Research Article

A Comparison of Parametric and Sample-Based Message Representation in Cooperative Localization

Jaime Lien,[1] Ulric J. Ferner,[2] Warakorn Srichavengsup,[3] Henk Wymeersch,[4] and Moe Z. Win[5]

[1] *Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA*
[2] *Research Laboratory of Electronics, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA*
[3] *Department of Computer Engineering, Thai-Nichi Institute of Technology, Bangkok 10250, Thailand*
[4] *Department of Signals and Systems, Chalmers University of Technology, Gothenburg 412 96, Sweden*
[5] *Laboratory for Information and Decision Systems, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA*

Correspondence should be addressed to Jaime Lien, jlien@stanford.edu

Location awareness is a key enabling feature and fundamental challenge in present and future wireless networks. Most existing localization methods rely on existing infrastructure and thus lack the flexibility and robustness necessary for large ad hoc networks. In this paper, we build upon SPAWN (sum-product algorithm over a wireless network), which determines node locations through iterative message passing, but does so at a high computational cost. We compare different message representations for SPAWN in terms of performance and complexity and investigate several types of cooperation based on censoring. Our results, based on experimental data with ultra-wideband (UWB) nodes, indicate that parametric message representation combined with simple censoring can give excellent performance at relatively low complexity.

1. Introduction

Location awareness has the potential to revolutionize a diverse array of present and future technologies. Accurate knowledge of a user's location is essential for a wide variety of commercial, military, and social applications, including next-generation cellular services [1, 2], sensor networks [3, 4], search-and-rescue [5, 6], military target tracking [7, 8], health care monitoring [9, 10], robotics [11, 12], data routing [13, 14], and logistics [15, 16]. Typically, only a small fraction of the nodes in the network, known as *anchors*, have prior knowledge about their location. The remaining nodes, known as *agents*, must determine their locations through a process of localization or positioning. The ad hoc and often dynamic nature of wireless networks requires distributed and autonomous localization methods. Moreover, location-aware wireless networks are frequently deployed in unknown environments and hence can rely only on minimal (if any) infrastructure, human maintenance, and a priori location information.

Cooperation is an emerging paradigm for localization in which agents take advantage of network connections and interagent measurements to improve their location estimates. Non-Bayesian cooperative localization in wireless sensor networks is discussed in [17]. Different variations of Bayesian cooperation have been considered, including Monte-Carlo sequential estimation [18] and nonparametric belief propagation in static networks [19]. For a comprehensive overview of Bayesian and non-Bayesian cooperative localization in wireless networks, we refer the reader to [20], which also introduces a distributed cooperative algorithm for large-scale mobile networks called SPAWN (sum-product algorithm over a wireless network). This message-passing algorithm achieves improved localization accuracy and coverage compared to other methods and will serve as the basic algorithm in this paper.

The complexity and cost associated with the SPAWN algorithm depend largely on how messages are represented for computation and transmission. As wireless networks typically operate under tight power and resource constraints, the

choice of message representation heavily impacts the feasibility and ease of implementation of the algorithm. The method of message representation and ensuing tradeoff between communication cost and localization performance are thus of great practical importance in the deployment of realistic localization systems. Particle methods do not necessarily lend themselves well in practice to be exchanged wirelessly between devices, due to their high computational complexity and communication overhead [21]. Other message-passing methods have been developed that rely on parametric message representation, thus alleviating these drawbacks but limiting representational flexibility. In particular, in [22], the expectation propagation algorithm is considered with Gaussian messages, while in [23], variational message passing with parametric messages is shown to exhibit low complexity. A variation of SPAWN combining GPS and UWB was evaluated in [24], using a collection of parametric distributions with ellipsoidal, conic, and cylindrical shapes.

This paper addresses the need for accurate, resource-efficient localization with an in-depth comparison of various message representations for SPAWN. We describe and evaluate different parametric and nonparametric message representations in terms of complexity and accuracy. Additionally, we analyze the performance of various cooperative schemes and message representations in a simulated large-scale ultra-wide bandwidth (UWB) network using experimental UWB ranging data. UWB is an attractive choice for ranging and communication due to its ability to resolve multipath [25, 26], penetrate obstacles [27], and provide high resolution distance measurements [28, 29]. Recent research advances in UWB signal acquisition [30, 31], multiuser interference [32, 33], multipath channels [34, 35], non-line-of-sight (NLOS) propagation [28, 29], and time-of-arrival estimation [36] increase the potential for highly accurate UWB-based localization systems in harsh environments. Consequently, significant attention has been paid to both algorithm design [37–42] and fundamental limits of accuracy [43–46] for UWB localization. It is expected that UWB will be exploited in future location-aware systems that utilize coexisting networks of sensors, controllers, and peripheral devices [47, 48].

2. Problem Formulation

We consider a wireless network of N nodes in an environment $\mathcal{E}$. Time is slotted with nodes moving independently from time slot to time slot. The position of node i at time t is described by the random variable $\mathbf{x}_i^{(t)}$; the vector of all positions is denoted by $\mathbf{x}^{(t)}$. At each time t, node i may collect internal position-related measurements $z_{i,\text{self}}^{(t)}$, for example, from an inertial measurement unit. The set of all internal measurements is denoted by $\mathbf{z}_{\text{self}}^{(t)}$. Within the network, nodes communicate with each other via wireless transmissions. We denote the set of nodes from which node i can receive transmissions at time t by $\mathcal{S}_{\rightarrow i}^{(t)}$. Note that the communication link may not be bidirectional; that is, $j \in \mathcal{S}_{\rightarrow i}^{(t)}$ does not imply $i \in \mathcal{S}_{\rightarrow j}^{(t)}$. Using packets received from $j \in \mathcal{S}_{\rightarrow i}^{(t)}$, node i may collect a set of relative measurements, represented by the vector $z_{j \rightarrow i}^{(t)}$, which we will limit to distance measurements. We denote the set of all relative measurements made in the network at time t by $\mathbf{z}_{\text{rel}}^{(t)}$. The full set of relative and internal measurements is denoted $\mathbf{z}^{(t)}$.

The objective of the localization problem is for each node i to determine the a posteriori distribution $p(\mathbf{x}_i^{(t)} \mid \mathbf{z}^{(1:t)})$ of its position $\mathbf{x}_i^{(t)}$ at each time t, given information up to and including t.

3. A Brief Introduction to SPAWN

In [20], we proposed a cooperative localization algorithm by factorizing the joint distribution $p(\mathbf{x}^{(0:T)} \mid \mathbf{z}^{(1:T)})$, formulating the problem as a factor graph with temporal and spatial constraints, and applying the sum-product algorithm. This leads to a distributed algorithm, known as SPAWN, presented in Algorithm 1. The aim of SPAWN is to compute a *belief* $b_i^{(t)}(\mathbf{x}_i^{(t)})$ available to node i at the end of any time slot t, which serves as an approximation of the marginal a posteriori distribution $p(\mathbf{x}_i^{(t)} \mid \mathbf{z}^{(1:t)})$. Note that each operation of SPAWN requires only information local to an individual node. Information is shared between nodes via physical transmissions. Each node can therefore perform the computations in Algorithm 1 using its local information and transmissions received from neighboring nodes.

Observe that Algorithm 1 contains a number of key steps.

(i) *Mobility update* (line 4), requiring knowledge of mobility models $p(\mathbf{x}_i^{(t)} \mid \mathbf{x}_i^{(t-1)})$ and self-measurement likelihood functions $p(z_{i,\text{self}}^{(t)} \mid \mathbf{x}_i^{(t-1)}, \mathbf{x}_i^{(t)})$.

(ii) *Message conversion* (line 10) of position information from neighboring devices to account for relative measurements, requiring knowledge of the neighbors and of relative measurement likelihood functions $p(z_{j \rightarrow i}^{(t)} \mid \mathbf{x}_i, \mathbf{x}_j)$.

(iii) *Belief update* (line 11), to fuse information from the mobility update with information from the current neighbors.

The first two operations can be interpreted as *message filtering*, while the latter operation is a *message multiplication*. How these operations can be implemented in practice will be the topic of Section 4.

4. Message Representation

4.1. Key Operations. In SPAWN, probabilistic information is exchanged and computed through *messages*. The manner in which these messages are represented for transmission between nodes and internal computation is closely related to the complexity and performance of the localization algorithm. In traditional communications problems, such as decoding, messages can be represented efficiently and exactly through, for instance, log-likelihood ratios [49]. In SPAWN, exact representation is impossible, so we must resort to different types of approximate message representations. Any representation must be able to capture the salient properties of the true message and must enable efficient computation

(1) Initialize belief $b^{(0)}(\mathbf{x}_i^{(0)}) = p(\mathbf{x}_i^{(0)})$, $\forall i$
(1) **for** $t = 1$ to T **do** {time index}
(2)   **for all** i **do** {mobility update}
(5)   Mobility update:
$$\tilde{b}_i^{(t)}(\mathbf{x}_i^{(t)}) \propto$$
$$\int p(\mathbf{x}_i^{(t)}|\mathbf{x}_i^{(t-1)})p(z_{i,\text{self}}^{(t)}|\mathbf{x}_i^{(t-1)},\mathbf{x}_i^{(t)})$$
$$\times b_i^{(t-1)}(\mathbf{x}_i^{(t-1)})\,d\mathbf{x}_i^{(t-1)} \qquad \text{(A-1)}$$
(6)   **end for**
(7)   Initialize $b_i^{(t)}(\mathbf{x}_i^{(t)}) = \tilde{b}_i^{(t)}(\mathbf{x}_i^{(t)})$, $\forall i$
(8)   **for** $l = 1$ to N_{it} **do** {iteration index; begin cooperative update}
(9)   **for all** i **do**
(10)   **for all** $j \in S_{\to i}^{(t)}$ **do**
(11)   Receive and convert $b_j^{(t)}(\mathbf{x}_j^{(t)})$ to a distribution $c_{j\to i}^{(l)}(\mathbf{x}_i^{(t)})$:
$$c_{j\to i}^{(l)}(\mathbf{x}_i^{(t)}) \propto$$
$$\int p(z_{j\to i}^{(t)}|\mathbf{x}_i^{(t)},\mathbf{x}_j^{(t)})b_j^{(t)}(\mathbf{x}_j^{(t)})d\mathbf{x}_j^{(t)} \qquad \text{(A-2)}$$
(12)   Update and broadcast $b_i^{(t)}(\mathbf{x}_i^{(t)})$:
$$b_i^{(t)}(\mathbf{x}_i^{(t)}) \propto \tilde{b}_i^{(t)}(\mathbf{x}_i^{(t)}) \prod_{k\in S_{\to i}^{(t)}} c_{k\to i}^{(l)}(\mathbf{x}_i^{(t)}) \qquad \text{(A-3)}$$
(13)   **end for**
(14)   **end for**
(15)   **end for** {end cooperative update}
(16) **end for** {end current time step}

ALGORITHM 1: SPAWN.

of the key steps in SPAWN, namely, message *filtering* (A-1)-(A-2) and message *multiplication* (A-3). We consider three types of message representation: discretized, sample-based, and parametric.

For convenience, we will introduce a set of new notations. For the filtering operation, the incoming message is denoted by $p_{\mathbf{X}}(\mathbf{x})$, the filtering operation by $h(\mathbf{x},\mathbf{y})$, and the outgoing message by $p_{\mathbf{Y}}(\mathbf{y})$, with

$$p_{\mathbf{Y}}(\mathbf{y}) \propto \int h(\mathbf{x},\mathbf{y})\,p_{\mathbf{X}}(\mathbf{x})d\mathbf{x}. \qquad (1)$$

For the multiplication operation, we assume M incoming messages $p_{\mathbf{X}}^{(i)}(\mathbf{x})$ ($i = 1,\dots,M$) over a single variable $\mathbf{X}$, and an outgoing message

$$\phi_{\mathbf{X}}(\mathbf{x}) \propto \prod_{i=1}^{M} p_{\mathbf{X}}^{(i)}(\mathbf{x}). \qquad (2)$$

Note that (1) maps to (A-1) through the following association:

$$\mathbf{x} \longrightarrow \mathbf{x}_i^{(t-1)},$$
$$\mathbf{y} \longrightarrow \mathbf{x}_i^{(t)},$$
$$h(\mathbf{x},\mathbf{y}) \longrightarrow p\left(\mathbf{x}_i^{(t)} \mid \mathbf{x}_i^{(t-1)}\right) p\left(z_{i,\text{self}}^{(t)} \mid \mathbf{x}_i^{(t-1)},\mathbf{x}_i^{(t)}\right), \qquad (3)$$
$$p_{\mathbf{X}}(\mathbf{x}) \longrightarrow b_i^{(t-1)}\left(\mathbf{x}_i^{(t-1)}\right).$$

Similarly, (1) maps to (A-2) through the following association:

$$\mathbf{x} \longrightarrow \mathbf{x}_j^{(t)},$$
$$\mathbf{y} \longrightarrow \mathbf{x}_i^{(t)},$$
$$h(\mathbf{x},\mathbf{y}) \longrightarrow p\left(z_{j\to i}^{(t)} \mid \mathbf{x}_i^{(t)},\mathbf{x}_j^{(t)}\right), \qquad (4)$$
$$p_{\mathbf{X}}(\mathbf{x}) \longrightarrow b_i^{(t)}\left(\mathbf{x}_i^{(t)}\right).$$

4.2. Discretized Message Representation. A naive but simple approach to represent a continuous distribution $p_{\mathbf{X}}(\mathbf{x})$ is to uniformly discretize the domain of $\mathbf{X}$, yielding a set of quantization points $Q = \{\mathbf{x}_1,\dots,\mathbf{x}_R\}$. The distribution is then approximated as a finite list of values, $\{p_{\mathbf{X}}(\mathbf{x}_k)\}_{k=1}^{R}$. The filtering operation then becomes

$$p_{\mathbf{Y}}(\mathbf{y}_k) \propto \sum_{l=1}^{R} h(\mathbf{x}_l,\mathbf{y}_k)\,p_{\mathbf{X}}(\mathbf{x}_l), \qquad (5)$$

requiring $\mathcal{O}(R^2)$ operations. The multiplication becomes

$$\phi_{\mathbf{X}}(\mathbf{x}_k) \propto \prod_{i=1}^{M} p_{\mathbf{X}}^{(i)}(\mathbf{x}_k), \qquad (6)$$

requiring $\mathcal{O}(RM)$ operations. Because R scales exponentially with the dimensionality of $\mathbf{X}$ and a large number of points are required in every dimension to capture fine features of the messages, discretization is impractical for SPAWN in UWB localization.

4.3. Sample-Based Message Representation. A sample-based message representation, as used in [19, 50], overcomes the drawback of discretization by representing messages as samples, concentrated where the messages have significant mass. Before describing the detailed implementation of the filtering and multiplication operations, we give a brief overview of generic sampling techniques (see also [51, 52]) and kernel density estimation (KDE).

4.3.1. Background: Sampling and Kernel Density Estimation. We say that a list of samples with associated weights $\{\mathbf{x}_k, w_k\}_{k=1}^{R}$ is a representation for a distribution $p_{\mathbf{X}}(\mathbf{x})$ if, for any integrable function $g(\mathbf{x})$, we have the following approximation:

$$I = \int g(\mathbf{x}) p_{\mathbf{X}}(\mathbf{x}) dx \approx \sum_{k=1}^{R} w_k g(\mathbf{x}_k). \tag{7}$$

Popular methods for obtaining the list of weighted samples include (i) direct sampling, where we draw R i.i.d. samples from $p_{\mathbf{X}}(\mathbf{x})$, each with weight $1/R$; and (ii) importance sampling, where we draw R i.i.d. samples from a distribution $q_{\mathbf{X}}(\mathbf{x})$, with a support that includes the support of $p_{\mathbf{X}}(\mathbf{x})$, and set the weight corresponding to sample $\mathbf{x}_k$ as $w_k = p_{\mathbf{X}}(\mathbf{x}_k)/q_{\mathbf{X}}(\mathbf{x}_k)$. In both cases, it can easily be verified that the approximation is unbiased with mean I and variance that reduces with R (and that depends on $q_{\mathbf{X}}(\mathbf{x})$, for importance sampling). Most importantly, the variance does not depend on the dimensionality of $\mathbf{x}$.

A variation of importance sampling that is not unbiased but that often has smaller variance is obtained by setting the weights as follows: $w_k \propto p_{\mathbf{X}}(\mathbf{x}_k)/q_{\mathbf{X}}(\mathbf{x}_k)$, $\sum_k w_k = 1$. This approach has the additional benefit that it does not require knowledge of the normalization constants of $p_{\mathbf{X}}(\mathbf{x})$ or $q_{\mathbf{X}}(\mathbf{x})$. A list of R equally weighted samples can be obtained from $\{\mathbf{x}_k, w_k\}_{k=1}^{R}$ through resampling, that is, by drawing (with repetition) R samples from the probability mass function defined by $\{\mathbf{x}_k, w_k\}_{k=1}^{R}$.

For numerical stability reasons, weights are often computed and stored in the logarithmic domain, that is, $\lambda_k = \log p_{\mathbf{X}}(\mathbf{x}_k) - \log q_{\mathbf{X}}(\mathbf{x}_k)$. When the distributions involved contain exponentials or products, the log-domain representation is also computationally efficient. Operations such as additions can be evaluated efficiently in the log-domain as well, using the Jacobian logarithm [49, pages 90–94]. Once all R log-domain weights are computed, they are translated, exponentiated, and normalized: $w_k \propto \exp(\lambda_k - \max_l \lambda_l)$.

Given a sample representation $\{\mathbf{x}_k, w_k\}_{k=1}^{R}$ of a distribution $p_{\mathbf{X}}(\mathbf{x})$, we obtain a kernel density estimate of $p_{\mathbf{X}}(\mathbf{x})$ as

$$\hat{p}_{\mathbf{X}}(\mathbf{x}) = \sum_{k=1}^{R} w_k K_\sigma(\mathbf{x} - \mathbf{x}_k), \tag{8}$$

where $K_\sigma(\mathbf{x})$ is the so-called *kernel* with bandwidth σ. The kernel is a symmetric distribution with a width parameter

that is tuned through σ. For instance, a two-dimensional Gaussian kernel is given by

$$K_\sigma(\mathbf{x}) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|\mathbf{x}\|^2}{2\sigma^2}\right). \tag{9}$$

While the choice of kernel affects the performance of the estimate to some limited extent (e.g., in an MMSE sense, where the error is $\int |p_X(x) - \hat{p}_X(x)|^2 p_X(x) dx$), the crucial parameter is the bandwidth σ, which needs to be estimated from the samples $\{\mathbf{x}_k, w_k\}_{k=1}^{R}$. A large choice of σ makes $\hat{p}_{\mathbf{X}}(\mathbf{x})$ smooth, but it may no longer capture the interesting features of $p_{\mathbf{X}}(\mathbf{x})$. When σ is too small, $\hat{p}_{\mathbf{X}}(\mathbf{x})$ may exhibit artificial structure not present in $p_{\mathbf{X}}(\mathbf{x})$ [53].

With this background in sampling techniques and KDE, we return to the problem at hand: filtering and multiplication of messages.

4.3.2. Message Filtering. We assume a message representation of $p_{\mathbf{X}}(\mathbf{x})$ as $\{\mathbf{x}_k, w_k\}_{k=1}^{R}$ and wish to obtain a message representation of $p_{\mathbf{Y}}(\mathbf{y}) \propto \int h(\mathbf{x}, \mathbf{y}) p_{\mathbf{X}}(\mathbf{x}) dx$. Let us interpret $h(\mathbf{x}, \mathbf{y})$ as a conditional distribution $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x})$, up to some arbitrary constant. Suppose we can draw samples $\{[\mathbf{x}_k, \mathbf{y}_k], w_k\}_{k=1}^{R} \sim p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x}) p_{\mathbf{X}}(\mathbf{x})$; then $\{\mathbf{y}_k, w_k\}_{k=1}^{R}$ will form a sample representation of $p_{\mathbf{Y}}(\mathbf{y})$. Now the problem reverts to drawing samples from $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x}) p_{\mathbf{X}}(\mathbf{x})$. This can be accomplished as follows: first, for every sample $\mathbf{x}_k$, draw $\mathbf{y}_k \sim q_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x}_k)$ from some distribution $q_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x}_k)$. Second, set the weight of sample $[\mathbf{x}_k, \mathbf{y}_k]$ as

$$v_k = w_k \frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}_k \mid \mathbf{x}_k)}{q_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}_k \mid \mathbf{x}_k)}. \tag{10}$$

Finally, renormalize the weights v_k to $v_k / \sum_l v_l$. The complexity of the filtering operation scales as $\mathcal{O}(R)$, a significant improvement from $\mathcal{O}(R^2)$ for discretization. In addition, R can generally be much smaller in a particle-based representation.

Let us consider some examples of the filtering operation in SPAWN.

(i) *Mobility update* (A-1): let $p_{\mathbf{X}}(\mathbf{x})$ be the belief before movement (represented by $\{\mathbf{x}_k, w_k\}_{k=1}^{R}$) and $p_{\mathbf{Y}}(\mathbf{y})$ the belief after movement. Assume that we are able to measure perfectly the distance traveled (given by z_{self}), but have no information regarding the direction, and furthermore that the direction is chosen uniformly in $(0, 2\pi]$. In that case,

$$h(\mathbf{x}, \mathbf{y}) \propto \delta(z_{\text{self}} - \|\mathbf{x} - \mathbf{y}\|), \tag{11}$$

where δ is a Dirac delta function, so that $q_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x}_k) = p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x}) \propto \delta(z_{\text{self}} - \|\mathbf{x}_k - \mathbf{y}\|)$ is a reasonable choice. For every $\mathbf{x}_k$, we can now draw values for $\mathbf{y} = \mathbf{x}_k + r \times [\cos\theta \sin\theta]^T$ by drawing $\theta \sim \mathcal{U}(0, 2\pi)$ and setting $r = z_{\text{self}}$, leading to $\mathbf{y}_k = \mathbf{x}_k + z_{\text{self}} \times [\cos\theta_k \sin\theta_k]^T$, with $v_k \propto w_k$.

(ii) *Ranging update* (A-2): let $p_{\mathbf{X}}(\mathbf{x})$ be a message (represented by $\{\mathbf{x}_k, w_k\}_{k=1}^{R}$) from a node with which we

have performed ranging, resulting in a range estimate z. Let $h(\mathbf{x}, \mathbf{y}) = p_{Z|D}(z \mid d)$, where $d = \|\mathbf{x} - \mathbf{y}\|$. Note that $p_{Z|D}(z \mid d)$ is a likelihood function, since the measurement z is known. Assume that we have a model for the ranging performance in the form of distributions $p_{Z|D}(z \mid d)$ for any value of d. We then sample $p_{\mathbf{Y}}(\mathbf{y})$ as follows: for every $\mathbf{x}_k$, draw $\mathbf{y} = \mathbf{x}_k + r \times [\cos\theta \sin\theta]^T$ by drawing $\theta \sim \mathcal{U}(0, 2\pi)$ and $r \sim q_{R|Z}(r \mid z)$, for some well-chosen $q_{R|Z}(r \mid z)$ (e.g., a Gaussian distribution with mean equal to the distance estimate, z, and a standard deviation that is sufficiently large with respect to the standard deviation of $p(z|d)$ for any d). The weights are set as

$$ v_k = w_k \frac{p_{Z|D}(z \mid r_k)}{q_{R|Z}(r_k z)}. \tag{12} $$

4.3.3. Message Multiplication. Here we assume message representations $\{\mathbf{x}_k^{(i)}, w_k^{(i)}\}_{k=1}^{R}$ for $p_{\mathbf{X}}^{(i)}(\mathbf{x}_k)$, $i = 1, \ldots, M$. In contrast to the discretization approach, we cannot directly compute $\prod_{i=1}^{M} p_{\mathbf{X}}^{(i)}(\mathbf{x})$ for arbitrary values of $\mathbf{x}$. Rather, for every message $p_{\mathbf{X}}^{(i)}(\mathbf{x}_k)$, we create a KDE $\hat{p}_{\mathbf{X}}^{(i)}(\mathbf{x}) = \sum_{k=1}^{R} w_k K_{\sigma^{(i)}}(\mathbf{x} - \mathbf{x}_k^{(i)})$ with a Gaussian kernel and a bandwidth estimated using the methods from [53]. Suppose we now draw R samples from a distribution $q_{\mathbf{X}}(\mathbf{x})$; then the weights are

$$ v_k \propto \frac{\prod_{i=1}^{M} \hat{p}_{\mathbf{X}}^{(i)}(\mathbf{x}_k)}{q_{\mathbf{X}}(\mathbf{x}_k)}, \tag{13} $$

which can be computed efficiently in the log-domain. A reasonable choice for $q_{\mathbf{X}}(\mathbf{x})$ could be one of the incoming messages $p_{\mathbf{X}}^{(i)}(\mathbf{x})$ (e.g., the one with the smallest entropy) or a mixture of the incoming messages. The computational complexity of the message multiplication operation scales as $\mathcal{O}(MR^2)$. This appears worse than the discretized case (complexity $\mathcal{O}(MR)$), but note that R is much smaller for sample-based representations than for discretization (e.g., $R = 10^3$ or $R = 10^4$ for the sample-based representation compared to $R = 10^8$ in the discretization).

4.4. Parametric Message Representation

4.4.1. Choosing a Suitable Parameterization. From the previous section, it is clear that the bottleneck of the sample-based message representation lies in the message multiplication, which scales quadratically with the number of samples. An alternative approach is to represent each message as a set of parameters (e.g., a Gaussian distribution characterized by a mean and covariance matrix). In contrast to the sample-based message representation, which can represent messages of any shape, parametric representations must be specially tailored to the problem at hand. For example, single two-dimensional Gaussian parametric messages are utilized in [22] for localization with both range and angle measurements. Our choice of parametric message is based on the following observations.

(i) For the filtering operation with a two-dimensional Gaussian input $p_{\mathbf{X}}(\mathbf{x})$, the output $p_{\mathbf{Y}}(\mathbf{y})$ can be approximated by a circular distribution with the same mean for both the mobility update (A-1) and the ranging update (A-2).

(ii) Multiplying Gaussian distributions yield a Gaussian distribution.

(iii) The multiplication of multiple circular distributions can be approximated by a Gaussian distribution or a mixture of Gaussian distributions.

We will use as a basic building block the following distribution in two dimensions:

$$ \mathcal{D}(\mathbf{x}; m_1, m_2, \sigma^2, \rho) $$

$$ = \frac{1}{C(\sigma^2, \rho)} \exp\left\{ -\frac{\left[\sqrt{(x_1 - m_1)^2 + (x_2 - m_2)^2} - \rho\right]^2}{2\sigma^2} \right\}, \tag{14} $$

where $[m_1, m_2]$ is the midpoint of the distribution, ρ is the radius, σ^2 is the variance, and $C(\sigma^2, \rho)$ is a normalization constant equal to

$$ C(\sigma^2, \rho) $$

$$ = 2\pi\sigma^2 \left[\exp\left(-\frac{\rho^2}{2\sigma^2}\right) + \frac{1}{2}\sqrt{\frac{2\pi\rho^2}{\sigma^2}} \left(1 + \mathrm{erf}\sqrt{\frac{\rho^2}{2\sigma^2}}\right) \right]. \tag{15} $$

As a special case, we note that, when $\rho = 0$, (14) reverts to a two-dimensional Gaussian. Moreover, we will represent all messages as a mixture of two distributions of the type (14), so that

$$ p_{\mathbf{X}}(\mathbf{x}) $$

$$ = \frac{1}{2}\mathcal{D}\left(\mathbf{x}; m_1^{(a)}, m_2^{(a)}, \sigma^2, \rho\right) + \frac{1}{2}\mathcal{D}\left(\mathbf{x}; m_1^{(b)}, m_2^{(b)}, \sigma^2, \rho\right), \tag{16} $$

which can be represented by the six-dimensional vector $[m_1^{(a)}, m_2^{(a)}, m_1^{(b)}, m_2^{(b)}, \rho, \sigma^2]$. We will denote the family of distributions of the form (16) by $\mathcal{D}_2$. Note that it is trivial to extend this distribution, which is designed for two-dimensional localization systems, for use in three-dimensional systems. Before we describe the message filtering and message multiplication operations, let us first show how the parameters of (14) can be estimated from a list of samples.

4.4.2. ML Estimation of the Parameters m_1, m_2, σ^2, ρ. Given a list of samples $\{\mathbf{x}_k\}_{k=1}^{R}$, we can estimate the parameters $[m_1, m_2, \sigma^2, \rho]$ as follows. The midpoint $\mathbf{m} = [m_1, m_2]$ is estimated by

$$ \hat{\mathbf{m}} = \frac{1}{R}\sum_{k=1}^{R} \mathbf{x}_k. \tag{17} $$

To find the radius ρ and variance σ^2 of $\mathcal{D}$-distribution, we use maximum likelihood (ML) estimation, assuming the R samples are independent. Introducing $\boldsymbol{\alpha} = [\rho \sigma^2]^T$, we find that

$$\hat{\boldsymbol{\alpha}}_{\mathrm{ML}} = \arg\max_{\boldsymbol{\alpha}} \underbrace{\sum_{k=1}^{R} \log p_{\mathbf{X}}(\mathbf{x}_k; \hat{\mathbf{m}}, \boldsymbol{\alpha})}_{\Lambda(\boldsymbol{\alpha})}. \qquad (18)$$

Treating the log-likelihood function (LLF) $\Lambda(\boldsymbol{\alpha})$ as an objective function, we find its maximum through the gradient ascent algorithm

$$\hat{\boldsymbol{\alpha}}^{(n+1)} = \hat{\boldsymbol{\alpha}}^{(n)} - \varepsilon \nabla_{\boldsymbol{\alpha}} \Lambda(\boldsymbol{\alpha}) \big|_{\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}^{(n)}}, \qquad (19)$$

where ε is a suitably small step size, and the gradient vector can be approximated using finite differences. To initialize (19), we consider two initial estimates for σ^2 and ρ: one assuming $\rho \ll \sigma$ and a second assuming $\rho \gg \sigma$. The LLF is evaluated for both preliminary solutions, and the one with the largest log-likelihood is used as the initial estimate in (19).

4.4.3. Message Filtering. To perform message filtering, we use the fact that sample-based message filtering is a low-complexity operation. We decompose $p_{\mathbf{X}}(\mathbf{x})$, represented in parametric form, into its two mixture components. From each component, we draw $R/2$ samples and perform sample-based message filtering, as outlined in Section 4.3.2. We can then estimate the new $\mathcal{D}$-parameters for each mixture component using the ML method described above. We thus have $p_{\mathbf{Y}}(\mathbf{y})$ in parametric form. The complexity of this operation scales as $\mathcal{O}(R)$.

4.4.4. Message Multiplication. The motivation for using the parametric message representation is to avoid the complexity associated with sample-based message multiplication. Given M distributions $p_{\mathbf{X}}^{(i)}(\mathbf{x}) \in \mathcal{D}_2$, our goal is to compute

$$\phi_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{M} p_{\mathbf{X}}^{(i)}(\mathbf{x}). \qquad (20)$$

Typically, $\phi_{\mathbf{X}}(\mathbf{x}) \notin \mathcal{D}_2$, so we will *approximate* $\phi_{\mathbf{X}}(\mathbf{x})$ by $q_{\mathbf{X}}^*(\mathbf{x}) \in \mathcal{D}_2$ by projecting $\phi_{\mathbf{X}}(\mathbf{x})$ onto the family $\mathcal{D}_2$:

$$q_{\mathbf{X}}^*(\cdot) = \arg\min_{q_{\mathbf{X}} \in \mathcal{D}_2} D_{\mathrm{KL}}(q_{\mathbf{X}} \| \phi_{\mathbf{X}}), \qquad (21)$$

where $D_{\mathrm{KL}}(\cdot \| \cdot)$ denotes the Kullback Leibler (KL) divergence, defined as

$$D_{\mathrm{KL}}(q_{\mathbf{X}} \| \phi_{\mathbf{X}}) = \int q_{\mathbf{X}}(\mathbf{x}) \log \frac{q_{\mathbf{X}}(\mathbf{x})}{\phi_{\mathbf{X}}(\mathbf{x})} d\mathbf{x}. \qquad (22)$$

Observe that all elements of $\mathcal{D}_2$ are characterized by the parameters $\mathbf{p} \triangleq [m_1^{(a)}, m_2^{(a)}, m_1^{(b)}, m_2^{(b)}, \rho, \sigma^2]^T$ and that the optimization (21) is therefore a six-dimensional problem over all possible $\mathbf{p}$. The divergence $D_{\mathrm{KL}}(q \| \phi)$ for an arbitrary

TABLE 1: Comparison of complexity of message representations, where R is the number samples taken from each distribution and M is the number of distributions or messages.

Approach	Operation	Complexity	Value of R
Discretized	Filtering	$\mathcal{O}(R^2)$	Large
Discretized	Multiplication	$\mathcal{O}(RM)$	Large
Sample-based	Filtering	$\mathcal{O}(R)$	Small
Sample-based	Multiplication	$\mathcal{O}(R^2 M)$	Small
Parametric	Filtering	$\mathcal{O}(R)$	Small
Parametric	Multiplication	$\mathcal{O}(RM)$	Small

$\mathbf{p} \in \mathbb{R}^4 \times \mathbb{R}_+^2$ can be determined using Monte-Carlo integration as follows. We rewrite (22) as

$$D_{\mathrm{KL}}(q_{\mathbf{x}} \| \phi_{\mathbf{x}}) = \int q_{\mathbf{X}}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}, \qquad (23)$$

where $f(\mathbf{x}) = \log q_{\mathbf{X}}(\mathbf{x}) - \sum_{i=1}^{M} \log p_{\mathbf{X}}^{(i)}(\mathbf{x})$. By drawing R weighted samples $\{w_k, \mathbf{x}_k\}_{k=1}^{R}$ from $q_{\mathbf{X}}(\mathbf{x})$ (e.g., through importance sampling), we can approximate (23) by

$$D_{\mathrm{KL}}(q_{\mathbf{x}} \| \phi_{\mathbf{x}}) \approx \sum_{k=1}^{R} w_k f(\mathbf{x}_k). \qquad (24)$$

Using this approximation, the six-dimensional optimization problem (21) is solved through gradient descent, similar to (19). The complexity of this operation scales as $\mathcal{O}(RM)$. The initial estimate of $\mathbf{p}$ is obtained through a set of heuristics: we first decide whether $\phi_{\mathbf{X}}(\mathbf{x})$ can reasonably be represented by a distribution in $\mathcal{D}_2$. If not, the outgoing message is not computed. Otherwise, we use a geometric argument to find at most two midpoints. The initial estimates for ρ and σ^2 are set to a small constant value.

4.5. Comparison of Message Representations. The complexities of the discretized, sample-based, and parametric message representations are compared in Table 1.

5. Performance Analysis

In this section, we compare the performance of the SPAWN algorithm with sample-based versus parametric message representation in a simulated wireless network. We also analyze the use of different subsets of information in the algorithm and its effect on localization performance.

5.1. Simulation Setup and Performance Measures. We simulate a large-scale ultra-wide bandwidth (UWB) network in a 100 m × 100 m homogeneous environment, with 100 uniformly distributed agents and 13 fixed anchors in a grid configuration. Each node is able to measure its range to other nodes within 20 meters. The simulated ranging measurements are independently drawn from the UWB ranging model developed in [54]. The model, based on data collected in a variety of indoor scenarios, consists of three component Gaussian densities, where the mean and variance

of each component are experimentally determined functions of the true distance between the ranging nodes. To decouple the effect of mobility with the message representation, we consider a single time slot, where every agent has a uniform a priori distribution over the environment $\mathcal{E}$. SPAWN was run for $N_{it} = 20$ iterations, though convergence was generally achieved well before 10 iterations. For the sample-based representation, the number of samples is set to $R = 2048$ unless otherwise stated.

We quantify localization performance using the complementary cumulative distribution function (CCDF) of the localization error $e = \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|$, where $\hat{\mathbf{x}}_i$ is the estimated location of node i, taken as the mean of the belief, similar to [18]. To estimate the CCDF, we consider 50 random network topologies and collect position estimates at every iteration for every agent. Note that a CCDF of 0.01 at an error of, say, $e = 1$ m means that 99% of the nodes have an error less than 1 meter.

5.2. Cooperation with Censoring. In Section 3, we considered processing messages between all neighboring pairs of nodes. However, information from neighbors may not always be useful: (i) when the receiving node's belief is already very informative (e.g., concentrated around the mean); or (ii) when the transmitting node's belief is very uninformative. To better understand how much cooperative information is beneficial to localization, we will consider varying the subset of nodes that broadcast and update their location beliefs at each iteration. We distinguish between these subsets by the *level of cooperative information* they induce in the algorithm. The level of cooperative information indicates how each node utilizes information from its neighbors at each iteration.

We introduce the following terminology: a distribution is said to be "sufficiently informative" when 95% of the probability mass is located within 2 m of the mean; a node becomes a *virtual anchor* when its belief is sufficiently informative; a *virtual bianchor* is a node with a bimodal belief, with each mode being sufficiently informative; a node that is neither a virtual anchor nor a virtual bianchor will be called a *blind agent*. We are now ready to introduce four levels of cooperative information at each iteration.

(i) *Level 1 (L1)*: virtual anchors broadcast their beliefs, while all other nodes censor their belief broadcast. Virtual anchors do not update their beliefs.

(ii) *Level 2 (L2)*: virtual anchors and virtual bianchors broadcast their beliefs, while blind nodes censor their belief broadcast. Virtual anchors do not update their beliefs.

(iii) *Level 3 (L3)*: all nodes broadcast their beliefs. Virtual anchors do not update their beliefs.

(iv) *Level 4 (L4)*: all nodes broadcast their beliefs. All nodes update their beliefs.

In terms of cooperation, note that L4 utilizes more cooperative information than L3, L3 utilizes more cooperative information than L2, and L2 utilizes more cooperative
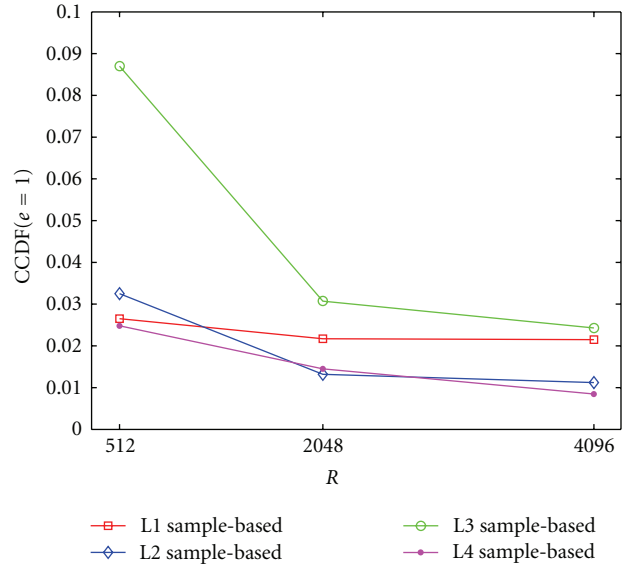

FIGURE 1: The effect of R and the level of cooperative information on the CCDF at $e = 1$ meter. All results are after $N_{it} = 10$ iterations.

information than L1. In this sense, the levels of cooperative information are strict subsets.

From previous sections, we know that the algorithm complexity scales linearly in M, the number of incoming messages in the multiplication operation. Hence, the level of cooperative information directly affects the algorithm's computational cost, with lower levels requiring less computation.

5.3. Numerical Results. We now examine how localization performance varies with the algorithm parameters. In particular, numerical results show the effect of message representation (sample-based or parametric) and level of cooperative information (L1, L2, L3, or L4) on the CCDF of the localization error.

We first consider the localization performance as a function of the number of samples R and level of cooperative information. Figure 1 displays the CCDF at $e = 1$ m after 10 iterations. As expected, for any level of cooperative information, the CCDF decreases as the number of samples is increased. However, the decrease in CCDF comes with a cost in computation time; as R is increased, the per-node complexity increases quadratically. Figure 1 also shows that levels L1, L2, and L4 are not as sensitive to R as L3 and that each generally outperforms L3. This effect is particularly pronounced when R is small. L3 broadcasts more complex distributions than L2 and L1, and these elaborate distributions are not accurately represented with a small number of samples.

Secondly, we investigate level of cooperative information and its effect on localization performance, with numerical results represented in Figures 2 and 3, after $N_{it} = 20$ iterations. Note that each curve exhibits a "floor" because there is always some subset of nodes that have insufficient information to localize without ambiguity. This may be due to lack of connectivity or large flip ambiguities. Let us
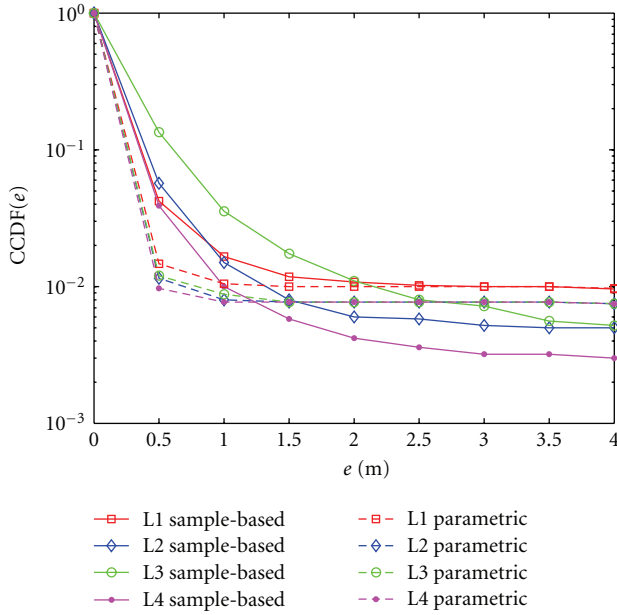
FIGURE 2: Effect of the level of cooperative information on local-ization accuracy for the distributed sample-based and parametric algorithms, when $N_{it} = 20$, and averaged over 50 networks.
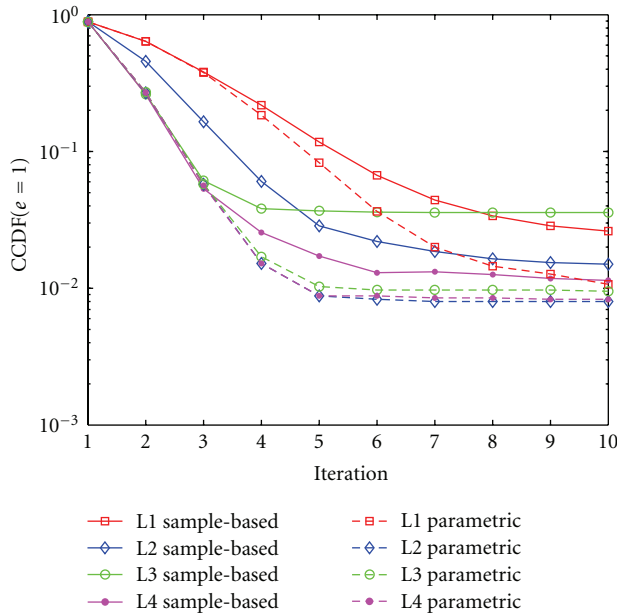

FIGURE 3: Effect of the level of cooperative information on the convergence speed for $e = 1$ m, for the distributed sample-based and parametric algorithms. Results are averaged over 50 networks.

focus on the sample-based representation in Figure 2 and consider the effect of the level of cooperative information on localization performance. In general, L4 has the best performance in terms of accuracy and floor. Intuitively, one might expect L3 to have the next best performance, followed by L2, and then L1. However, Figure 2 demonstrates that in some cases L3 has poorer accuracy than L2 and a similar

floor. This effect can be explained as follows. Agents that do not become virtual anchors within $N_{it} = 20$ tend to have large localization errors, creating a floor. Such agents comprise 1.7% of the total nodes for L1 and 0.3% for both L2 and L3. Since L2 and L3 have a similar fraction of agents that do not become virtual anchors, they have similar floors. In addition, the accuracy of beliefs belonging to agents that have become virtual anchors turns out to be highest for L1, followed by L2, and then L3. This is because L3 uses less reliable information than L2, which in turn is less reliable than L1. The final CCDF depends both on the fraction of virtual anchors (lowest for L1) and the accuracy of those virtual anchors (highest for L1). Note that we cannot compare L4 in this context, since there is no concept of a virtual anchor in L4.

We now move on to the parametric representation, still in Figure 2. We observe that L4 has the lowest overall CCDF for any e, for both types of message representation. For the parametric messages, the differences among different levels of cooperative information are smaller, and we generally obtain better performance (for $e < 1$ m) compared to sample-based messages.

Finally, in Figure 3, we evaluate the convergence speed of the different message representations and levels of coopera-tion, for a fixed error of 1 meter. We see that the parametric messages generally lead to faster convergence and lower CCDF than their sample-based counterparts. Levels L2, L3, and L4 all converge in around 5 iterations with a final CCDF at $e = 1$ m of around 0.01 for the parametric representation. Our results show that more cooperative information leads to faster improvement in terms of accuracy. The lowest level of cooperative information, L1, is consistently slower to converge and less accurate. However, higher levels of cooperative information also require the computation and representation of more complicated distributions. As a possible consequence, convergence issues may occur for levels L3 and L4. We also see that the parametric message representation performs approximately equal to or better than the sample-based messages in terms of both conver-gence and accuracy, while requiring much less execution time. Overall, parametric message representations yield a better performance/complexity tradeoff. This is due to the fact that the parametric distributions are well tailored to the localization problem and the homogeneous simulation environment.

6. Conclusions and Extensions

In this paper, we considered different message represen-tations for Bayesian cooperative localization in wireless networks: a generic sample-based representation and a tailored parametric representation. We used experimentally derived UWB ranging models to evaluate the performance of SPAWN as a function of message representation and level of cooperative information. Our results show that the tradeoffs between message representation, cooperative information, localization accuracy, and algorithm convergence are not straightforward and should be tailored to the scenario.

Through large-scale network simulations, we demonstrated that more cooperative information may improve localization accuracy but also increase the complexity of messages. Higher levels of cooperative information do not always correspond to an improvement in localization accuracy or convergence rate. As complicated distributions associated with location-uncertain nodes are computed and transmitted, the resulting increases in computational complexity and signal interference can actually reduce localization performance. It may therefore be advantageous to broadcast only confident information in cooperative localization networks, especially considering the resources saved by a node censorship policy.

We also demonstrated that though parametric messages have less representational flexibility, they can outperform nonparametric message representation at a much lower computational cost. In our simulations, the parametric representation achieved a lower probability of outage for errors under 1 meter while converging in equal or fewer iterations than the sample-based representation. Clearly, a parametric representation well tailored to the localization scenario is desirable in terms of both resource efficiency and localization accuracy.

The use of parametric distributions for localization can be extended to (i) different ranging models; (ii) different types of measurements; (iii) more general scenarios. In terms of ranging models, the proposed distributions can be applied as long as typical distributions in SPAWN roughly resemble a distribution in the $\mathcal{D}_2$-family. Note that a Gaussian ranging error satisfies this criteria, as would many other, more realistic, models. Other models, such as those derived from received signal strength, will require different types of parametric distributions. The same comment applies to the use of different types of measurements. For instance, with angle-of-arrival measurements, the parametric distributions should include a collection of linear distributions. Finally, more general scenarios may require tailor-made distributions. With NLOS measurements that can be modeled as biased Gaussians [20], for example, mixtures of $\mathcal{D}_2$ distributions would easily accommodate LOS/NLOS propagation, without relying on explicit NLOS identification.

References

[1] J. J. Caffery and G. L. Stüber, "Overview of radiolocation in CDMA cellular systems," *IEEE Communications Magazine*, vol. 36, no. 4, pp. 38–45, 1998.

[2] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 24–40, 2005.

[3] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications (WSNA '02)*, pp. 88–97, ACM Press, September 2002.

[4] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom '03)*, pp. 81–95, September 2003.

[5] K. Pahlavan, X. Li, and J. P. Mäkelä, "Indoor geolocation science and technology," *IEEE Communications Magazine*, vol. 40, no. 2, pp. 112–118, 2002.

[6] S. J. Ingram, D. Harmer, and M. Quinlan, "Ultrawideband indoor positioning systems and their use in emergencies," in *Proceedings of the Position Location and Navigation Symposium (PLANS '04)*, pp. 706–715, April 2004.

[7] C.-Y. Chong and S. P. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003.

[8] H. Yang and B. Sikdar, "A protocol for tracking mobile targets using sensor networks," in *Proceedings of the 1st IEEE International Workshop on Sensor Network Protocols and Applications*, pp. 71–81, May 2003.

[9] X. Ji and H. Zha, "Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling," in *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '04)*, vol. 4, pp. 2652–2661, March 2004.

[10] R. A. Marjamaa, P. M. Torkki, M. I. Torkki, and O. A. Kirvelä, "Time accuracy of a radio frequency identification patient tracking system for recording operating room timestamps," *Anesthesia & Analgesia*, vol. 102, no. 4, pp. 1183–1186, 2006.

[11] D. Fox, W. Burgard, F. Dellaert, and S. Thrun, "Monte Carlo localization: efficient position estimation for mobile robots," in *Proceedings of the 16th National Conference on Artificial Intelligence (AAAI '99)*, pp. 343–349, Orlando, Fla, USA, July 1999.

[12] J. J. Leonard and H. F. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons," *IEEE Transactions on Robotics and Automation*, vol. 7, no. 3, pp. 376–382, 1991.

[13] R. Jain, A. Puri, and R. Sengupta, "Geographical routing using partial information for wireless ad hoc networks," *IEEE Personal Communications*, vol. 8, no. 1, pp. 48–57, 2001.

[14] H. Frey, "Scalable geographic routing algorithms for wireless ad hoc networks," *IEEE Network*, vol. 18, no. 4, pp. 18–22, 2004.

[15] R. J. Fontana and S. J. Gunderson, "Ultra-wideband precision asset location system," in *Proceedings of IEEE Conference on Ultra Wideband Systems and Technologies (UWBST '02)*, vol. 21, no. 1, pp. 147–150, Baltimore, Md, USA, May 2002.

[16] W. C. Chung and D. Ha, "An accurate ultra wideband (UWB) ranging for precision asset location," in *Proceedings of IEEE Conference on Ultra Wideband Systems and Technologies (UWBST '03)*, pp. 389–393, November 2003.

[17] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero III, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 54–69, 2005.

[18] M. Castillo-Effen, W. A. Moreno, M. A. Labrador, and K. P. Valavanis, "Adapting sequential Monte-Carlo estimation to cooperative localization in wireless sensor networks," in *Proceedings of IEEE International Conference on Mobile Ad Hoc and Sensor Sysetems (MASS '06)*, pp. 656–661, Vancouver, Canada, October 2006.

[19] A. T. Ihler, J. W. Fisher III, R. L. Moses, and A. S. Willsky, "Nonparametric belief propagation for self-localization of sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 809–819, 2005.

[20] H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 427–450, 2009.

[21] A. T. Ihler, J. W. Fisher, and A. S. Willsky, "Particle filtering under communications constraints," in *Proceedings of the 13th IEEE/SP Workshop on Statistical Signal Processing*, pp. 89–94, Bordeaux, France, July 2005.

[22] M. Welling and J. J. Lim, "A distributed message passing algorithm for sensor localization," in *Proceedings of the 17th International Conference on Artificial Neural Networks (ICANN '07)*, pp. 767–775, September 2007.

[23] C. Pedersen, T. Pedersen, and B. H. Fleury, "A variational message passing algorithm for sensor self-localization in wireless networks," in *Proceedings of IEEE International Symposium on Information Theory Proceedings (ISIT '11)*, pp. 2158–2162, Saint-Petersburg, Russia, July 2011.

[24] M. Caceres, F. Penna, H. Wymeersch, and R. Garello, "Hybrid cooperative positioning based on distributed belief propagation," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 10, pp. 1948–1958, 2011.

[25] M. Z. Win and R. A. Scholtz, "Characterization of ultrawide bandwidth wireless indoor channels: a communication-theoretic view," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1613–1627, 2002.

[26] M. Z. Win and R. A. Scholtz, "On the robustness of ultra-wide bandwidth signals in dense multipath environments," *IEEE Communications Letters*, vol. 2, no. 2, pp. 51–53, 1998.

[27] A. F. Molisch, J. R. Foerster, and M. Pendergrass, "Channel models for ultrawideband personal area networks," *IEEE Wireless Communications*, vol. 10, no. 6, pp. 14–21, 2003.

[28] J.-Y. Lee and R. A. Scholtz, "Ranging in a dense multipath environment using an UWB radio link," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1677–1683, 2002.

[29] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win, "Ranging with ultrawide bandwidth signals in multipath environments," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 404–425, 2009.

[30] W. Suwansantisuk and M. Z. Win, "Multipath aided rapid acquisition: optimal search strategies," *IEEE Transactions on Information Theory*, vol. 53, no. 1, pp. 174–193, 2007.

[31] H. Xu and L. Yang, "Timing with dirty templates for low-resolution digital UWB receivers," *IEEE Transactions on Wireless Communications*, vol. 7, no. 1, pp. 54–59, 2008.

[32] M. Z. Win, P. C. Pinto, and L. A. Shepp, "A mathematical theory of network interference and its applications," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 205–230, 2009.

[33] N. C. Beaulieu and D. J. Young, "Designing time-hopping ultrawide bandwidth receivers for multiuser interference environments," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 255–284, 2009.

[34] M. Z. Win, G. Chrisikos, and A. F. Molisch, "Wideband diversity in multipath channels with nonuniform power dispersion profiles," *IEEE Transactions on Wireless Communications*, vol. 5, no. 5, pp. 1014–1022, 2006.

[35] M. Z. Win, G. Chrisikos, and N. R. Sollenberger, "Performance of Rake reception in dense multipath channels: implications of spreading bandwidth and selection diversity order," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 8, pp. 1516–1525, 2000.

[36] C. Falsi, D. Dardari, L. Mucchi, and M. Z. Win, "Time of arrival estimation for UWB localizers in realistic environments," *EURASIP Journal on Advances in Signal Processing*, vol. 2006, Article ID 32082, pp. 1–13, 2006.

[37] S. Venkatesh and R. M. Buehrer, "Non-line-of-sight identification in ultra-wideband systems based on received signal statistics," *IET Microwaves, Antennas & Propagation*, vol. 1, no. 6, pp. 1120–1130, 2007.

[38] D. B. Jourdan, J. J. Deyst Jr., M. Z. Win, and N. Roy, "Monte Carlo localization in dense multipath environments using UWB ranging," in *Proceedings of IEEE International Conference on Ultra-Wideband (ICU '05)*, pp. 314–319, September 2005.

[39] S. Venkatesh and R. M. Buehrer, "Multiple-access design for ad hoc UWB position-location networks," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC '06)*, vol. 4, pp. 1866–1873, Las Vegas, Nev, USA, April 2006.

[40] D. Dardari, A. Conti, J. Lien, and M. Z. Win, "The effect of cooperation on UWB-based positioning systems using experimental data," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, Article ID 513873, 2008.

[41] S. Venkatesh and R. M. Buehrer, "Power control in UWB position-location networks," in *Proceedings of IEEE International Conference on Communications (ICC '06)*, vol. 9, pp. 3953–3959, Istanbul, Turkey, June 2006.

[42] W. C. Headley, C.R.C.M. da Suva, and R. M. Buehrer, "Indoor location positioning of non-active objects using ultra-wideband radios," in *Proceedings of IEEE Radio and Wireless Symposium (RWS '07)*, pp. 105–108, Long Beach, Calif, USA, January 2007.

[43] D. B. Jourdan, D. Dardari, and M. Z. Win, "Position error bound for UWB localization in dense cluttered environments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 2, pp. 613–628, 2008.

[44] Y. Shen and M. Z. Win, "Fundamental limits of wideband localization—part I: a general framework," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4956–4980, 2010.

[45] Y. Shen, H. Wymeersch, and M. Z. Win, "Fundamental limits of wideband localization—part II: cooperative networks," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4981–5000, 2010.

[46] Y. Shen and M. Z. Win, "On the accuracy of localization systems using wideband antenna arrays," *IEEE Transactions on Communications*, vol. 58, no. 1, pp. 270–280, 2010.

[47] S. Gezici, Z. Tian, G. B. Giannakis et al., "Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, 2005.

[48] I. Oppermann, M. Hämäläinen, and J. Iinatti, *UWB Theory and Applications*, John Wiley & Sons, 2004.

[49] H. Wymeersch, *Iterative Receiver Design*, Cambridge University Press, 2007.

[50] D. Fox, W. Burgard, H. Kruppa, and S. Thrun, "A Monte Carlo algorithm for multi-robot localization," Tech. Rep. CMU-CS-99-120, Computer Science Department, Carnegie Mellon University, Pittsburgh, Pa, USA, 1999.

[51] D. MacKay, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, 2003.

[52] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[53] Z. Botev, *Nonparametric Density Estimation Via Diffusion Mixing*, Postgraduate Series, The University of Queensland, 2007, http://espace.library.uq.edu.au/view/UQ:120006.

[54] J. Lien, *A framework for cooperative localization in ultra-wideband wireless networks [M.S. thesis]*, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Mass, USA, 2007.