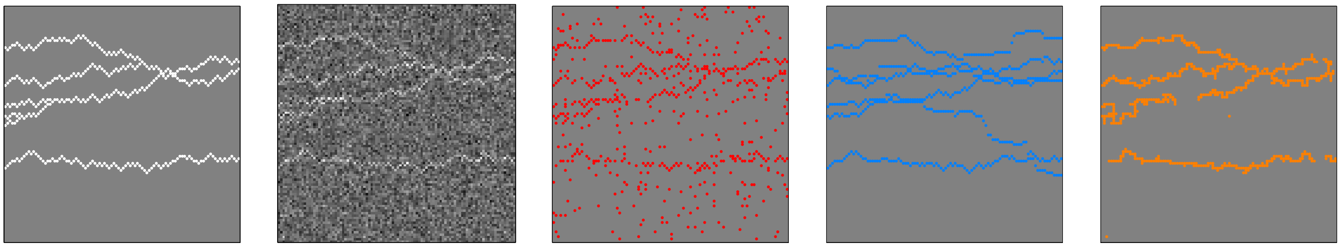

(a) Original image (b) Noisy input (c) Sparse approximation (d) EMD [4] (e) Proposed method

Fig. 4. *Feature extraction on a synthetic image. (a) Original image. (b) The image is contaminated with a heavy amount of noise (SNR = -5dB). (c) Feature extraction via sparse approximation introduces several isolated false positives. (d) Feature extraction via the EMD-based approach [4] introduces spurious linear features. (e) Feature extraction via our proposed method robustly recovers the correct features.*

4. NUMERICAL EXPERIMENTS

We now present the results of several experiments on synthetic and real test images to demonstrate the benefits of our proposed PCST-based algorithm. In our experiments, we performed parameter tuning to achieve the best visually acceptable performance for our proposed method, sparse approximation, and the method of [4]. First, we study the performance of our algorithm in the context of feature extraction from *synthetic* images. Figure 4 displays the results of a numerical experiment detailing the performance of different algorithms on a randomly generated test image of size 100 × 100. (Fig. 4(a)). The image is contaminated with a large amount of i.i.d. Gaussian noise (equivalent to -5dB), resulting in Fig. 4(b). We use this as the input to three different algorithms – sparse approximation; the Earth Movers Distance (EMD)-based approach of [4]; and our proposed PCST-based algorithm (with parameters $\lambda = 0.75, \gamma = 4$). We depict the locations of the detected feature pixels by each algorithm in Figs. 4(c,d,e).

While both the sparse approximation (Fig. 4(c)) and the EMD-based approach (Fig. 4(d)) perform reasonably well in identifiying the features, both lead to spurious outliers. The sparse approximation method ignores the connectivity constraint and therefore produces several isolated false positives. On the other hand, the EMD based method of [4] assumes a constant number of features per column and therefore hallucinates two spurious layer features. In contrast, our proposed method leads to stable recovery of the original features in (a) with very few false positives or negatives.

Next, we study the performance of our method on a real seismic image. Figure 1(a) displays the test image (of resolution 150 × 300) corresponding to a section of the Southeastern Malay Basin [5]. It is visually evident that the image contains a single dominant unconformity feature, and is corrupted by a considerable amount of noise. In order to mitigate the noise, we first pre-process the input image via a median filter with window size 5×5. Then, we use this denoised image and apply both sparse approximation (with 10% sparsity) as well as our proposed PCST-based method (with parameters $\lambda = 1.5, \gamma = 22$). Again, our proposed method is able to successfully recover the unconformity without outliers.

Finally, we study the performance of our algorithm on a second real seismic image containing a variety of faults, unconformities, and other events. The test image is of resolution 200×200 and displayed as Fig. 5(a). Again, we first use a median denoising pre-processing step and then run various feature identification methods. Our method (with parameters $\lambda = 0.25, \gamma = 1.7$) recovers several features in the image and is able to ignore the noise (present in the bottom right part of the image). In contrast, both conventional sparse approximation and the EMD-based method of [4] pick up several spurious outliers.

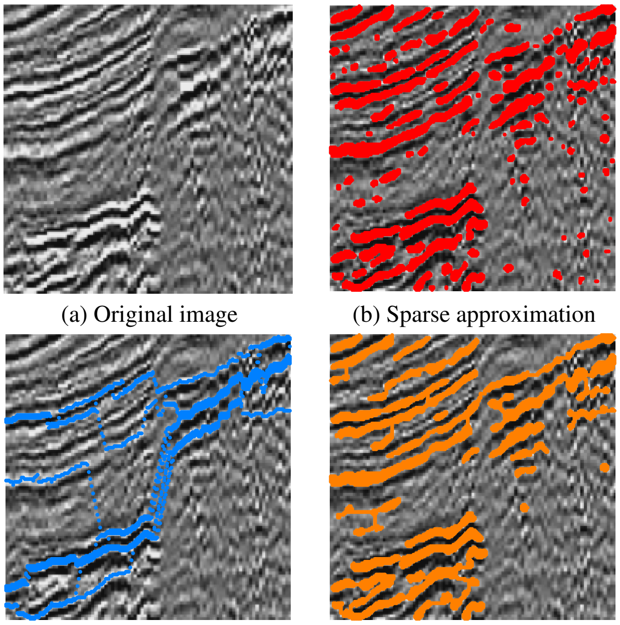

(a) Original image (b) Sparse approximation
(c) EMD [4] (d) Proposed method

Fig. 5. *Feature extraction on (a) section of a real seismic image. (b) The sparse approximation approach produces several spurious outliers around small clusters of high-intensity pixels that do not always represent important features in the image. (c) The EMD-based approach [4] fails to identify several significant layers and incorrectly connects some of the layers. (d) In contrast, our proposed method correctly identifies most of the important features.*

5. CONCLUSIONS

In this paper, we have proposed a new algorithm for extraction of "interesting" features in seismic images. Our approach involves formulating the feature extraction task as an instance of the Prize-Collecting Steiner Tree (PCST) problem, which we can solve using efficient existing methods. We have demonstrated its applicability to a number of synthetic and real test seismic images. We mention in passing that while we focus exclusively on 2D seismic images in this paper, our techniques are potentially applicable for other signal and image processing domains.

6. REFERENCES

[1] T. Ideker, O. Ozier, B. Schwikowski, and A. Siegel, "Discovering regulatory and signalling circuits in molecular interaction networks," *Bioinformatics*, 2002.

[2] D. Bienstock, M. Goemans, D. Simchi-Levi, and D. Williamson, "A note on the prize collecting traveling salesman problem," *Mathematical Programming*, 1993.

[3] M. Goemans and D. Williamson, "A general approximation technique for constrained forest problems," *SIAM Journal on Computing*, 1995.

[4] L. Schmidt, C. Hegde, P. Indyk, J. Kane, L. Lu, and D. Hohl, "Automatic fault localization using the Generalized Earth Movers Distance," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, 2014.

[5] K. Ngah, "Structural framework of Southeastern Malay Basin," *Search and Discovery*, 2000.

[6] D. Hale, "Methods to compute fault images, extract fault surfaces, and estimate fault throws from 3D seismic images," *Geophysics*, vol. 78, no. 2, pp. O33–O43, 2013.

[7] X. Wu and D. Hale, "3D seismic image processing for unconformities," Tech. Rep. CWP-813, Colorado School of Mines, 2014.

[8] T. Randen, E. Monsen, C. Signer, A. Abrahamsen, J. Hansen, T. Sæter, and J. Schlaf, "Three-dimensional texture attributes for seismic data analysis," in *SEG Annual Meeting*, 2000.

[9] K. Marfurt, R. Kirlin, S. Farmer, and M. Bahorich, "3D seismic attributes using a semblance-based coherency algorithm," *Geophysics*, vol. 63, no. 4, pp. 1150–1165, 1998.

[10] S. Pederson, T. Skov, T. Randen, and L. Sønneland, "Automatic fault extraction using artificial ants," *Math. Methods and Modeling in Hydrocarbon Explor.*, vol. 7, pp. 107–116, 2005.

[11] T. Randen, S. Pederson, C. Signer, and L. Sønneland, "Image processing tools for geological unconformity extraction," in *Proc. IEEE Nordic Signal Processing Symposium*, 1999.

[12] D. Williamson and D. Shmoys, *The Design of Approximation Algorithms*, 2011.

[13] Richard Karp, "Reducibility among combinatorial problems," in *Complexity of Computer Computations*, The IBM Research Symposia Series, pp. 85–103. 1972.

[14] D. Johnson, M. Minkoff, and S. Phillips, "The prize collecting Steiner tree problem: Theory and practice," in *SODA*, 2000.

[15] P. Feofiloff, C. G. Fernandes, C. E. Ferreira, and J. Coelho de Pina, "A note on Johnson, Minkoff and Phillips' algorithm for the prize-collecting Steiner tree problem," *CoRR*, vol. abs/1004.1437, 2010.