

Autonomous Positioning Techniques Based on Cramér-Rao Lower Bound Analysis

Mats Rydström,¹ Andreu Urruela,² Erik G. Ström,¹ and Arne Svensson¹

¹Department of Signals and Systems, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

²Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain

Received 31 May 2005; Revised 6 October 2005; Accepted 11 October 2005

We consider the problem of autonomously locating a number of asynchronous sensor nodes in a wireless network. A strong focus lies on reducing the processing resources needed to solve the relative positioning problem, an issue of great interest in resource-constrained wireless sensor networks. In the first part of the paper, based on a well-known derivation of the Cramér-Rao lower bound for the asynchronous sensor positioning problem, we are able to construct optimal preprocessing methods for sensor clock-offset cancellation. A cancellation of unknown clock-offsets from the asynchronous positioning problem reduces processing requirements, and, under certain reasonable assumptions, allows for statistically efficient distributed positioning algorithms. Cramér-Rao lower bound theory may also be used for estimating the performance of a positioning algorithm. In the second part of this paper, we exploit this property in developing a distributed algorithm, where the global positioning problem is solved suboptimally, using a divide-and-conquer approach of low complexity. The performance of this suboptimal algorithm is evaluated through computer simulation, and compared to previously published algorithms.

Copyright © 2006 Hindawi Publishing Corporation. All rights reserved.

1. INTRODUCTION

Large-scale wireless sensor networks (WSNs) have been proposed for a multitude of applications ranging from passive information gathered in remote and/or hostile environments to active automotive safety applications. Many interesting problems arise from implementation aspects, for instance, hard constraints on resources such as battery capacity, bandwidth, or production cost. A “must-have” property of many WSNs is the ability to autonomously position individual nodes, that is, without relying on surrounding fixed infrastructure, such as beacons, base-stations, or satellites. Autonomous positioning algorithms have been proposed based on a number of different techniques, where, recently, time-of-flight (ToF) based techniques have seen the most attention, see, for example, [1, 2], and the references cited therein. Positioning technique based on ToF measurements between nodes is made more complicated if the nodes cannot be assumed synchronized in time, a property not feasible in large-scale sensor networks. Further, the complexity of the autonomous positioning problem grows rapidly as sensor networks scale in number of nodes and/or connectivity.

The Cramér-Rao lower bound (CRB) is a lower bound on the variance of all unbiased estimators which can be derived for most estimation problems. In this paper, we employ CRB

theory for two reasons. First, it offers a measure of how accurately we can estimate a set of unknown parameters, given a vector of measurements. This measure is useful in positioning algorithm design, since it allows us to investigate the effects on the best possible performance, in a mean-squared-error (MSE) sense, of an estimator, if we first transform the measurement vector in some way. Also, it offers a practical performance estimator, intuitive and easy to implement, that can be used in positioning algorithms.

Drawing on CRB theory, we present two methods of optimally canceling a set of unknown clock-offsets from the autonomous relative coordinate estimation problem. Optimality is measured with respect to the Fisher information [3] of the relative coordinate estimation problem. The main reason for wanting to cancel unknown clocks from the problem is, of course, a reduction in the unknown parameter space. The first method, called the QR-method, building on previous work in [4], is suited for a centralized approach to the relative coordinate estimation problem, while the second clock-cancellation method, based on the assumption that measurement noise variance is similar on forward and reverse channels between two nodes, is well suited for distributed positioning algorithms.

In some WSN applications, such as environmental monitoring, accuracies predicted by the CRB are often not needed

to fulfill the requirements of the served application. Positioning algorithms operating in resource-constrained WSNs should therefore only spend as much processing resources as needed in order to fulfill the current requirements of the served application. Based on a performance estimator, given by the CRB computed in estimated node coordinates, we, based on work in [5], implement a suboptimal algorithm of low complexity, employing a divide-and-conquer approach, that is capable of increasing its positioning accuracy step-wise, conserving energy in scenarios where demands on accuracy are varying and the full power of statistically efficient estimators is not needed.

2. PROBLEM FORMULATION

Given a set of asynchronous internode delay measurements between sensor nodes, we wish to infer the relative two-dimensional layout of the wireless sensor network.

2.1. Signal model

We assume global node identification, similar to the unique addressing of Ethernet network interface cards, is available for each node in the WSN. If one arbitrary node transmits a message containing its node ID, all other nodes in range of the transmitting node can measure the arrival time of this message, relative to their local clocks. Since no synchronization is assumed between nodes, each delay measurement will be affected by unknown clock-offsets at both transmitting and receiving nodes. In this framework, an asynchronous delay, or pseudo-time-of-arrival (pTOA), measurement between nodes i and j , measured at node j , can be written

$$\tau_{i,j} = \Delta_i - \Delta_j + d(\mathbf{x}_i, \mathbf{x}_j)/c + v_{i,j}, \quad (1)$$

where Δ_n is the unknown clock-offset of node n , $d(\mathbf{x}_i, \mathbf{x}_j)$ is the distance between nodes i and j , in meters, as a function of their relative node coordinates $\mathbf{x}_n = [x_n \ y_n]^T$, c is the electromagnetic propagation speed, and $v_{i,j}$ is zero-mean Gaussian noise with variance $\sigma_{i,j}^2$, where we assume $\sigma_{i,j}^2$ is known. The assumption of Gaussian measurement noise with known variance greatly simplifies further developments. Most concepts described in this work are, however, applicable also in the case of non-Gaussian noise with unknown variance.

The clock-offsets Δ_n are considered as deterministic, albeit unknown, and measured relative to one node in the network, an N node network therefore has at most $N - 1$ unknown clock-offsets. Without loss of generality, we take the clock-offset of node 1 to be the reference clock in the network and set it equal to zero. Further, the clock-drift is assumed to be negligible in the relatively short period between node ID transmissions. We also assume, that if a pTOA measurement between nodes i and j is successfully made, the nodes are in transmission range of each other, and therefore the reverse pTOA measurement between nodes j and i is also available.

We denote the two pTOA measurements made between two nodes a pTOA measurement pair. It should be noted that the number of internode distances in a network of N nodes is $N(N - 1)/2$, so that the maximum number of pTOA pairs

in a network of N nodes is $M_{\max} = N(N - 1)/2$. From (1), assuming M pTOA measurement pairs have been made in a network of N nodes, we can write the pair-wise ordered pTOA measurement vector as

$$\boldsymbol{\tau} = \mathbf{H}_t \boldsymbol{\Delta} + \mathbf{H}_d \mathbf{d}(\mathbf{x}) + \mathbf{v} \in \mathbb{R}^{2M}, \quad (2)$$

where

$$\boldsymbol{\tau} = \left[\tau_{i(1),j(1)} \ \tau_{j(1),i(1)} \ \cdots \ \tau_{i(M),j(M)} \ \tau_{j(M),i(M)} \right]^T. \quad (3)$$

The indexing functions $i(k)$ and $j(k)$ denote the transmitting and receiving nodes of the k th pTOA measurement pair. Further,

$$\boldsymbol{\Delta} = [\Delta_2 \ \cdots \ \Delta_N]^T \in \mathbb{R}^{N-1}, \quad (4)$$

$$\mathbf{H}_t = \mathbf{H}_0 \otimes \begin{bmatrix} 1 \\ -1 \end{bmatrix} \in \mathbb{R}^{2M \times (N-1)}, \quad (5)$$

$$\mathbf{H}_0 = [\mathbf{h}_{t_1} \ \mathbf{h}_{t_2} \ \cdots \ \mathbf{h}_{t_M}]^T \in \mathbb{R}^{M \times (N-1)}, \quad (6)$$

$$\mathbf{H}_d = \frac{1}{c} \mathbf{I}_M \otimes [1 \ 1]^T \in \mathbb{R}^{2M \times M}, \quad (7)$$

$$\mathbf{d}(\mathbf{x}) = [d(\mathbf{x}_{i(1)}, \mathbf{x}_{j(1)}) \ \cdots \ d(\mathbf{x}_{i(M)}, \mathbf{x}_{j(M)})]^T \in \mathbb{R}^M,$$

$$\mathbf{x} = [\mathbf{x}_1^T \ \mathbf{x}_2^T \ \cdots \ \mathbf{x}_N^T]^T \in \mathbb{R}^{2N}, \quad (8)$$

$$\mathbf{x}_l = [x_l \ y_l]^T \in \mathbb{R}^2.$$

The matrix \mathbf{I}_M denotes the $M \times M$ identity matrix and \otimes denotes the Kronecker product. Each column vector $\mathbf{h}_{t_n} \in \mathbb{R}^{N-1}$, in (6), selects one or two clock-offsets from $\boldsymbol{\Delta}$ for each element in $\boldsymbol{\tau}$ according to (1), that is, for $\tau_{1,j(l)}$ and $\tau_{i(k),1}$, $[\mathbf{h}_{t_l}]_{j(l)-1} = -1$ and $[\mathbf{h}_{t_k}]_{i(k)-1} = 1$, respectively, with zeros elsewhere, because node 1 is the clock-reference node. If node 1 is not involved in pair n , \mathbf{h}_{t_n} would select two clock-offsets, with opposite signs, from $\boldsymbol{\Delta}$, that is, for $\tau_{i(n),j(n)}$ $[\mathbf{h}_{t_n}]_{i(n)-1} = 1$, $[\mathbf{h}_{t_n}]_{j(n)-1} = -1$, with zeros elsewhere. Because $\boldsymbol{\tau}$ is pair-wise ordered, the $(2n)$ th row of \mathbf{H}_t will be the negative of the $(2n - 1)$ th row. If some measurement pair m cannot for some reason be obtained and is missing from $\boldsymbol{\tau}$, that is, $M < M_{\max}$, the corresponding internode distance element in \mathbf{d} is removed, and the dimension of \mathbf{H}_d and \mathbf{H}_t is reduced accordingly. Also, if synchronized nodes are present in the network, the dimensions of $\boldsymbol{\Delta}$ and \mathbf{H}_t are reduced, that is, vector $\boldsymbol{\Delta}$ will always contain only unknown parameters. The measurement noise \mathbf{v} is assumed zero-mean Gaussian with covariance matrix \mathbf{V} . We assume \mathbf{V} is known. If \mathbf{V} is not known, it will have to be estimated with a possible degradation in estimation performance as a consequence. We further assume \mathbf{V} to be symmetric and positive definite, which will always be true for nondeterministic pTOA measurements.

2.2. Cramér-Rao lower bound

Due to the Gaussian properties of $\boldsymbol{\tau}$, and our assumptions on \mathbf{V} , the Fisher information matrix \mathbf{J} of (2) is given as [3],

$$[\mathbf{J}(\mathbf{z})]_{i,j} = \left[\frac{\partial \boldsymbol{\mu}_{\boldsymbol{\tau}}(\boldsymbol{\Delta}, \mathbf{x})}{\partial z_i} \right]^T \mathbf{V}^{-1} \left[\frac{\partial \boldsymbol{\mu}_{\boldsymbol{\tau}}(\boldsymbol{\Delta}, \mathbf{x})}{\partial z_j} \right], \quad (9)$$

where $\mathbf{z} = [\Delta^T \mathbf{x}_u^T]^T$, vector \mathbf{x}_u contain the unknown elements in \mathbf{x} , and $\boldsymbol{\mu}_\tau = \mathbb{E}[\boldsymbol{\tau}] = \mathbf{H}_t \Delta + \mathbf{H}_d \mathbf{d}(\mathbf{x})$. Partial derivatives are evaluated at the true value of $[\Delta^T \mathbf{x}^T]^T$. The CRB on the variance of any unbiased estimator of unknown relative node coordinates and unknown clock-offsets, based on a set of measured pTOAs as modeled by (2), given as the inverse of the Fisher information matrix \mathbf{J} , is therefore

$$\text{Var} \left(\begin{bmatrix} \hat{\Delta} \\ \hat{\mathbf{x}}_u \end{bmatrix} \right) \geq \left[\begin{bmatrix} \mathbf{H}_t^T \\ [\mathbf{H}_d \nabla \mathbf{d}]^T \end{bmatrix} \mathbf{V}^{-1} \begin{bmatrix} \mathbf{H}_t & \mathbf{H}_d \nabla \mathbf{d} \end{bmatrix} \right]^{-1}, \quad (10)$$

where $\text{Var}(\mathbf{x}) = \mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T]$, a matrix inequality on the form $\mathbf{M}_1 \geq \mathbf{M}_2$ should be interpreted as $\mathbf{M}_1 - \mathbf{M}_2$ being nonnegative definite, and the matrix $\nabla \mathbf{d} \in \mathbb{R}^{M \times 2N-3}$, assuming $2N - 3$ unknown coordinates, and M pTOA measurement pairs, is given by the Jacobian [3] of $\mathbf{d}(\mathbf{x})$,

$$\nabla \mathbf{d} = \frac{\partial \mathbf{d}(\mathbf{x})}{\partial \mathbf{x}_u}. \quad (11)$$

The Fisher information matrix \mathbf{J} quantifies the amount of information a measurement data-set contains about the unknown parameters that index the joint PDF of the data-set [6]. The original data-set obviously offers maximum information. If the data is preprocessed in some way, we can measure the ‘‘information-loss’’ due to the preprocessing operation in terms of the Fisher information. If the Fisher information about a subset of parameters is unchanged after preprocessing, we, following [6], denote this preprocessor an *invariant preprocessor*.

3. CLOCK-OFFSET CANCELLATION METHODS

In this section, we develop two invariant preprocessors that remove unknown clock-offsets from (2).

In canceling clock-offsets $\mathbf{H}_t \Delta$ from (2), we wish to find a matrix \mathbf{H}_t^\perp that is orthogonal to \mathbf{H}_t , that is, $\mathbf{H}_t^\perp \mathbf{H}_t = \mathbf{0}$. Many such matrices exist, but, in order to ensure invariant preprocessing, we need to find \mathbf{H}_t^\perp such that the Fisher information about \mathbf{x} is the same in $\boldsymbol{\tau}(\Delta, \mathbf{x})$ as in $\boldsymbol{\tau}_x(\mathbf{x}) = \mathbf{H}_t^\perp \boldsymbol{\tau}$.

3.1. QR-cancellation

We can obtain \mathbf{H}_t^\perp from a QR-factorization of the sparse matrix \mathbf{H}_t , $\mathbf{H}_t = \mathbf{Q}\mathbf{R}$, such that $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$. For the case of an N node network, where one clock-offset is defined to be the global clock reference and the other $N - 1$ clocks are unknown, the rank of \mathbf{H}_t is $N - 1$, assuming $2M$ pTOA measurements are available and that $2M > N - 1$. Matrices \mathbf{Q} and \mathbf{R} can therefore be divided into submatrices such that $\mathbf{H}_t = [\mathbf{Q}_1 \ \mathbf{Q}_2][\mathbf{R}_1^T \ \mathbf{0}]^T = \mathbf{Q}_1 \mathbf{R}_1$, where $\mathbf{Q}_1 \in \mathbb{R}^{2M \times (N-1)}$, $\mathbf{Q}_2 \in \mathbb{R}^{2M \times (2M-N+1)}$, $\mathbf{R}_1 \in \mathbb{R}^{(N-1) \times (N-1)}$. From this we conclude that

$$\mathbf{Q}_2^T \mathbf{H}_t = \mathbf{Q}_2^T [\mathbf{Q}_1 \mathbf{R}_1 + \mathbf{Q}_2 \mathbf{0}] = \mathbf{Q}_2^T \mathbf{Q}_1 \mathbf{R}_1 = \mathbf{0}, \quad (12)$$

since $\mathbf{Q}_2^T \mathbf{Q}_1 = \mathbf{0}$, that is, a possible choice of \mathbf{H}_t^\perp is $\mathbf{H}_t^\perp = \mathbf{Q}_2^T$. Multiplying (2) from the left by \mathbf{Q}_2^T , we obtain

$$\boldsymbol{\tau}_{\text{QR}} = \mathbf{Q}_2^T \boldsymbol{\tau} = \mathbf{Q}_2^T \mathbf{H}_d \mathbf{d}(\mathbf{x}) + \mathbf{Q}_2^T \mathbf{v}.$$

This preprocessed measurement vector $\boldsymbol{\tau}_{\text{QR}} \in \mathbb{R}^{2M-(N-1)}$ is Gaussian with mean $\boldsymbol{\mu}_{\text{QR}}(\mathbf{x}) = \mathbf{Q}_2^T \mathbf{H}_d \mathbf{d}(\mathbf{x})$ and covariance matrix $\mathbf{V}_{\text{QR}} = \mathbf{Q}_2^T \mathbf{V} \mathbf{Q}_2$.

For an N node network, the autonomous relative coordinate estimation problem is now a problem of estimating a maximum of $2N - 3$ unknown parameters given a data-set with a maximum size of $(N - 1)^2$, that is compared to the original problem stated in Section 2, we have reduced the number of parameters by one third and decreased the original data-set with a maximum of $N(N - 1)$ elements for $M = M_{\text{max}}$ by $N - 1$ elements.

Again, using (9), and the Gaussian properties of $\boldsymbol{\tau}_{\text{QR}}$, we can derive the CRB of the preprocessed problem as

$$\text{Var}(\hat{\mathbf{x}}_{\text{QR}}) \geq \left[[\mathbf{H}_d \nabla \mathbf{d}]^T \mathbf{Q}_2 [\mathbf{Q}_2^T \mathbf{V} \mathbf{Q}_2]^{-1} \mathbf{Q}_2^T \mathbf{H}_d \nabla \mathbf{d} \right]^{-1}, \quad (13)$$

where we find that the bound in (13) is equal to the lower right block of the bound in (10) for all parameter vectors \mathbf{x} and Δ and all positive definite noise covariance matrices \mathbf{V} , proof is given in Appendix A. The full Fisher information about \mathbf{x} in $\boldsymbol{\tau}$ is therefore preserved in $\boldsymbol{\tau}_{\text{QR}}$ and so this cancellation method represents an invariant preprocessing method.

It should be noted that, in general, the elements of the preprocessed measurement vector will be correlated, making a distributed positioning algorithm more difficult to implement. As such, the QR-method is more suited for centralized solutions to the autonomous positioning problem.

3.2. Σ -cancellation

To make a distributed positioning scheme feasible after preprocessing, we wish to find an invariant preprocessor \mathbf{H}_t^\perp such that the effect of clock-offsets is eliminated from $\boldsymbol{\tau}$, while the transformed problem can be distributed evenly among the nodes in the network, reducing the need for long distance, multiple-hop communication.

If we assume that the pTOA measurement noise variance only depends on the range between nodes and on system parameters such as bandwidth (see, e.g., [7] for justification of this assumption), we can assume that the pTOA measurement noise variance on the forward and reverse channels between two nodes are equal. With this key assumption, assuming a pair-wise ordered data-set $\boldsymbol{\tau} \in \mathbb{R}^{2M}$, $\mathbf{V} = \text{diag}(\sigma_1^2, \sigma_1^2, \sigma_2^2, \sigma_2^2, \dots, \sigma_M^2, \sigma_M^2)$, where $\sigma_k^2 = \sigma_{i(k),j(k)}^2$.

Then, upon inspection of the joint PDF of $\boldsymbol{\tau}$,

$$p(\boldsymbol{\tau}; \mathbf{x}, \Delta) = \frac{1}{\sqrt{(2\pi)^{2M} |\mathbf{V}|}} \times \exp \left[-\frac{1}{2} (\boldsymbol{\tau} - \boldsymbol{\mu}_\tau)^T \mathbf{V}^{-1} (\boldsymbol{\tau} - \boldsymbol{\mu}_\tau) \right], \quad (14)$$

we find that, under the assumption of pair-wise equal variances, we can, as derived in Appendix B, factor the PDF as

$$p(\boldsymbol{\tau}; \mathbf{x}, \Delta) = \frac{1}{\sqrt{(2\pi)^{2M} |\mathbf{V}|}} \exp \left[-\frac{1}{2}(A - B) \right] \times \exp \left[-\frac{1}{2}(C - D + E) \right], \quad (15)$$

where

$$\begin{aligned} A &= [\mathbf{H}_d \mathbf{d}(\mathbf{x})]^T \mathbf{V}^{-1} \mathbf{H}_d \mathbf{d}(\mathbf{x}), \\ B &= 2\mathbf{d}^T(\mathbf{x}) \mathbf{V}_2^{-1} \mathbf{H}_d^T \boldsymbol{\tau}, \\ C &= \boldsymbol{\tau}^T \mathbf{V}^{-1} \boldsymbol{\tau}, \end{aligned} \quad (16)$$

$$D = 2[\mathbf{T} \mathbf{H}_t \Delta]^T \mathbf{V}_2^{-1} \mathbf{D} \boldsymbol{\tau},$$

$$E = [\mathbf{H}_t \Delta]^T \mathbf{V}^{-1} \mathbf{H}_t \Delta,$$

$$\mathbf{V}_2 = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_M^2) \in \mathbb{R}^{M \times M}, \quad (17)$$

$$\mathbf{D} = \mathbf{I}_M \otimes \begin{bmatrix} 1 & -1 \end{bmatrix} \in \mathbb{R}^{M \times 2M}, \quad (18)$$

$$\mathbf{T} = \mathbf{I}_M \otimes \begin{bmatrix} 1 & 0 \end{bmatrix} \in \mathbb{R}^{M \times 2M}. \quad (19)$$

Also, from (5) and (7) and the properties of the Kronecker product, $[\mathbf{A} \otimes \mathbf{B}]^T = \mathbf{A}^T \otimes \mathbf{B}^T$, $[\mathbf{A} \otimes \mathbf{B}][\mathbf{C} \otimes \mathbf{D}] = \mathbf{AC} \otimes \mathbf{BD}$, where it is assumed that all matrix products exist, we have

$$\begin{aligned} \mathbf{H}_d^T \mathbf{H}_t &= \left[\frac{1}{c} \mathbf{I}_M \otimes \begin{bmatrix} 1 & 1 \end{bmatrix} \right]^T \left[\mathbf{H}_0 \otimes \begin{bmatrix} 1 & -1 \end{bmatrix} \right]^T \\ &= \frac{1}{c} \mathbf{I}_M \mathbf{H}_0 \otimes \left(\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right) = \mathbf{0}, \end{aligned} \quad (20)$$

that is, \mathbf{H}_d^T is orthogonal to matrix \mathbf{H}_t . The PDF is now on the form $p(\boldsymbol{\tau}; \mathbf{x}, \Delta) = f(S(\boldsymbol{\tau}; \mathbf{x})h(\boldsymbol{\tau}; \Delta))$, where $S(\boldsymbol{\tau}) = \mathbf{H}_d^T \boldsymbol{\tau}$, that is, $\boldsymbol{\tau}_\Sigma(\mathbf{x}) = c\mathbf{H}_d^T \boldsymbol{\tau}(\mathbf{x}, \Delta)$, the sum of forward and reverse pTOAs is, under the above mentioned assumption of pair-wise equal noise variances, a partially sufficient statistic [3, 8] for the estimation of relative node coordinates \mathbf{x} . That it is also complete meaning there is only one function of $\boldsymbol{\tau}_\Sigma(\mathbf{x})$ that is an unbiased estimator of $\mathbf{d}(\mathbf{x})$, follows from the fact that the PDF in (14) is a member of the exponential family of PDFs [3]. It follows from the partial sufficiency of $\boldsymbol{\tau}_\Sigma$, that the full Fisher information about \mathbf{x} in $\boldsymbol{\tau}$ is preserved in $\boldsymbol{\tau}_\Sigma$ [3, 8], and the preprocessor \mathbf{H}_d^T is therefore invariant. Further, since $\mathbf{H}_d^T \mathbf{H}_d = 2\mathbf{I}/c^2$, the mean of the Gaussian vector $\boldsymbol{\tau}_\Sigma$ is $\mathbb{E}[\boldsymbol{\tau}_\Sigma] = 2\mathbf{d}(\mathbf{x})/c$, that is, one half of a measured round-trip time, multiplied by c , corresponds to the internode distance. We can now formulate the ML estimator of relative node coordinates, operating on $\boldsymbol{\tau}_\Sigma$ as

$$\hat{\mathbf{x}}_\Sigma = \arg \min_{\mathbf{x}} (c\boldsymbol{\tau}_\Sigma - 2\mathbf{d}(\mathbf{x}))^T [\mathbf{H}_d \mathbf{V} \mathbf{H}_d^T]^{-1} (c\boldsymbol{\tau}_\Sigma - 2\mathbf{d}(\mathbf{x})). \quad (21)$$

This problem is equivalent to minimizing the energy in a system of point-masses and springs, where the springs obey Hooke's law. It is shown in [4] that a distributed algorithm based on this analogy can indeed be considered statistically efficient under a range of reasonable assumptions.

4. A DISTRIBUTED POSITIONING ALGORITHM

In this section, a distributed algorithm is presented, that divides the global asynchronous relative positioning problem into a set of separate subproblems distributed across the network. CRB theory is then relied upon to fuse the solutions to the subproblems, increasing accuracy step-wise up to the desired performance, while keeping computational complexity low.

4.1. The kernel algorithm

The kernel algorithm is an extension of the classic TDOA positioning technique, widely employed and well known throughout the positioning community. In a classic TDOA positioning algorithm, pTOA measurements are made by three fixed and synchronized reference stations, with respect to the mobile node. The estimated position of the mobile node is then obtained as the intersection of two hyperbolic curves, resulting from a difference operation on the three measured pTOAs [9, 10]. The kernel algorithm extends this concept to the case where there are no fixed synchronized reference stations, and more than three pTOA measurements are available.

Basically, the kernel algorithm operates in three phases;

- (i) Partition the network into groups of at least three nodes (kernels). For each kernel, define a local coordinate system.
- (ii) Using standard time-difference-of-arrival (TDOA) techniques, estimate the coordinates of all other nodes in transmission range of the kernel.
- (iii) For each positioned node outside the kernel, estimate the accuracy in relative coordinates. If the accuracy is found to be inadequate for the application at hand, use the accuracy estimate in a fusion process with other kernels in order to improve on position estimates.

Forming a kernel

To form a kernel, we first need to partition the network into groups of three nodes. Due to the varying geometric properties of different network partitions [2], the choice of partition will influence the accuracy of the position estimates. We are, however, not assuming any prior knowledge of node locations and therefore partition the network randomly, that is, without any attempts at optimization. It should be noted that an initial random partition of the network does not have to be complete in the sense that every node will be a member of exactly one kernel, for the kernel algorithm to produce valid coordinate estimates. Some nodes may be members of zero, two or more kernels in an initial run of the algorithm, the extension of our algorithm to this case being trivial. We assume hard-wired global node identification is available, and denote the coordinates of the i th node in the local coordinate system of kernel k as $\mathbf{x}_{k,i} = [x_{k,i} \ y_{k,i}]^T$. The indices of the three nodes in the k th kernel are denoted k_1, k_2 , and k_3 . Assuming pTOA measurements have been exchanged by the three nodes in kernel k , we first assign the center coordinates

and a zero clock-offset to node k_1 , that is, $\hat{\mathbf{x}}_{k,k_1} = [0 \ 0]^T$, and $\Delta_{k,k_1} = 0$. The estimated internode distances \hat{d}_{k_1,k_2} , \hat{d}_{k_1,k_3} , and \hat{d}_{k_2,k_3} are obtained from the sum of two corresponding pTOA distance measurements, $\hat{d}_{k_i,k_j} = c(\tau_{k_i,k_j} + \tau_{k_j,k_i})/2$, where $\tau_{i,j}$ is given by (1), eliminating the unknown clock-offsets. It should be noted that as long as the measurement error characteristics are similar on the forward and reverse link between two nodes, this fusion of pTOA measurements represents a sufficient statistic and therefore does not represent any information loss, as derived in the previous section. We assign the coordinates $\hat{\mathbf{x}}_{k,k_2} = [0 \ \hat{d}_{k_1,k_2}]^T$ to the second node within our kernel, fixing it on the y -axis of the local coordinate system. Finally, we, using standard trigonometric identities, estimate the remaining unknown kernel coordinates $\hat{\mathbf{x}}_{k,k_3} = [\hat{x}_{k,k_3} \ \hat{y}_{k,k_3}]^T$ as

$$\hat{y}_{k,k_3} = \frac{\hat{d}_{k_1,k_3}^2 + \hat{d}_{k_1,k_2}^2 - \hat{d}_{k_2,k_3}^2}{2\hat{d}_{k_1,k_2}}, \quad (22)$$

$$\hat{x}_{k,k_3} = \begin{cases} \pm \sqrt{\hat{d}_{k_1,k_3}^2 - \hat{y}_{k,k_3}^2}, & \text{if } \hat{d}_{k_1,k_3}^2 - \hat{y}_{k,k_3}^2 > 0 \\ 0, & \text{otherwise.} \end{cases}$$

If $\hat{d}_{k_1,k_3}^2 - \hat{y}_{k,k_3}^2 < 0$, it is assumed that the third node is located very close to the y -axis. When this happens, the nodes in kernel k are almost colinear, resulting in poor locationing performance, due to the high geometric dilution-of-precision (GDOP) [2]. However, this poor performance is easily detectable. We note a mirror ambiguity when forming a kernel. This ambiguity may be resolved if at least two fixed nodes, or other prior information, are available within the system, but, since we are only interested in the relative location of nodes, the algorithm is able to resolve this ambiguity in the fusion process described in Section 4.3.

We also estimate the error covariance matrix $\mathbf{C}_{k,k_i} = \mathbb{E}[(\hat{\mathbf{x}}_{k,k_i} - \mathbf{x}_{k,k_i})(\hat{\mathbf{x}}_{k,k_i} - \mathbf{x}_{k,k_i})^T]$ of the i th node in the k th kernel in units of m^2 . Since node k_1 is taken as reference for kernel k , the covariance matrix $\mathbf{C}_{k,k_1} = \mathbf{0}$. The covariance matrix \mathbf{C}_{k,k_2} of node k_2 will, assuming pair-wise equal pTOA noise variances, have a variance in the y -direction corresponding to half of the pTOA measurement variance σ_{k_1,k_2}^2 , translated into distance, that is, $\mathbf{C}_{k,k_2} = c^2 \text{diag}(0, \sigma_{k_1,k_2}^2/2)$. We finally estimate the covariance matrix $\hat{\mathbf{C}}_{k,k_3}$, of kernel member k_3 , as the CRB on node coordinate estimates $[\hat{x}_{k,k_3} \ \hat{y}_{k,k_3}]^T$, computed in estimated coordinates. The estimate is given as the projection of internode distance variances on the reference system formed by node k_1 and k_2 of kernel k [11],

$$\hat{\mathbf{C}}_{k,k_3} = c^2 \left[\frac{\hat{\mathbf{u}}_{k,k_1,k_3} \hat{\mathbf{u}}_{k,k_1,k_3}^T}{\sigma_{k_1,k_3}^2/2} + \frac{\hat{\mathbf{u}}_{k,k_2,k_3} \hat{\mathbf{u}}_{k,k_2,k_3}^T}{\sigma_{k_2,k_3}^2} \right]^{-1}, \quad (23)$$

where

$$\hat{\mathbf{u}}_{k,i,j} = \frac{\hat{\mathbf{x}}_{k,j} - \hat{\mathbf{x}}_{k,i}}{\|\hat{\mathbf{x}}_{k,j} - \hat{\mathbf{x}}_{k,i}\|} \quad (24)$$

is the estimate of a unit vector in the direction of node j from node i in the coordinate system of kernel k . The contribution

from node k_2 in (23) has a greater distance variance due to the uncertainty in location of this node. An extension of covariance estimators to nonequal pTOA variances is trivial.

Finally, kernel nodes k_2 and k_3 tune their local clocks to the clock of node k_1 , using computed internode distances and measured pTOAs.

4.2. Obtaining relative locations using information available within a kernel

Now that we have, in a relative sense, fixed our kernel, achieved approximate synchronization within the kernel and estimated the accuracy in kernel positions, we move on to position the remaining nodes of the network. To locate some node l , not a member of kernel k , we use pTOA measurements τ_{l,k_1} , τ_{l,k_2} , and τ_{l,k_3} available within the kernel. As noted above, the pTOA measurements are affected by Gaussian noise with variance σ_{l,k_i}^2 . Taking worst-case uncertainties in kernel locations and clock-offsets into account, we estimate the covariance matrix of the stacked pTOA measurements $\mathbf{p}_{l,k} = [\tau_{l,k_1} \ \tau_{l,k_2} \ \tau_{l,k_3}]^T$ as [11],

$$\hat{\mathbf{Q}}_{l,k} = \text{diag} \left(\sigma_{l,k_1}^2, \sigma_{l,k_2}^2 + 2 \frac{\text{tr}(\hat{\mathbf{C}}_{k,k_2})}{c^2}, \sigma_{l,k_3}^2 + 2 \frac{\text{tr}(\hat{\mathbf{C}}_{k,k_3})}{c^2} \right). \quad (25)$$

The three-element vector of pTOA measurements can be combined into two TDOA measurements,

$$\mathbf{t}_{l,k} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \mathbf{p}_{l,k} = \mathbf{H} \mathbf{p}_{l,k} = \begin{bmatrix} \tau_{l,k_1} - \tau_{l,k_2} \\ \tau_{l,k_1} - \tau_{l,k_3} \end{bmatrix}, \quad (26)$$

canceling the unknown clock-offset Δ_l of node l , with respect to the clock of kernel node k_1 . The estimated covariance matrix of transformed measurements is $\hat{\mathbf{R}}_{l,k} = \mathbf{H} \hat{\mathbf{Q}}_{l,k} \mathbf{H}^T$.

Now, the kernel k estimator of the two-dimensional coordinates $\hat{\mathbf{x}}_{k,l}$ of node l in the network is given by

$$\hat{\mathbf{x}}_{k,l} = \arg \min_{\mathbf{x}_{k,l}} \left(\mathbf{t}_{l,k} - \frac{\mathbf{f}(\mathbf{x}_{k,l})}{c} \right)^T \hat{\mathbf{R}}_{l,k}^{-1} \left(\mathbf{t}_{l,k} - \frac{\mathbf{f}(\mathbf{x}_{k,l})}{c} \right), \quad (27)$$

where

$$\mathbf{f}(\mathbf{x}_{k,l}) = \begin{bmatrix} \|\mathbf{x}_{k,l} - \hat{\mathbf{x}}_{k,k_1}\| - \|\mathbf{x}_{k,l} - \hat{\mathbf{x}}_{k,k_2}\| \\ \|\mathbf{x}_{k,l} - \hat{\mathbf{x}}_{k,k_1}\| - \|\mathbf{x}_{k,l} - \hat{\mathbf{x}}_{k,k_3}\| \end{bmatrix}. \quad (28)$$

A minimizer of (27) is a solution, should at least one exist, to $\mathbf{f}(\hat{\mathbf{x}}_{k,l})/c = \mathbf{t}_{l,k}$, derived in [9]. Zero, one, or two solutions may exist, corresponding to zero, one, or two intersections of the TDOA hyperbolas. If two solutions exist, both locations are remembered and one is later discarded based on information from other kernels. On some occasions, there is no closed form solution to $\mathbf{f}(\hat{\mathbf{x}}_{k,l})/c = \mathbf{t}_{l,k}$; then the algorithm presented in [9] does not produce a minimizer in (27), and the node is considered unfixed. Unfixed nodes are assigned high variance estimates, excluding them from future steps of the algorithm. A possibility not considered in this work is a numerical minimization of (27) when no closed-form solution to $\mathbf{f}(\hat{\mathbf{x}}_{k,l})/c = \mathbf{t}_{l,k}$ exists. For the case of a kernel

not having access to a complete set of three pTOA measurements with respect to some node, due to, for instance, signal strength issues, the node is also considered unfixed.

Under the assumption that $\mathbf{t}_{l,k}$ is Gaussian with mean $\mathbb{E}[\mathbf{H}\mathbf{p}_{l,k}]$, and covariance matrix $\mathbf{R}_{l,k}$, the CRB for an unbiased estimator of $\mathbf{x}_{k,l}$ is given by the inverse of the Fisher information matrix in (9). We estimate the covariance matrix $\hat{\mathbf{C}}_{k,l}$ of $\hat{\mathbf{x}}_{k,l}$ as the CRB evaluated at the estimated coordinates $\hat{\mathbf{x}}_{k,l}$, using the estimated measurement covariance matrix $\hat{\mathbf{R}}_{l,k}$, that is,

$$\hat{\mathbf{C}}_{k,l} = \left[\nabla_{\mathbf{p}_{l,k}}^T \mathbf{H}^T \hat{\mathbf{R}}_{l,k}^{-1} \mathbf{H} \nabla_{\mathbf{p}_{l,k}} \right]^{-1}, \quad (29)$$

where $\nabla_{\mathbf{p}_{l,k}} = [\hat{\mathbf{u}}_{k,k_1,l} \ \hat{\mathbf{u}}_{k,k_2,l} \ \hat{\mathbf{u}}_{k,k_3,l}]^T/c$, and $\hat{\mathbf{u}}_{k,i,l}$ is given by (24). This approach to variance estimation, that is, using the CRB calculated in estimated coordinates to estimate the accuracy in a position estimate based on TDOA measurements, as a rule of thumb, yields accurate estimates as long as the true variance is reasonably small [3, 10]. If the true variance is large, the error of the estimated variance will be large, but so will the estimated variance, making this approach suitable for fusion purposes.

4.3. Fusion of kernel estimates

In order for one kernel r to share its positioning information with another kernel k , the estimate has to be transformed so as to fit into the local coordinate system of kernel k . Since both kernels will have different nodes located in the origin and also different nodes fixed on the y -axis, a bias as well as a rotation will separate the two estimates. As noted above, a mirror ambiguity may also separate the two estimates. To find this bias, rotation angle and possible ambiguity, nodes that have a low location variance should be given more weight than nodes that are poorly located or not fixed at all. We wish to find the rotation matrix

$$\mathbf{G}(\alpha_{r-k}) = \begin{bmatrix} \cos \alpha_{r-k} & \sin \alpha_{r-k} \\ -\sin \alpha_{r-k} & \cos \alpha_{r-k} \end{bmatrix} \quad (30)$$

and the bias \mathbf{b}_{r-k} that, based on an ML estimator of rotation angle and bias, derived under the assumption of zero-mean Gaussian single kernel positioning errors, minimizes the weighted sum of squared Euclidean distances,

$$\begin{bmatrix} \hat{\alpha}_{r-k} \\ \hat{\mathbf{b}}_{r-k} \end{bmatrix} = \arg \min_{\alpha, \mathbf{b}} \sum_{i=1}^N (\hat{w}_{k,i}^2 + \hat{w}_{r,i}^2)^{-1} \left\| \hat{\mathbf{x}}'_{r,i} - \hat{\mathbf{x}}_{k,i} \right\|^2, \quad (31)$$

where $\hat{\mathbf{x}}'_{r,i} = \mathbf{G}(\alpha)\hat{\mathbf{x}}_{r,i} + \mathbf{b}$ is the rotated and translated kernel r estimate of node i , and $\hat{w}_{k,i}^2 = \text{tr}(\hat{\mathbf{C}}_{k,i})$. If mirror ambiguities have not been resolved beforehand, they can be resolved by trying both possible orientations of kernel r in (31), and selecting the orientation with maximum likelihood, that is, the orientation with the best weighted MS fit. It can be shown

[11] that the angle α_{r-k} that minimizes (31) is given by

$$\hat{\alpha} = \arctan \left(\frac{\sum_{i=1}^N (\hat{w}_{k,i}^2 + \hat{w}_{r,i}^2)^{-1} (\hat{\tilde{x}}_{k,i} \hat{\tilde{y}}_{r,i} - \hat{\tilde{y}}_{k,i} \hat{\tilde{x}}_{r,i})}{\sum_{l=1}^N (\hat{w}_{k,l}^2 + \hat{w}_{r,l}^2)^{-1} (\hat{\tilde{x}}_{k,l} \hat{\tilde{y}}_{r,l} + \hat{\tilde{y}}_{k,l} \hat{\tilde{x}}_{r,l})} \right), \quad (32)$$

where the function $\arctan(a, b)$ is the four-quadrant inverse tangent function, $\hat{\tilde{x}}_{k,i}$, and $\hat{\tilde{y}}_{k,i}$ are the x -coordinate and y -coordinate of node i in kernel estimate k , centered with respect to the weighted center of gravity at kernel k , $[\hat{\tilde{x}}_{k,i} \ \hat{\tilde{y}}_{k,i}]^T = [\hat{x}_{k,i} \ \hat{y}_{k,i}]^T - \bar{\mathbf{x}}_k$, and

$$\bar{\mathbf{x}}_k = \left[\sum_{i=1}^N (\hat{w}_{k,i}^2 + \hat{w}_{r,i}^2)^{-1} \right]^{-1} \sum_{l=1}^N (\hat{w}_{k,l}^2 + \hat{w}_{r,l}^2)^{-1} \hat{\mathbf{x}}_{k,l}. \quad (33)$$

Likewise, the estimated weighted bias $\hat{\mathbf{b}}_{r-k}$, separating kernel k and r estimates, and minimizing (31), is given by [11],

$$\hat{\mathbf{b}}_{r-k} = - \left[\sum_{i=1}^N (\hat{w}_{k,i}^2 + \hat{w}_{r,i}^2)^{-1} \right]^{-1} \sum_{l=1}^N (\hat{w}_{k,l}^2 + \hat{w}_{r,l}^2)^{-1} \hat{\mathbf{e}}_l, \quad (34)$$

where $\hat{\mathbf{e}}_l = \mathbf{G}(\hat{\alpha}_{r-k})\hat{\mathbf{x}}_{r,l} - \hat{\mathbf{x}}_{k,l}$ is the error of each node after rotation.

Once the rotation angle and bias minimizing (31) has been found, we also have to apply the rotation matrix $\mathbf{G}(\hat{\alpha}_{r-k})$ to the covariance matrix of the kernel estimate subject to rotation,

$$\hat{\mathbf{C}}'_{r,i} = \mathbf{G}(\hat{\alpha}_{r-k})\hat{\mathbf{C}}_{r,i}\mathbf{G}^T(\hat{\alpha}_{r-k}), \quad \forall i \in [1, N]. \quad (35)$$

When kernel estimates have been rotated into a common frame, the merged estimate is obtained as a straightforward fusion of Gaussian variables [3],

$$\hat{\mathbf{x}} = \left[\hat{\mathbf{C}}_r'^{-1} + \hat{\mathbf{C}}_k^{-1} \right]^{-1} \left[\hat{\mathbf{C}}_r'^{-1} \hat{\mathbf{x}}'_r + \hat{\mathbf{C}}_k^{-1} \hat{\mathbf{x}}_k \right], \quad (36)$$

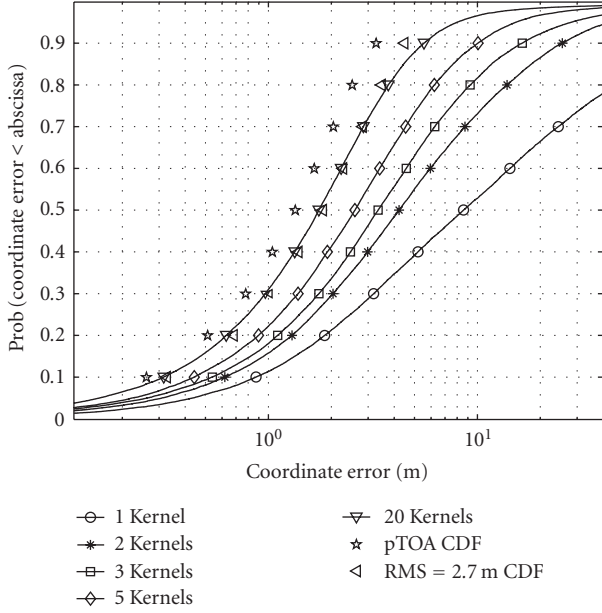
where covariance estimators $\hat{\mathbf{C}}'_r = \text{diag}(\hat{\mathbf{C}}'_{r,1}, \dots, \hat{\mathbf{C}}'_{r,N})$ and $\hat{\mathbf{C}}_k = \text{diag}(\hat{\mathbf{C}}_{k,1}, \dots, \hat{\mathbf{C}}_{k,N})$, and coordinate estimators $\hat{\mathbf{x}}_k = [\hat{\mathbf{x}}_{k,1}^T \ \dots \ \hat{\mathbf{x}}_{k,N}^T]^T$ and $\hat{\mathbf{x}}'_r = [\hat{\mathbf{x}}'_{r,1}^T \ \dots \ \hat{\mathbf{x}}'_{r,N}^T]^T$. The fused estimate will, from the covariance matrix of a weighted mean of two Gaussian vectors, weighted by the inverse of their respective covariance matrices, have covariance matrix estimate

$$\hat{\mathbf{C}}_{r+k} = \left[\hat{\mathbf{C}}_r'^{-1} + \hat{\mathbf{C}}_k^{-1} \right]^{-1}. \quad (37)$$

If the estimates of more than two kernels are to be fused, the process is repeated for each additional kernel using the previous merging as base of rotation. The reason for merging kernels in a successive manner is an increase in rotation accuracy.

4.4. Simulation results

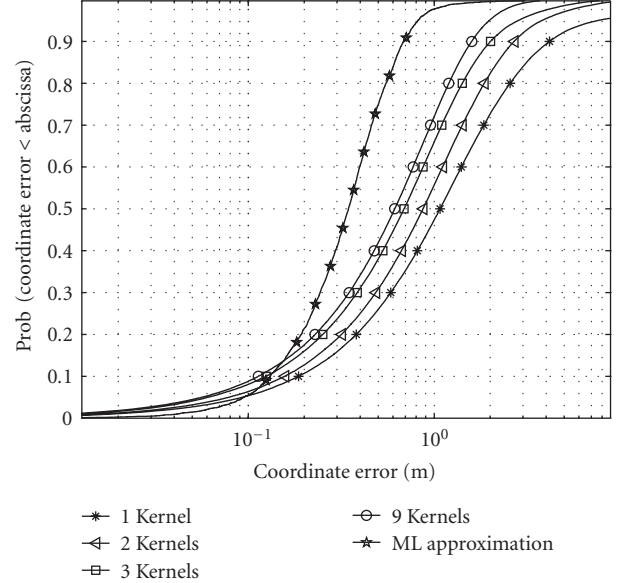
To evaluate the performance of our proposed algorithm, random node coordinates in networks of varying sizes were generated from a uniform distribution, constrained within a

FIGURE 1: 60 nodes random network layout, $c\sigma = 2$ m.

square with side 500 m. For each node, a clock-offset was generated from a zero-mean Gaussian distribution with a variance of 1 s^2 . Based on the node coordinates and the clock-offsets, true pTOA distance measurements were calculated and zero-mean Gaussian noise with a standard deviation of $c\sigma = 2$ m was added, that is, the noise variance was assumed equal for all pTOAs. The nodes in the network were randomly grouped into kernels, each containing three nodes, and the algorithm was run to produce estimates, fusing a varying number of kernel estimates. The node location errors were saved and the process was repeated 100 times. In Figure 1, the cumulative distribution function (CDF) of the location error is plotted for different numbers of merged kernels. The cumulative effect in accuracy is clear from Figure 1, adding information from more kernels produces estimates of higher accuracy. If none of the merged kernels have a solution for some node, this node remains unfixed with infinite variance. Obviously, the number of unfixed nodes decreases drastically with the number of merged kernels. For comparison purposes, the CDF of the pTOA measurement noise, used in the simulations, has been included. The CDF of a Gaussian positioning error with a root-mean-square (RMS) value of 2.7 m is also included. In [2], RMS locationing accuracies between 0.9 and 2.7 m are reported for a TOA measurement standard deviation of around 1.83 m. The comparison to [2] being somewhat unfair since a smaller network, including fixed reference nodes and oriented in a square grid pattern, was implemented in [2].

If the simulation results are investigated in more detail, we find that nodes located on the outskirts of the network are often located with less accuracy than nodes situated near the center. The same phenomenon is noted and explained in [2].

If the measurement noise variance σ^2 is reduced, we experience a substantial performance gain. The main reason,

FIGURE 2: 27 nodes square network layout, $c\sigma = 1$ m.

of course, is more accurate kernel estimates, but also a more accurate covariance matrix estimate $\hat{\mathbf{C}}_k$, yielding a more efficient fusion process. In Figure 2, the effect of using a square node deployment pattern is exemplified, 27 nodes were placed on a grid pattern, kernel assignments were random, and the measurement accuracy was set to $c\sigma = 1$ m. Compared to Figure 1, we note an improvement, especially for a smaller number of fused kernels. This is mainly due to the lower average GDOP, experienced by single kernel estimators. For comparison purposes, we also plot the performance of an approximation to the ML estimator of relative node coordinates, given by (21), discussed in [1] and also in [4].

The robustness of the algorithm was verified in each simulation run. We investigate the relationship γ between instantaneous squared error and estimated MSE,

$$\gamma = \frac{1}{2N-3} \mathbf{e}^T \hat{\mathbf{C}}_{\sum_{k=1}^K}^{-1} \mathbf{e}, \quad (38)$$

where \mathbf{e} is a column vector of the stacked node location errors and $\hat{\mathbf{C}}_{\sum_{k=1}^K}$ is the covariance matrix estimate when K kernels have been fused. Simulation runs producing values of γ below one indicate a pessimistic estimate of the node coordinate errors while values greater than one indicate an optimistic estimate. In Figure 3, the CDF of γ is plotted for simulation setups, all with 27 nodes and $K = 9$, distributed uniformly within a square area with side 500 m, but with different pTOA measurement variances. Simulations were also made for a scenario with 27 nodes located in a square grid pattern, and a pTOA standard deviation of $c\sigma = 1$ m. Each simulation was run for 1000 network layout and measurement noise realizations. The obtained results indicate a robust algorithm. From simulation results, we note, that if the measurement noise variance is low, or the network has a low GDOP layout, yielding more accurate coordinate estimates,

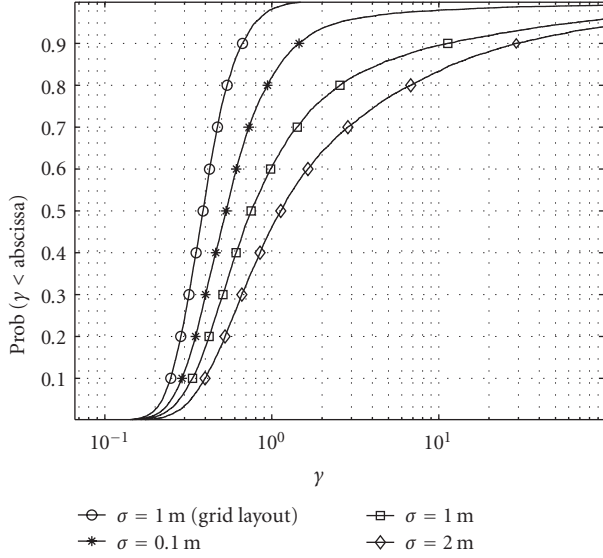


FIGURE 3: Indication of algorithm robustness.

the kernel algorithm produces somewhat pessimistic accuracy predictions. This is most likely due to the worst-case assumption in (25). Extreme values of γ , especially common for higher noise variances, are caused by poorly located or unfixed nodes, where both estimated and true errors are large. From a data fusion point of view, erroneous error estimates have little effect, as long as both true and estimated error is large, in which case the estimates are heavily down-weighted in the fusion process.

5. CONCLUSION

This paper, based on analysis of the Cramér-Rao lower bound (CRB) of the asynchronous and autonomous relative coordinate estimation problem, has derived two methods for canceling unknown clock-offsets at individual nodes from the coordinate estimation problem. Both methods were shown to represent invariant preprocessors, that is, neither method altered the CRB of the original estimation problem, and the methods fit well together with centralized or distributed ML-type coordinate estimators, described in [1, 2, 4]. It was also argued that CRB-type expressions may be used in estimating the performance of a positioning algorithm. This concept was exploited in a distributed, suboptimal algorithm, that had the ability to increase performance step-wise, according to requirements from the served application.

APPENDICES

A. INFORMATION PRESERVATION OF THE QR METHOD

The proposition that the right-hand side of (13) is equal to the lower right block of the right-hand side of (10), for all parameter vectors \mathbf{x} and $\mathbf{\Delta}$, and all symmetric and positive

definite noise covariance matrices \mathbf{V} , may be stated as

$$\mathbf{TC}_j\mathbf{T}^T = \mathbf{C}_s, \quad (\text{A.1})$$

where $\mathbf{T} = [\mathbf{0}_{N-1} \ \mathbf{I}_C]$,

$$\mathbf{C}_j = \begin{bmatrix} \mathbf{H}_t^T \mathbf{V}^{-1} \mathbf{H}_t & \mathbf{H}_t^T \mathbf{V}^{-1} \mathbf{H}_d \nabla_d \\ [\mathbf{H}_d \nabla_d]^T \mathbf{V}^{-1} \mathbf{H}_t & [\mathbf{H}_d \nabla_d]^T \mathbf{V}^{-1} \mathbf{H}_d \nabla_d \end{bmatrix}^{-1}, \quad (\text{A.2})$$

$$\mathbf{C}_s = \left[[\mathbf{H}_d \nabla_d]^T \mathbf{Q}_2 [\mathbf{Q}_2^T \mathbf{V} \mathbf{Q}_2]^{-1} \mathbf{Q}_2^T \mathbf{H}_d \nabla_d \right]^{-1},$$

\mathbf{H}_t is given by (5), \mathbf{H}_d is given in (7), ∇_d is given by (11), and \mathbf{Q}_2 is defined in Section 3.1. Now, the matrix inversion lemma states that

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1} \mathbf{B} \mathbf{S}_A^{-1} \mathbf{C} \mathbf{A}^{-1} & -\mathbf{A}^{-1} \mathbf{B} \mathbf{S}_A^{-1} \\ -\mathbf{S}_A^{-1} \mathbf{C} \mathbf{A}^{-1} & \mathbf{S}_A^{-1} \end{bmatrix}, \quad (\text{A.3})$$

where $\mathbf{S}_A = (\mathbf{D} - \mathbf{C} \mathbf{A}^{-1} \mathbf{B})$ is the Schur complement of \mathbf{A} . The matrix $\mathbf{TC}_j\mathbf{T}^T$ can therefore be written as

$$\begin{aligned} \mathbf{TC}_j\mathbf{T}^T &= \left[\left([\mathbf{H}_d \nabla_d]^T \mathbf{V}^{-1} \mathbf{H}_d \nabla_d \right) - \left([\mathbf{H}_d \nabla_d]^T \mathbf{V}^{-1} \mathbf{H}_t \right) \right. \\ &\quad \left. \times \left(\mathbf{H}_t^T \mathbf{V}^{-1} \mathbf{H}_t \right)^{-1} \left(\mathbf{H}_t^T \mathbf{V}^{-1} \mathbf{H}_d \nabla_d \right) \right]^{-1}. \end{aligned} \quad (\text{A.4})$$

The equality in (A.1) holds, that is, $\mathbf{TC}_j\mathbf{T}^T = \mathbf{C}_s$ if

$$\mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{H}_t \left(\mathbf{H}_t^T \mathbf{V}^{-1} \mathbf{H}_t \right)^{-1} \mathbf{H}_t^T \mathbf{V}^{-1} = \mathbf{Q}_2 \left[\mathbf{Q}_2^T \mathbf{V} \mathbf{Q}_2 \right]^{-1} \mathbf{Q}_2^T, \quad (\text{A.5})$$

where \mathbf{V} is positive definite and symmetric, which implies that positive definite and symmetric matrices $\mathbf{V}^{1/2}$ and $\mathbf{V}^{-1/2}$ exist such that $\mathbf{V}^{1/2} \mathbf{V}^{1/2} = \mathbf{V}$ and $\mathbf{V}^{-1/2} \mathbf{V}^{-1/2} = \mathbf{V}^{-1}$.

The left-hand side of (A.5) can be written as

$$\begin{aligned} &\mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{H}_t \left(\mathbf{H}_t^T \mathbf{V}^{-1} \mathbf{H}_t \right)^{-1} \mathbf{H}_t^T \mathbf{V}^{-1} \\ &= \mathbf{V}^{-1/2} \left(\mathbf{I} - \mathbf{A} \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T \right) \mathbf{V}^{-1/2} \\ &= \mathbf{V}^{-1/2} \boldsymbol{\pi}_A^{\perp} \mathbf{V}^{-1/2}, \end{aligned} \quad (\text{A.6})$$

where $\mathbf{A} = \mathbf{V}^{-1/2} \mathbf{H}_t$, and $\mathbf{A}^T = [\mathbf{V}^{-1/2} \mathbf{H}_t]^T = \mathbf{H}_t^T \mathbf{V}^{-1/2}$, following from the symmetry of $\mathbf{V}^{-1/2}$. The matrix $\boldsymbol{\pi}_A^{\perp} = \mathbf{I} - \mathbf{A} \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T$ is a projection matrix onto the orthogonal complement subspace of range (\mathbf{A}) [3, page 232]. The right-hand side of (A.5) is

$$\begin{aligned} &\mathbf{Q}_2 \left[\mathbf{Q}_2^T \mathbf{V} \mathbf{Q}_2 \right]^{-1} \mathbf{Q}_2^T \\ &= \mathbf{V}^{-1/2} \mathbf{V}^{(1/2)} \mathbf{Q}_2 \left[\mathbf{Q}_2^T \mathbf{V}^{1/2} \mathbf{V}^{1/2} \mathbf{Q}_2 \right]^{-1} \mathbf{Q}_2^T \mathbf{V}^{1/2} \mathbf{V}^{-1/2} \\ &= \mathbf{V}^{-1/2} \boldsymbol{\pi}_B \mathbf{V}^{-1/2}, \end{aligned} \quad (\text{A.7})$$

where $\boldsymbol{\pi}_B$ is a projection matrix onto the space spanned by the columns of $\mathbf{B} = \mathbf{V}^{1/2} \mathbf{Q}_2$.

Now, since the projection matrix onto a subspace is unique, it suffices to show that $(\text{range}(\mathbf{A}))^\perp = \text{range}(\mathbf{B})$. It is well known that the left null space of a matrix is the orthogonal complement of the column space (or the range), that is,

$$(\text{range}(\mathbf{A}))^\perp = \text{null}(\mathbf{A}^T) = \{\mathbf{x} : \mathbf{H}_t^T \mathbf{V}^{-1/2} \mathbf{x} = \mathbf{0}\}. \quad (\text{A.8})$$

Also, since $\mathbf{A} \in \mathbb{R}^{2M \times (N-1)}$ has full column rank,

$$\dim(\text{range}(\mathbf{A}))^\perp = 2M - \text{rank}(\mathbf{A}) = 2M - N + 1. \quad (\text{A.9})$$

Further, we have $\text{range}(\mathbf{B}) = \text{range}(\mathbf{V}^{1/2} \mathbf{Q}_2) = \{\mathbf{y} : \mathbf{y} = \mathbf{V}^{1/2} \mathbf{Q}_2 \mathbf{z}, \mathbf{z} \in \mathbb{R}^{2M-N+1}\}$, and

$$\dim \text{range}(\mathbf{B}) = \text{rank}(\mathbf{Q}_2) = 2M - N + 1. \quad (\text{A.10})$$

For all $\mathbf{y} \in \text{range}(\mathbf{B})$, since $\mathbf{H}_t^T \mathbf{Q}_2 = \mathbf{0}$, we have

$$\begin{aligned} \mathbf{A}^T \mathbf{y} &= \mathbf{H}_t^T \mathbf{V}^{-1/2} \mathbf{V}^{1/2} \mathbf{Q}_2 \mathbf{z} = \mathbf{H}_t^T \mathbf{Q}_2 \mathbf{z} \\ &= \mathbf{0} \implies \text{range}(\mathbf{B}) \subset \text{null}(\mathbf{A}^T). \end{aligned} \quad (\text{A.11})$$

Comparing the dimension of subspaces, we have, from (A.8), (A.9), and (A.10); $\dim \text{null}(\mathbf{A}^T) = \dim \text{range}(\mathbf{B}) = 2M - N + 1$. We therefore conclude that $\text{range}(\mathbf{B}) = \text{null}(\mathbf{A}^T) = (\text{range}(\mathbf{A}))^\perp$, that is, $\mathbf{V}^{-1/2} \mathbf{B}_A^\perp \mathbf{V}^{-1/2} = \mathbf{V}^{-1/2} \mathbf{B}_B \mathbf{V}^{-1/2}$, and (A.5) holds for all parameter vectors Δ and \mathbf{x} , and all positive definite and symmetric covariance matrices \mathbf{V} .

B. FACTORIZATION OF THE JOINT PDF OF $\boldsymbol{\tau}$

Consider the sum of squares in the exponent of (14),

$$\begin{aligned} &(\boldsymbol{\tau} - \boldsymbol{\mu}_\tau)^T \mathbf{V}^{-1} (\boldsymbol{\tau} - \boldsymbol{\mu}_\tau) \\ &= (\boldsymbol{\tau} - \mathbf{H}_d \mathbf{d}(\mathbf{x}))^T \mathbf{V}^{-1} (\boldsymbol{\tau} - \mathbf{H}_d \mathbf{d}(\mathbf{x})) \\ &\quad - 2(\boldsymbol{\tau} - \mathbf{H}_d \mathbf{d}(\mathbf{x}))^T \mathbf{V}^{-1} \mathbf{H}_t \Delta + [\mathbf{H}_t \Delta]^T \mathbf{V}^{-1} \mathbf{H}_t \Delta. \end{aligned} \quad (\text{B.1})$$

We may factor and rewrite the first term in (B.1) as

$$\begin{aligned} &(\boldsymbol{\tau} - \mathbf{H}_d \mathbf{d}(\mathbf{x}))^T \mathbf{V}^{-1} (\boldsymbol{\tau} - \mathbf{H}_d \mathbf{d}(\mathbf{x})) \\ &= \boldsymbol{\tau}^T \mathbf{V}^{-1} \boldsymbol{\tau} - 2[\mathbf{H}_d \mathbf{d}(\mathbf{x})]^T \mathbf{V}^{-1} \boldsymbol{\tau} \\ &\quad + [\mathbf{H}_d \mathbf{d}(\mathbf{x})]^T \mathbf{V}^{-1} \mathbf{H}_d \mathbf{d}(\mathbf{x}). \end{aligned} \quad (\text{B.2})$$

Due to the special shape of \mathbf{V} , we may rewrite the second term of (B.2) as

$$2[\mathbf{H}_d \mathbf{d}(\mathbf{x})]^T \mathbf{V}^{-1} \boldsymbol{\tau} = 2\mathbf{d}(\mathbf{x})^T \mathbf{V}_2^{-1} \mathbf{H}_d^T \boldsymbol{\tau}, \quad (\text{B.3})$$

where \mathbf{V}_2 is given by (17), and $\mathbf{H}_d^T \boldsymbol{\tau}$ contain the sums of corresponding pTOAs. Again, from the shape of \mathbf{V} , we may rewrite the second term in (B.1) as

$$2(\boldsymbol{\tau} - \mathbf{H}_d \mathbf{d}(\mathbf{x}))^T \mathbf{V}^{-1} \mathbf{H}_t \Delta = 2[\mathbf{TH}_t \Delta]^T \mathbf{V}_2^{-1} \mathbf{D} \boldsymbol{\tau}, \quad (\text{B.4})$$

where \mathbf{T} , given by (19), selects every second element of $\mathbf{H}_t \Delta$, and \mathbf{D} , given by (18), takes the difference of corresponding pTOAs. The exponent in (B.1) can therefore be written

$$\begin{aligned} &(\boldsymbol{\tau} - \boldsymbol{\mu}_\tau)^T \mathbf{V}^{-1} (\boldsymbol{\tau} - \boldsymbol{\mu}_\tau) \\ &= [\mathbf{H}_d \mathbf{d}(\mathbf{x})]^T \mathbf{V}^{-1} \mathbf{H}_d \mathbf{d}(\mathbf{x}) - 2\mathbf{d}^T(\mathbf{x}) \mathbf{V}_2^{-1} \mathbf{H}_d^T \boldsymbol{\tau} \\ &\quad + \boldsymbol{\tau}^T \mathbf{V}^{-1} \boldsymbol{\tau} - 2[\mathbf{TH}_t \Delta]^T \mathbf{V}_2^{-1} \mathbf{D} \boldsymbol{\tau} + [\mathbf{H}_t \Delta]^T \mathbf{V}^{-1} \mathbf{H}_t \Delta. \end{aligned}$$

ACKNOWLEDGMENT

This work has been partially funded by Vinnova, project no. 2003-02803.

REFERENCES

- [1] R. L. Moses, D. Krishnamurthy, and R. M. Patterson, "A self-localization method for wireless sensor networks," *EURASIP Journal on Applied Signal Processing*, vol. 2003, no. 4, pp. 348–358, 2003.
- [2] N. Patwari, A. O. Hero III, M. Perkins, N. S. Correal, and R. J. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137–2148, 2003.
- [3] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall PTR, Upper Saddle River, NJ, USA, 1993.
- [4] M. Rydström, E. G. Ström, and A. Svensson, "Clock-offset cancellation methods for positioning in asynchronous sensor networks," in *Proceedings of IEEE International Conference on Wireless Networks, Communications, and Mobile Computing (WirelessCom '05)*, vol. 2, pp. 981–986, Maui, Hawaii, USA, June 2005.
- [5] M. Rydström, A. Urruela, E. G. Ström, and A. Svensson, "A low complexity algorithm for distributed sensor localization," in *Proceedings of the 11th European Wireless Conference (EW '05)*, vol. 2, pp. 714–718, Nicosia, Cyprus, April 2005.
- [6] L. L. Scharf and L. T. McWhorter, "Geometry of the Cramer-Rao bound," in *Proceedings of IEEE 6th SP Workshop on Statistical Signal and Array Processing*, pp. 5–8, Victoria, BC, Canada, October 1992.
- [7] Y. Qi, *Wireless geolocation in a non-line-of-sight environment*, Ph.D. thesis, Princeton University, Princeton, NJ, USA, 2003.
- [8] V. P. Bhapkar, "Estimating functions, partial sufficiency and q -sufficiency in the presence of nuisance parameters," in *Selected Proceedings of the Symposium on Estimating Functions*, Athens, Ga, USA, March 1996.
- [9] B. T. Fang, "Simple solutions for hyperbolic and related position fixes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 5, pp. 748–753, 1990.
- [10] A. Urruela and J. Riba, "Novel closed-form ML position estimator for hyperbolic location," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, vol. 2, pp. 149–152, Montreal, Quebec, Canada, May 2004.

- [11] M. Rydström, "Positioning and tracking in asynchronous wireless sensor networks," Tech. Rep. R027/2005, Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, October 2005.

Mats Rydström was born in Stockholm, Sweden, in 1978. He received his M.S. degree in computer engineering from Chalmers University of Technology, Göteborg, Sweden, in 2003. Mats was also enrolled at the Electrical Engineering Department at the University of Illinois at Chicago, during 2002, under a full scholarship. He is currently working toward his Ph.D. degree at the Communication Systems Group at Chalmers University of Technology, where his research interests include autonomous positioning algorithms for wireless sensor networks, and wireless networks for traffic safety applications.



Andreu Urruela was born in Castellbisbal, Barcelona, Spain, in 1978. He received the M.S. degree in telecommunications engineering in 2001 from the Technical University of Catalonia (UPC), Barcelona. Since September 2001, he has been a Graduate Research Assistant in the Signal Processing for Communications Group at UPC under the Spanish Government predoctoral scholarship FPU. He has been involved in the IST EMILY (European Mobile Integrated Location sYstem) project for the development of advanced algorithms for wireless location as a Member of the Signal Processing Group at UPC. He is currently working toward the Ph.D. degree. His research interests include high-accuracy time-delay estimators, closed-form algorithms for wireless location, sensor-network location, and the development of wireless location schemes for cellular-networks robust to multipath and nonline of sight.



Erik G. Ström received the M.S. degree from the Royal Institute of Technology (KTH), Stockholm, Sweden, in 1990, and the Ph.D. degree from the University of Florida, Gainesville, in 1994, both in electrical engineering. He accepted a Postdoctoral position at the Department of Signals, Sensors, and Systems at KTH in 1995. In February 1996, he was appointed Assistant Professor at KTH, and in June 1996, he joined the Department of Signals and Systems at Chalmers University of Technology, Göteborg, Sweden, where he is now a Professor in communication systems since June 2003 and Head of the Communication Systems Group since 2005. He received the Chalmers' Pedagogical Prize in 1998. Since 1990, he has acted as a Consultant for the Educational Group for Individual Development, Stockholm, Sweden. He is a contributing Author and Associate Editor for Roy Admiralty Publishers' FesGas-series, and was a Coguest Editor for the special issue of the IEEE Journal on Selected Areas in Communications on Signal Synchronization in Digital Transmission Systems, 2001. His research interests include code-division multiple access, synchronization, and wireless communications, and he has published more than 60 conference and journal papers.



Arne Svensson was born in Vedåkra, Sweden, on October 22, 1955. He received the M.S. (Civilingenjör) degree in electrical engineering from the University of Lund, Sweden in 1979, and the Dr.Ing. (Teknisk Licentiat) and Dr.Techn. (Teknisk Doktor) degrees at the Department of Telecommunication Theory, University of Lund, in 1982 and 1984, respectively. Currently, he is with the Department of Signal and Systems at Chalmers University of Technology, Gothenburg, Sweden, where he was appointed Professor and Chair in Communication Systems in April 1993 and Head of department from January 2005. Before 1987, he was with Department of Telecommunication Theory, University of Lund, Lund, Sweden, and between 1987 and 1994, he was with Ericsson Radio Systems AB and Ericsson Radar Electronics AB, both in Mölndal, Sweden. His current interest is wireless communication systems with special emphasis on physical layer design and analysis. He is the Coauthor of *Coded Modulation Systems* (Norwell, MA: Kluwer Academic/Plenum, 2003). He has also published four book chapters, 34 journal papers/letters, and more than 150 conference papers. He received the IEEE Vehicular Technology Society Paper of the Year Award in 1986. He is a Fellow of IEEE, an Editor for IEEE Transactions on Wireless Communications, and a Member of the council of NRS (Nordic Radio Society).

