A comparison of Bayesian localization methods in the presence of outliers

Citation for the original published paper (version of record):
Ferrara, N., Wymeersch, H., Lohan, E. et al (2017). A comparison of Bayesian localization methods in the presence of outliers. 2017 13th International Wireless Communications and Mobile Computing Conference, IWCMC 2017: 1546-1551. http://dx.doi.org/10.1109/IWCMC.2017.7986514

(article starts on next page)

A Comparison of Bayesian Localization Methods in the Presence of Outliers

Nunzia Giorgia Ferrara*, Henk Wymeersch†, Elena Simona Lohan* and Jari Nurmi*
*Department of Electronics and Communications Engineering, Tampere University of Technology, Tampere, Finland
†Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden
Email: {nunzia.ferrara, elena-simona.lohan, jari.nurmi}@tut.fi, henkw@chalmers.se

***Abstract*—Localization of a user in a wireless network is challenging in the presence of malfunctioning or malicious reference nodes, since if they are not accounted for, large localization errors can ensue. We evaluate three Bayesian methods to statistically identify outliers during localization: an exact method, an expectation maximization (EM) method proposed earlier, and a new method based on Variational Bayesian EM (VBEM). Simulation results indicate similar performance for the latter two schemes, with the VBEM algorithm able to provide a statistical description of the user location, rather than an estimate as in the simpler EM case. In contrast to previous studies, we find that there is a significant gap between the approximate methods and the exact method, the cause of which is discussed.**

***Index Terms*—localization, outliers, Bayesian methods, wireless networks.**

I. Introduction

Localization of nodes in a wireless network has many applications, ranging from personal navigation, over military tracking and monitoring, to automated driving. Localization can be in either an absolute or relative frame of reference. In absolute localization, users (referred to as agents) collect measurements with respect to fixed reference nodes (or anchors). These anchors have a known location, so that measurements in the form of distance or angle can provide absolute location information to the agent.

Under normal operating conditions, anchors would provide reliable information, both in terms of their own location, but also in terms of the measurements. However, in certain conditions, anchors may be compromised or malfunctioning, leading them to provide incorrect location or measurement information. If not accounted for properly, these anchors can lead the agent to compute an erroneous location, with an ensuing impact on the higher-level application. For this reason, outlier detection in wireless networks has received great attention, especially for safety-critical applications.

In order to mitigate the effect of malfunctioning/malicious anchors, several methods to detect them have been proposed [1]–[8]. In particular, [1] and [2] consider generic outlier detection for wireless sensor networks and provide a taxonomy of different approaches. In the context of localization, [3] applies convex optimization followed by a geometric approach to identify malfunctioning anchors. As an example of attack-resistant localization, [4] resorts to a cooperative approach, using optimization with dedicated Huber loss function. Using linear equations to describe the localization problem with an l_1 norm, linear programming can be used to avoid outliers measurements in the agent location estimation [5]. Alternatively, since the effect of outliers can be treated as errors in the transmitted data, error correcting codes have also been proposed [6], where iteratively, the region of interest is split into regions and a hypothesis test is performed at the fusion center, which receives data from each sensor. In contrast to the above optimization approaches, Bayesian methods were considered in [7], [8]: [7] considered message passing algorithm based on probabilistic graphical models for identifying outliers, while [8] applied the EM algorithm to perform joint localization and outlier detection and reported similar performance to [7] with reduced complexity. However, the EM method does not provide a distribution of the agent location, only an estimate.

In this paper, we extend [8] to a VBEM method, which provides a statistical description of the agent location, rather than only an estimate as in the simpler EM case. Since, through the VBEM approach, a complete distribution is obtained, an estimate of the user location but also information about the uncertainty in such estimate are provided. We also make an in-depth comparison between three methods: an exact method, an EM method, and a method based on VBEM. Each method is evaluated in terms of the localization performance and also in terms of the ability to identify outliers. In contrast to [8], we find a significant performance gap between the exact method and the two approximate methods. Our analysis provides insight into the underlying reason for this gap as well as suggestions on how it may be closed.

II. System model

We consider a wireless network with a single agent and N anchors deployed over region $\mathcal{R} \subseteq \mathbb{R}^2$. Let $\boldsymbol{\theta} \in \mathcal{R}$ be the agent position with Gaussian prior $p(\boldsymbol{\theta})$, with mean $\boldsymbol{\mu}_p$ and covariance matrix $\boldsymbol{\Sigma}_p$, and $\boldsymbol{a}_n \in \mathcal{R}$ be the n-th anchor position in the two-dimensional space. The agent localizes itself by estimating the distance $\|\boldsymbol{a}_n - \boldsymbol{\theta}\|$ with respect to every anchor. Different types of measurement techniques can be employed, including distance estimation from received signal strength, time-of-arrival, or angle-of-arrival. We denote by ρ_n the distance estimated by the agent with respect to anchor n. Each anchor is in one of the two states $s_n \in \{0, 1\}$, where $s_n = 0$ represents the state of a well-functioning anchor, whereas $s_n = 1$ means that the anchor is not functioning properly.

The state s_n is assumed to have a Bernoulli distribution, and it is independent across anchors:

$$p(s_n) = \delta^{s_n}(1-\delta)^{1-s_n}. \tag{1}$$

where δ is the probability of an anchor to be malfunctioning. The ranging error model depends on the state of the anchors. In particular, we assume

$$p(\rho_n|s_n = 0;\boldsymbol{\theta}) = f(\rho_n, \boldsymbol{\theta}) \triangleq \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\big(-\frac{(\rho_n - \|\boldsymbol{a}_n - \boldsymbol{\theta}\|)^2}{2\sigma_n^2}\big), \tag{2}$$

while $p(\rho_n|s_n = 1;\boldsymbol{\theta}) = c$, where c is an appropriate constant (for instance $c = 1/R_{\max}$, where $R_{\max}$ is the maximum value of ρ_n). In other words, for well-functioning anchors, the distance estimation error has a zero-mean Gaussian distribution, while, for malfunctioning anchors, distance estimates can fall with equal probability anywhere within a certain range.

After collecting the distance estimates $\boldsymbol{\rho} = [\rho_1, \rho_2, \ldots, \rho_N]^T$ and ranging qualities $\boldsymbol{\sigma} = [\sigma_1, \sigma_2, \ldots, \sigma_N]^T$ from the anchors, the agent has a goal to determine its posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{\rho})$.

We will assume a positioning method is available that can compute $p(\boldsymbol{\theta}|\boldsymbol{\rho})$ when there are no malfunctioning anchors. Hence, such a method is able to efficiently compute

$$p(\boldsymbol{\theta}|\boldsymbol{\rho}) = p(\boldsymbol{\theta}) \prod_{n=1}^{N} f(\rho_n, \boldsymbol{\theta}) \tag{3}$$

in the form of a suitable approximation, e.g., a Gaussian distribution.

III. Three Bayesian Methods

In this section, we present three Bayesian methods for localizing the agent: an exact method, an EM method, and a VBEM method.

A. Exact Method

Introducing $\mathbf{s} = [s_1, \ldots, s_N]^T$, we can determine $p(\boldsymbol{\theta}|\boldsymbol{\rho})$ by taking the expectation over the state of each anchor. This leads to

$$\begin{aligned} p(\boldsymbol{\theta}|\boldsymbol{\rho}) &= \sum_{\mathbf{s}\in\{0,1\}^N} p(\boldsymbol{\theta}, \mathbf{s}|\boldsymbol{\rho}) \\ &\propto \sum_{\mathbf{s}\in\{0,1\}^N} p(\boldsymbol{\rho}|\boldsymbol{\theta}, \mathbf{s}) p(\boldsymbol{\theta}) p(\mathbf{s}) \\ &= p(\boldsymbol{\theta}) \sum_{\mathbf{s}\in\{0,1\}^N} \prod_{n=1}^{N} p(\rho_n|\boldsymbol{\theta}, s_n) p(s_n) \\ &= p(\boldsymbol{\theta}) \prod_{n=1}^{N} \sum_{s_n\in\{0,1\}} p(\rho_n|\boldsymbol{\theta}, s_n) p(s_n) \\ &= p(\boldsymbol{\theta}) \prod_{n=1}^{N} (\delta c + (1-\delta) f(\rho_n, \boldsymbol{\theta})), \end{aligned} \tag{4}$$

where we have used the a priori independence of the anchors' states and the agent's location, as well as the mutual independence of the anchors' states.

Remark 1. *Expression* (4) *is not part of the standard form. On inspection, we see that $p(\boldsymbol{\theta}|\boldsymbol{\rho})$ can be expressed as mixture of 2^N components of the form* (3)*. Hence, it is hard to evaluate for large N.*

B. Expectation Maximization Method

The EM method is an iterative method for determining the MAP estimate of $\boldsymbol{\theta}$ in the presence of a hidden variable, under the assumption that estimation would be easy if the hidden variable was known [9, Chapter 9]. In our case, $\mathbf{s}$ is the hidden variable. The EM algorithm starts from an initial estimate $\hat{\boldsymbol{\theta}}_0$ and produces estimates $\hat{\boldsymbol{\theta}}_t$ with $p(\hat{\boldsymbol{\theta}}_t|\boldsymbol{\rho}) \geq p(\hat{\boldsymbol{\theta}}_{t-1}|\boldsymbol{\rho})$.

During iteration t, in the E-step, we determine

$$\begin{aligned} q^{(t)}(\boldsymbol{\theta}) &\propto \exp\left(\sum_{\mathbf{s}\in\{0,1\}^N} q_{\mathbf{s}}^{(t)}(\mathbf{s}) \ln p(\boldsymbol{\theta}|\mathbf{s}, \boldsymbol{\rho})\right) \\ &\propto p(\boldsymbol{\theta}) \exp\left(\sum_{\mathbf{s}\in\{0,1\}^N} q_{\mathbf{s}}^{(t)}(\mathbf{s}) \ln p(\boldsymbol{\rho}|\boldsymbol{\theta}, \mathbf{s})\right) \\ &= p(\boldsymbol{\theta}) \prod_{n=1}^{N} \exp\left(\sum_{s_n\in\{0,1\}} q_{s_n}^{(t)}(s_n) \ln p(\rho_n|\boldsymbol{\theta}, s_n)\right) \\ &= p(\boldsymbol{\theta}) \prod_{n=1}^{N} \exp\left(q_{s_n}^{(t)}(0) \ln f(\rho_n, \boldsymbol{\theta}) + q_{s_n}^{(t)}(1) \ln c\right) \\ &\propto p(\boldsymbol{\theta}) \prod_{n=1}^{N} f(\rho_n, \boldsymbol{\theta})^{q_{s_n}^{(t)}(0)}, \end{aligned} \tag{5}$$

where $q_{\mathbf{s}}^{(t)}(\mathbf{s}) = p(\mathbf{s}|\hat{\boldsymbol{\theta}}_{t-1}, \boldsymbol{\rho})$ and $q_{s_n}^{(t)}(s_n)$ is the marginal distribution of $q_{\mathbf{s}}^{(t)}(\mathbf{s})$ for s_n. These are related by

$$\begin{aligned} q_{\mathbf{s}}^{(t)}(\mathbf{s}) &= p(\mathbf{s}|\hat{\boldsymbol{\theta}}_{t-1}, \boldsymbol{\rho}) \\ &\propto p(\mathbf{s}) p(\boldsymbol{\rho}|\hat{\boldsymbol{\theta}}_{t-1}, \mathbf{s}) \\ &= \prod_{n=1}^{N} \underbrace{p(s_n) p(\rho_n|\hat{\boldsymbol{\theta}}_{t-1}, s_n)}_{\propto q_{s_n}^{(t)}(s_n)}. \end{aligned} \tag{6}$$

In the M-step, we maximize $q^{(t)}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}}_t = \arg\max_{\boldsymbol{\theta}} q^{(t)}(\boldsymbol{\theta}). \tag{7}$$

Remark 2. *We notice that in* (5)*, we can express*

$$f(\rho_n, \boldsymbol{\theta})^{q_{s_n}^{(t)}(0)} \propto \exp\big(-\frac{(\rho_n - \|\boldsymbol{a}_n - \boldsymbol{\theta}\|)^2}{2\sigma_n^2/q_{s_n}^{(t)}(0)}\big). \tag{8}$$

In other words, the ranging variance is increased by a factor $1/q_{s_n}^{(t)}(0)$. This implies that $q^{(t)}(\boldsymbol{\theta})$ can be computed efficiently using (3)*. The EM method can thus be easily implemented, but fails to provide a posterior distribution of the user location.*

Remark 3. *The EM algorithm will consider an anchor more likely to be malfunctioning when* $q_{s_n}^{(t)}(0) < q_{s_n}^{(t)}(1)$. *It is readily verified that this happens when*

$$\frac{\left(\rho_n - \|\mathbf{a}_n - \hat{\boldsymbol{\theta}}_t\|\right)^2}{2\sigma^2} > \log\left(\frac{1-\delta}{\delta c\sqrt{2\pi\sigma^2}}\right). \tag{9}$$

Hence, in order to have reasonable positioning performance, the estimate $\hat{\boldsymbol{\theta}}_t$ *should be such that for at least 3 good anchors, the condition* (9) *is satisfied.*

C. Variational Bayesian EM Method

The VBEM method [9, Chapter 10] aims to approximate $p(\boldsymbol{\theta}, \mathrm{s}|\boldsymbol{\rho})$ by a distribution of the form $q_{\boldsymbol{\theta}}(\boldsymbol{\theta})q_{\mathbf{s}}(\mathrm{s})$ in such a way as to minimize the Kullback-Leibler divergence [10, Chapter 2]:

$$\begin{aligned} &D(q_{\boldsymbol{\theta}}(\boldsymbol{\theta})q_{\mathbf{s}}(\mathrm{s})\|p(\boldsymbol{\theta}, \mathrm{s}|\boldsymbol{\rho})) \\ &= \sum_{\mathbf{s}} \int q_{\boldsymbol{\theta}}(\boldsymbol{\theta})q_{\mathbf{s}}(\mathrm{s}) \log \frac{q_{\boldsymbol{\theta}}(\boldsymbol{\theta})q_{\mathbf{s}}(\mathrm{s})}{p(\boldsymbol{\theta}, \mathrm{s}|\boldsymbol{\rho})} \mathrm{d}\boldsymbol{\theta}, \end{aligned} \tag{10}$$

so that the optimal solution is given by solving

$$\begin{aligned} &[q_{\boldsymbol{\theta}}^*(\boldsymbol{\theta}), q_{\mathbf{s}}^*(\mathrm{s})] \\ &= \arg \min_{q_{\boldsymbol{\theta}}(\boldsymbol{\theta}), q_{\mathbf{s}}(\mathrm{s})} D(q_{\boldsymbol{\theta}}(\boldsymbol{\theta})q_{\mathbf{s}}(\mathrm{s})\|p(\boldsymbol{\theta}, \mathrm{s}|\boldsymbol{\rho})). \end{aligned} \tag{11}$$

The problem (11) can be solved iteratively, starting from an initial guess $q_{\boldsymbol{s}}^{(0)}(\boldsymbol{s})$. It can be verified that the solution takes the following fixed-point expression:

$$q_{\boldsymbol{\theta}}^{(t)}(\boldsymbol{\theta}) \propto p(\boldsymbol{\theta}) \prod_{n=1}^{N} (f(\rho_n, \boldsymbol{\theta}))^{q_{s_n}^{(t-1)}(0)}, \tag{12}$$

where $q_{\mathbf{s}}^{(t)}(\mathrm{s}) = \prod_{n=1}^{N} q_{s_n}^{(t)}(s_n)$, and

$$q_{s_n}^{(t)}(s_n) \propto p(s_n) \exp\left[\int \ln p(\rho_n|s_n, \boldsymbol{\theta}) q_{\boldsymbol{\theta}}^{(t)}(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta}\right]. \tag{13}$$

Note that (13) evaluates to

$$\begin{aligned} &q_{s_n}^{(t)}(s_n) \\ &= \begin{cases} (1-\delta) \exp\left[\int \ln f(\rho_n, \boldsymbol{\theta}) q_{\boldsymbol{\theta}}^{(t-1)}(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta}\right] & s_n = 0 \\ \delta c & s_n = 1. \end{cases} \end{aligned} \tag{14}$$

which can be solved efficiently using Monte Carlo integration.

Remark 4. *The VBEM can rely on the standard operation* (3) *to compute* $q_{\boldsymbol{\theta}}^{(t)}(\boldsymbol{\theta})$. *The update rules shown in* (13)–(12) *converge to a local minimum of* $D(q_{\boldsymbol{\theta}}(\boldsymbol{\theta})q_{\mathbf{s}}(\mathrm{s})\|p(\boldsymbol{\theta}, \mathrm{s}|\boldsymbol{\rho}))$. *The VBEM algorithm is a variational generalization of the EM algorithm. The difference is that in the EM algorithm the maximization step computes a point estimate rather than a distribution over the parameter* $\boldsymbol{\theta}$. *If, at every step in the VBEM method,* $q_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ *is restricted to a Dirac delta function, the algorithm reduces to EM.*

Remark 5. *The VBEM algorithm will consider an anchor more likely to be malfunctioning when* $q_{s_n}^{(t)}(0) < q_{s_n}^{(t)}(1)$. *It is readily verified that this happens when*

$$\mathbb{E}\left\{\frac{(\rho_n - \|\mathbf{a}_n - \boldsymbol{\theta}\|)^2}{2\sigma^2}\right\} > \log\left(\frac{1-\delta}{\delta c\sqrt{2\pi\sigma^2}}\right), \tag{15}$$

in which the expectation is with respect to $q_{\boldsymbol{\theta}}^{(t)}(\boldsymbol{\theta})$. *In other words, when* $q_{\boldsymbol{\theta}}^{(t)}(\boldsymbol{\theta})$ *is such that, for good anchors, the expected ranging error is above a certain threshold, the anchor will (incorrectly) be considered to be malfunctioning. In particular, a very broad* $q_{\boldsymbol{\theta}}^{(t)}(\boldsymbol{\theta})$ *will cause the VBEM algorithm to consider all anchors to be malfunctioning.*

IV. Performance Evaluation

We have carried out Monte Carlo simulations to evaluate the performance of the three methods in terms of localization accuracy and ability to identify outliers.

A. Performance Metrics

The three algorithms eventually localize the agent as follows: for the exact method, the estimated user location $\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\boldsymbol{\rho})$, for the EM method $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}^{(t)}$ for a sufficiently large value of t (see later), while for the VBEM method, $\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} q_{\boldsymbol{\theta}}^{(t-1)}(\boldsymbol{\theta})$ for a sufficiently large value of t (see later). The localization accuracy is measured based on the deviation of the estimated location $\hat{\boldsymbol{\theta}}$ from the true location $\boldsymbol{\theta}$. The final performance metric is the RMSE, defined as

$$\mathrm{RMSE} = \sqrt{\mathbb{E}\{\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|^2\}}, \tag{16}$$

where the expectation is over the random simulation parameters (anchor locations, agent locations, noise realizations).

The outlier identification capability is measured in terms of the probability of incorrectly identifying an outlier. To this end, the performance metric is the error probability, defined as

$$P_{\mathrm{err}} = \delta P_{\mathrm{md}} + (1-\delta) P_{\mathrm{fa}}, \tag{17}$$

in which P_{md} and P_{fa} denote the missed detection and false alarm probabilities, respectively, based on the MAP decisions of s, i.e., $\hat{\mathrm{s}} = [\hat{s}_1, \ldots, \hat{s}_N]^T$ in which

$$\hat{s}_n = \arg \max_{s_n \in \{0,1\}} p(s_n|\boldsymbol{\rho}), \tag{18}$$

for the exact method and

$$\hat{s}_n = \arg \max_{s_n \in \{0,1\}} q_{s_n}^{(t)}(s_n), \tag{19}$$

for the EM and VBEM methods.

We note that (18) can be computed as follows:

$$\begin{aligned} &p(s_n|\boldsymbol{\rho}) \\ &\propto \begin{cases} (1-\delta) \int \frac{f(\rho_n, \boldsymbol{\theta}) p(\boldsymbol{\theta}|\boldsymbol{\rho})}{(\delta c + (1-\delta) f(\rho_n, \boldsymbol{\theta}))} \mathrm{d}\boldsymbol{\theta} & s_n = 0 \\ \delta c \int \frac{p(\boldsymbol{\theta}|\boldsymbol{\rho})}{(\delta c + (1-\delta) f(\rho_n, \boldsymbol{\theta}))} \mathrm{d}\boldsymbol{\theta} & s_n = 1, \end{cases} \end{aligned} \tag{20}$$

in which the integral can be evaluated through standard Monte Carlo integration.

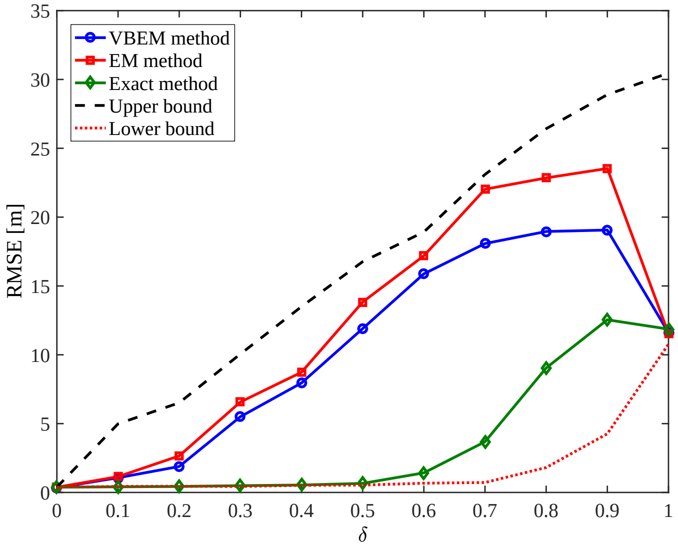

Fig. 1. Localization RMSE as a function of the malfunctioning probability δ.

B. Simulation setup

We consider a network with $N = 30$ anchors randomly distributed over a square region of side length $L = 50$ m. At every simulation, the single agent position is drawn from a prior Gaussian distribution with mean $\boldsymbol{\mu}_p = [0\ 0]^T$, corresponding to the center of the region, and standard deviation $\sigma_p = L/6$ m in both dimensions. The square region is discretized in a grid of resolution 0.1 m, in order to carry out numerical integration and thus enable a fair comparison of the three methods. The performance is evaluated for $\delta \in [0, 1]$. The measurements from well-functioning anchors are generated according to (2) with $\sigma_n = 1$, whereas the measurements from malfunctioning anchors are generated from the distribution $p(\rho_n | s_n = 1; \boldsymbol{\theta}) = c$, where c is set equal to $\frac{1}{\sqrt{2}L}$. For the EM method, we set $\hat{\boldsymbol{\theta}}_0 = \boldsymbol{\mu}_p$, while for the VBEM method, for all n, $q^{(0)}_{s_n}(s_n) = 0.5$, $s_n \in \{0, 1\}$. The number of iterations in EM and VBEM are adaptive, based on the differences[1] between successive iterations. We observed that, on average, the VBEM method requires a slightly higher number of iterations compared to the EM approach.

C. Localization performance

The results in terms of RMSE for the three methods are shown in Fig. 1. The plot also contains the curves for upper and lower bounds. The upper bound on the RMSE corresponds to the worst-case scenario in which all the anchors, even the malfunctioning ones, are considered to be well-functioning and, hence, are all given the same (unitary) weight when computing $q_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. The lower bound, instead, corresponds to the best-case scenario where all the malfunctioning nodes are discarded and only the well-functioning ones are used, with the same (unitary) weight, when computing $q_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. For the EM and VBEM methods, $q_{s_n}(s_n)$ determines the weight that the n-th anchor's measurement is given when computing respectively (5) and (12). The higher (closer to 1) $q_{s_n}(0)$ is, the more strongly the n-th anchor is deemed to be well-functioning, and therefore a higher weight is given to its measurement.

First of all, we notice a large gap between the bounds. The lower bound is approximately constant at around 0.5 m until $\delta = 0.7$, while the upper bound suffers from large degradations, starting already at $\delta = 0.1$. The exact method is able to achieve the lower bound until $\delta = 0.5$, after which it quickly degrades to up to around 12 m RMSE. The VBEM and EM methods achieve similar performance, much better than the upper bound for low values of δ, reducing the degradation that the malfunctioning anchors would cause if not accounted for, but still leading to considerable errors (e.g., 5 m for $\delta = 0.3$). Note that (9) is satisfied when $\rho_n - \|\mathbf{a}_n - \hat{\boldsymbol{\theta}}_t\|$ is greater than 3.17 m (for $\delta = 0.1$) down to 1.12 m (for $\delta = 0.9$). Hence, if the initial estimate is more than 3.17 m away from the true value, the EM algorithm will consider the corresponding anchor to be malfunctioning. This highlights the role of having a good initial estimate. A similar reasoning can be carried out for the VBEM algorithm where an initial $q^{(0)}_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ greatly affects the quality of the result. In all three algorithms, the increasing trend of the error as a function of δ is interrupted at $\delta = 0.9$. When all the reference nodes are malfunctioning ($\delta = 1$), (4) (for the exact method), (5) (for EM) and (12) (for VBEM) revert to the prior $p(\boldsymbol{\theta})$, so that in all cases, $\hat{\boldsymbol{\theta}} = \boldsymbol{\mu}_p$. We note that these results are in contrast to [8], which observed that EM had similar performance to the exact method. This is because of difference in initial estimates. In our case, both EM and VBEM are plagued with poor initial estimates, thus leading many anchors that are considered to be malfunctioning, and in turn to larger RMSE.

While the RMSE may indicate that EM and VBEM yield poor performance, a more nuanced view is offered in Fig. 2. We see that VBEM leads to low localization errors for $\delta < 0.3$ at least 99% of the time. Similar observations hold for EM. This implies that the RMSE results are dominated by a very small fraction of large errors. To have a deeper understanding, we now look to the error detection probability.

D. Outlier detection performance

Fig. 3 shows the error probability, averaged over all the N anchors, for the three analyzed methods. It can be observed that in all the three cases, the maximum value occurs for $\delta = 0.7$. The VBEM and EM methods show similar performance, with the former slightly outperforming the latter for $\delta > 0.5$. For δ below 0.2, EM and VBEM have a detection probability similar to the exact method, in contrast to their respective RMSE performance. To understand this, we note that the erroneous detection of an anchor state can be due either to a false alarm or to a missed detection. Fig. 4 shows a histogram of the false alarms and missed detections for each network realization (i.e., how many of the N anchors were incorrectly detected as malfunctioning and normal, respectively)

[1]The algorithms are assumed to have converged when $\max_n |q^{(t)}_{s_n}(0) - q^{(t-1)}_{s_n}(0)| \leq \epsilon$ and $\max_{\boldsymbol{\theta}} |q^{(t)}_{\boldsymbol{\theta}}(\boldsymbol{\theta}) - q^{(t-1)}_{\boldsymbol{\theta}}(\boldsymbol{\theta})| \leq \epsilon$, for VBEM, and $|\hat{\boldsymbol{\theta}}_t - \hat{\boldsymbol{\theta}}_{t-1}| \leq \epsilon$ for EM. In our simulations, we used $\epsilon = 10^{-3}$.

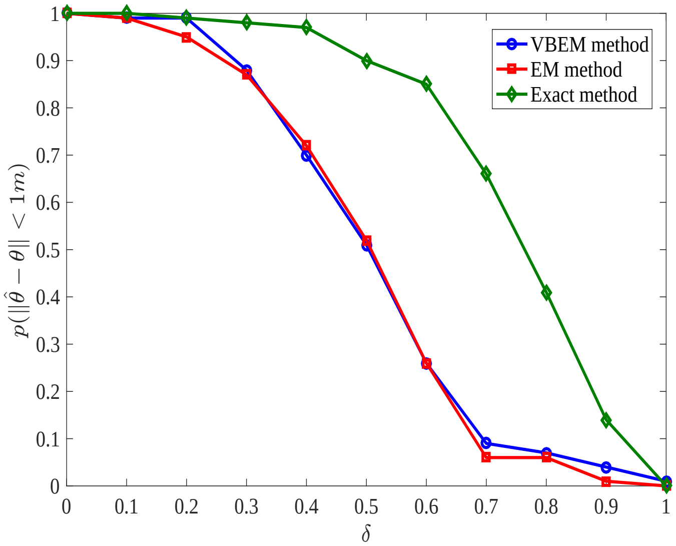

Fig. 2. Probability of having localization error less than 1 m as a function of the malfunctioning probability δ.

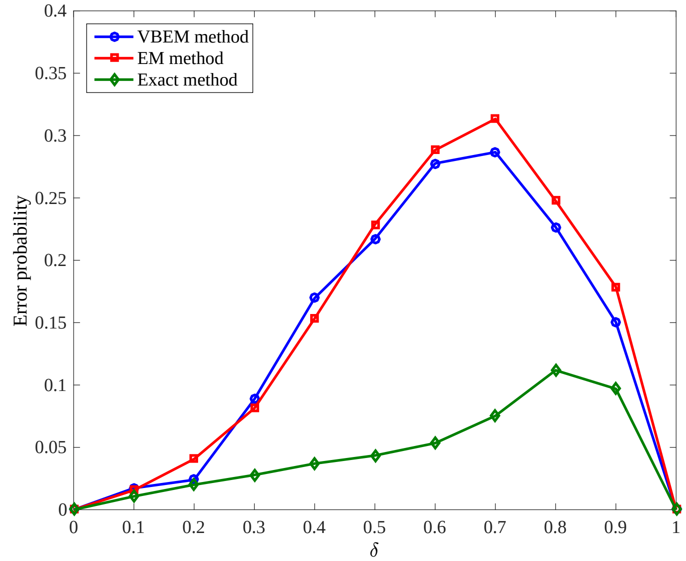

Fig. 3. Probability of incorrectly estimating the states of the anchors as a function of the malfunctioning probability δ.

for $\delta = 0.4$ for the VBEM method (similar observations hold for EM). We observe that most of the error events with a small number of errors (say, less than 5 errors) correspond to missed detections (i.e., malfunctioning anchors that are considered operating normally). Since they are few in number, their impact on the positioning is limited. In contrast, the large error events (10–20 errors) are almost exclusively false alarms (so that we discard many good anchors, of which there are on average only $(1 - \delta)N = 18$). This means that when the number of detected malfunctioning anchors far exceeds δN, the algorithm in fact knows that the initial estimates were poor and can try to re-run VBEM or EM with better initial estimates. Such a strategy was not explored in the current paper.

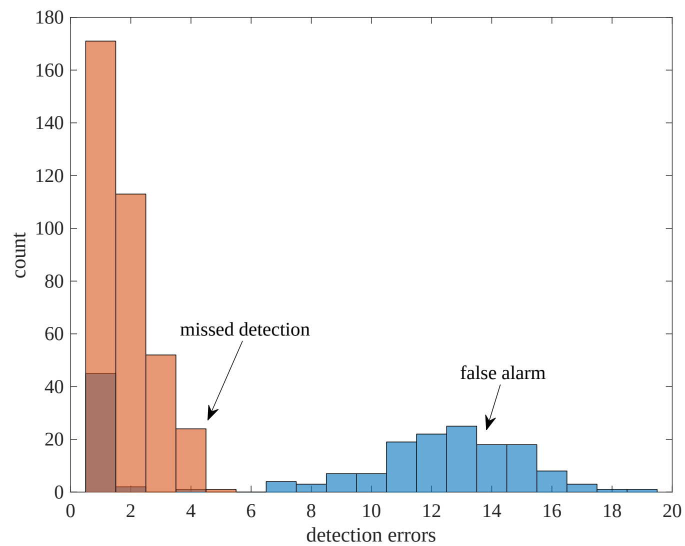

Fig. 4. Histogram of detection errors for VBEM when $\delta = 0.4$.

V. Conclusions

In this paper, we have evaluated three Bayesian methods to deal with the problem of user localization in wireless networks in the presence of malfunctioning anchors: an exact method, an EM method, and a VBEM method. Their performance in terms of localization accuracy and ability to identify outliers have been analyzed. In contrast to the literature, we observe that the two approximate methods, which show similar performance (with VBEM slightly outperforming EM), generally lead to poor performance compared to the exact method. This poor performance is mainly due to false alarms, which in turn are due to poor initialization of the EM and VBEM methods.

Acknowledgment

The EU FP7 Marie Curie Initial Training Network MULTI-POS (Multi-technology Positioning Professionals) supported this work financially under Grant Number 316528.

References

[1] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier detection techniques for wireless sensor networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 2, pp. 159–170, 2010.

[2] V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial intelligence review*, vol. 22, no. 2, pp. 85–126, 2004.

[3] S. Misra, G. Xue, and S. Bhardwaj, "Secure and robust localization in a wireless ad hoc environment," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 3, pp. 1480–1489, 2009.

[4] D. Wang, L. Yang, and X. Cheng, "A low-complexity cooperative algorithm for robust localization in wireless sensor networks," in *International Conference on Computing, Networking and Communications (ICNC)*, 2016.

[5] J. S. Picard and A. J. Weiss, "Bounds on the number of identifiable outliers in source localization by linear programming," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2884–2895, 2010.

[6] A. Vempaty, Y. S. Han, and P. K. Varshney, "Target localization in wire-

less sensor networks using error correcting codes," *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 697–712, 2014.

[7] Y. Li, D. Liu, and H. Wymeersch, "Bayesian outlier detection in location-aware wireless networks," in *Workshop on Positioning Navigation and Communication (WPNC)*, 2011, pp. 39–44.

[8] Y. Xiong, N. Wu, H. Wang, and J. Kuang, "Joint localization and cooperative detection in location-aware wireless networks in the presence of ranging outliers," in *International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2016.

[9] C. M. Bishop, *Pattern recognition and Machine Learning*. Springer, 2006.

[10] D. J. MacKay, *Information theory, inference and learning algorithms*. Cambridge university press, 2003.