VARIATIONAL INFERENCE COOPERATIVE NETWORK LOCALIZATION WITH NARROWBAND RADIOS

Nikos Fasarakis-Hilliard, Panos N. Alevizos, and Aggelos Bletsas
School of Electronic and Computer Engineering, Technical University of Crete, Chania, Greece
Emails: {nfasarakis,palevizos}@.isc.tuc.gr, aggelos@telecom.tuc.gr

ABSTRACT

Distributed cooperative network location estimation in motionless and mobile narrowband wireless sensor networks (WSNs) is studied, where *agents* (to be localized) are several hops away from the *anchors* (with a priori known location). This work proposes a reduced-communication cooperative and distributed particle filtering (CoopPF) approach based on variational inference (VI) Gaussian mixture modeling (GMM), where network nodes exchange information locally, i.e. only with neighboring terminals; each node transmits the parameters of the estimated Gaussian mixture, instead of the whole posterior density, offering tremendous reduction in communication overhead, as required in narrowband applications (e.g. underwater communications). The proposed VI approach *jointly* estimates the number of required Gaussians and their parameters, in sharp contrast to standard expectation-maximization techniques, where the number of components must be estimated first with other techniques (e.g. clustering). Accuracy comparable to state-of-the-art PF cooperative localization is demonstrated, with an order of magnitude reduction in communication overhead.

1. INTRODUCTION

Cooperative localization, i.e localization where agents (of unknown location) perform ranging and exchange measurements with other neighboring agents and possibly with anchors (of a priori known location), has been the focus of intense research over the last years. Distributed location estimation techniques include distributed particle filtering (PF) [1, 2] as well as factor graphs (FG) [3]. Estimation accuracy offered by the aforementioned setups comes at the cost of increased communication overhead, since the transmission of posterior location estimate distributions is communication bandwidth-intensive.

As a response, distributed PF implementations with posterior approximations have been proposed, that reduce communication overhead. Such approximations are usually accomplished by parametric techniques with Gaussian mixture modeling (GMM). For example, work in [4] proposes a distributed PF utilizing GMM and expectation-maximization (EM) assuming apriori known number of Gaussian components, while work in [5] approximates posterior distributions with a single Gaussian distribution (assuming availability of a global navigation satellite system-GNSS). Relevant work can also be found in [6], where the authors proposed parametric representation of posteriors as a multivariate Gaussian distribution. Finally, works in [7, 8] utilize clustering and sample-based techniques to estimate the number of Gaussian components, and then through EM, the posterior location estimates are approximated as GMMs. If the EM algorithm is executed with an excessively large number of mixture components, it may produce communication redundancy (due to large number of utilized mixtures), while a smaller number of components may offer reduced accuracy [8].

In sharp contrast to the above, this work proposes GMM based on *variational inference*, in order to *minimize* the information communicated between nodes *and* infer the required number of Gaussians in the mixture that approximates the node posteriors for motionless and mobile network localization. More importantly, simulation results quantify the impact of accuracy-communication trade-off within the GMM framework and corroborate the efficiency of the proposed variational approach. The latter offers comparable accuracy compared to state-of-the-art nonparametric cooperative PF (CoopPF) localization.

The paper is organized as follows: Section 2 provides the system model and basic assumptions, Section 3 presents the CoopPF procedure and then, offers the proposed parametric variational GMM PF-based cooperative localization algorithm. Section 4 discusses the numerical results and work is concluded in Section 5.

2. SYSTEM MODEL

A network of $\mathcal{N} + \mathcal{N}_{\mathcal{A}}$ nodes is assumed, with $\mathcal{N}$ *agents* of unknown locations (to be estimated) and $\mathcal{N}_{\mathcal{A}}$ *anchors* with a priori known coordinates. Time is slotted and the set of all nodes at time instant t is denoted as $H^{(t)}$, while the set of nodes with a communication link to an arbitrary node i is denoted as $H^{(t)}(i)$. The true 2-dimensional (2D) coordinates of node i at time t are represented by a time-dependent state vector $\mathbf{x}_i^{(t)} = [x_i^{(t)} \;\; y_i^{(t)}]^T$ (i.e., dimension $\mathcal{D} = 2$).

It is assumed that at time t each node i can measure vector $\mathbf{z}_i^{(t)} \triangleq [(\mathbf{z}_{i,\text{self}}^{(t)})^T \;\; (\mathbf{z}_{i,\text{rel}}^{(t)})^T]^T$, where $\mathbf{z}_{i,\text{self}}^{(t)}$ is based on internal measurements (e.g. odometer) and $\mathbf{z}_{i,\text{rel}}^{(t)}$ is the vector of range

The work of P. Alevizos and A. Bletsas (Sections 1., 2., and 4.) was supported by the European Union (European Social Fund - ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework through the Research Funding Program "Thales - Investing in knowledge society through the European Social Fund."

measurements between node i and neighbors in $H^{(t)}(i)$. In this work, the ranging measurements between nodes i and j are modeled as:

$$\widehat{d}_{j\to i}^{(t)} = d_{j\to i}^{(t)} + w_{j\to i}^{(t)}, \quad (1)$$

where $d_{j\to i}^{(t)} = \|\mathbf{x}_i^{(t)} - \mathbf{x}_j^{(t)}\|_2$ is the Euclidean distance between nodes i and j, while $w_{j\to i}^{(t)}$ are independent, zero-mean Gaussian random variables with variance σ_r^2, assumed independent of past and present states. Each agent i aims to calculate the posterior distribution of its state $\mathbf{x}_i^{(t)}, \forall t$, conditioned on all measurements up to time t, $p(\mathbf{x}_i^{(t)}|\mathbf{z}_i^{(1:t)}), i = 1, 2, \ldots, \mathcal{N}$. Calculating the mean of the posterior offers the minimum mean squared error estimate of the node's position at time t.

3. A VARIATIONAL INFERENCE GMM APPROACH FOR COOPERATIVE LOCALIZATION

3.1. Cooperative Particle Filtering (CoopPF) Revisited

The cooperative particle filtering (CoopPF) approach tracks the state of each node as it evolves over time and involves iterative exchange of location estimates among neighboring nodes [8]. In a distributed implementation, each node i maintains a population of M weights $\mathbf{w}_{i,t} = \{w_{i,t}^{[m]}\}_{m=1}^{M}$, as well as M location samples $\mathcal{X}_i^{(t)} = \{\mathbf{x}_{i,t}^{[m]}\}_{m=1}^{M}$ - the *particles* - that represent the possible locations of that node at time t, with possibility of particle $\mathbf{x}_{i,t}^{[m]}$ proportional to its corresponding weight $w_{i,t}^{[m]}$. The posterior $p(\mathbf{x}_i^{(t)}|\mathbf{z}_i^{(1:t)})$ can be constructed from particles and weights using the histogram approach. The posterior is approximated using a discrete grid of regions of area δ^2 each; the posterior value at each grid region is proportional to the sum of weights from all particles that fall within that region. We denote by $\mathcal{H}_{\text{ist}}(\mathbf{w}_{i,t}, \mathcal{X}_i^{(t)})$ the function that produces the histogram.

Assuming that nodes move independently according to a memoryless walk, the particle set $\mathcal{X}_i^{(t)}$, at time t, is constructed through $\mathcal{X}_i^{(t-1)}$. Specifically, the new m-th particle of node i is generated according to the (known) state transition distribution, which is based on the previous m-th particle:

$$\mathbf{x}_{i,t}^{[m]} \propto p\left(\mathbf{x}_i^{(t)} = \mathbf{x}_{i,t}^{[m]} |\, \mathbf{x}_i^{(t-1)} = \mathbf{x}_{i,t-1}^{[m]} \in \mathcal{X}_i^{(t-1)}\right). \quad (2)$$

Note that the aforementioned step (referred to as *prediction*) can be performed individually by each node.

For each new particle $\mathbf{x}_{i,t}^{[m]}$ of each node i, a new weight $w_{i,t}^{[m]}$ must be calculated for the set of collected ranging measurements, as well as internal measurements $\mathbf{z}_{i,\text{self}}^{(t)}$ in what is known as the *correction* operation. Assuming independence between ranging and internal measurements, as well as independence between ranging measurements, the following holds:

$$w_{i,t}^{[m]} = p\left(\mathbf{z}_i^{(t)} |\, \mathbf{x}_{i,t}^{[m]}\right) = p\left(\mathbf{z}_{i,\text{self}}^{(t)} |\, \mathbf{x}_{i,t}^{[m]}\right) \prod_{j\in H^{(t)}(i)} p\left(\widehat{d}_{j\to i}^{(t)} |\, \mathbf{x}_{i,t}^{[m]}\right)$$

$$= \left(\int p\left(\mathbf{z}_{i,\text{self}}^{(t)} |\, \mathbf{x}_i^{(t-1)}, \mathbf{x}_{i,t}^{[m]}\right) \widehat{p}\left(\mathbf{x}_i^{(t-1)} |\mathbf{z}_i^{(1:t-1)}\right) d\mathbf{x}_i^{(t-1)}\right) \cdot$$

$$\cdot \prod_{j\in H^{(t)}(i)} \left(\int p\left(\widehat{d}_{j\to i}^{(t)} |\, \mathbf{x}_{i,t}^{[m]}, \mathbf{x}_j^{(t)}\right) \widehat{p}\left(\mathbf{x}_j^{(t)} |\mathbf{z}_j^{(1:t)}\right)^{(l-1)} d\mathbf{x}_j^{(t)}\right). \quad (3)$$

For the calculation of $p\big(\widehat{d}_{j\to i}^{(t)} |\mathbf{x}_{i,t}^{[m]}\big)$ node i integrates the known conditional probability density function (p.d.f.) of the measured distance $\widehat{d}_{j\to i}^{(t)}$ between itself and neighbor $j \in H^{(t)}(i)$ at time t and the posterior estimate of node j. During each time step t, each node stores a representation of its most up to date posterior denoted as $\widehat{p}(\mathbf{x}_i^{(t)} |\mathbf{z}_i^{(1:t)})^{(l)}$. Within the same time step, nodes broadcast their posteriors (in parallel) and receive posteriors from neighbors. The received posteriors are then utilized in conjunction with the ranging measurements at time t and through Eq. (3) all nodes update the histogram representation of their own posterior. The posteriors are then broadcasted back to the network, hence the superscript l above. After N_{iter} iterations, the parallel iterative correction operation terminates and the node posteriors are offered by $p(\mathbf{x}_i^{(t)} |\mathbf{z}_i^{(1:t)}) \equiv \widehat{p}(\mathbf{x}_i^{(t)} |\mathbf{z}_i^{(1:t)})^{(N_{\text{iter}})}, \ \forall i \in \mathcal{N}$.

Functions $p(\mathbf{z}_{i,\text{self}}^{(t)} |\, \mathbf{x}_i^{(t-1)}, \mathbf{x}_{i,t}^{[m]})$ and $p(\mathbf{x}_i^{(t-1)} |\mathbf{z}_i^{(1:t-1)})$ (second line of Eq. (3)) are used for the calculation of p.d.f. $p\big(\mathbf{z}_{i,\text{self}}^{(t)} |\, \mathbf{x}_{i,t}^{[m]}\big)$. Both of them, are local to node i and can be calculated with sensor i's own information. The former is the conditional p.d.f. of the internal measurement of node i at time t while the later is simply the posterior distribution of node i's state computed during time $t-1$. CoopPF algorithm also includes a *resampling* of the particles and in this work, the *low variance sampler* is utilized [9]. The resulted weights of the new particles have intelligently incorporated the measurements, offering for $M \to \infty$ a representation of the unknown posterior.

3.2. Variational Inference for GMMs in CoopPF

The iterative exchange of posteriors between nodes is a communication intensive process due to the fact that each transmitted posterior maps $\mathbb{R}^{\mathcal{D}} \mapsto \mathbb{R}$ and thus, significant amount of communication overhead is required for transmission. The information communicated between nodes can be minimized through parametric approximations of the broadcasted posteriors, offering tremendous communication gains. Gaussian mixture models (GMM), defined as a weighted combination of Gaussian densities, are an attractive choice as flexible parametric approximations to any given density.

During each iteration l and for each agent i, variational CoopPF algorithm approximates the posterior $\widehat{p}(\mathbf{x}_i^{(t)} |\mathbf{z}_i^{(1:t)})^{(l)}$ as a two-dimensional GMM of K components with p.d.f.

$$\widehat{p}\left(\mathbf{x}_i^{(t)} |\mathbf{z}_i^{(1:t)}\right)^{(l)} \approx \sum_{k=1}^{K} \pi_{ik}\, \mathcal{N}\left(\mathbf{x}_i^{(t)}; \boldsymbol{\mu}_{ik}, (\boldsymbol{\Lambda}_{ik})^{-1}\right), \quad (4)$$

Algorithm 1: Variational CoopPF (V-CoopPF)

(1): **Input**: $\mathcal{X}_i^{(0)}$, $\forall i \in H^{(t=0)}$ and the number of particles M.
(2): **for** $t = 1 : T$ (time index)
(3): **for** $\forall\, i \in H^{(t)}$ **in parallel**
(4): Construct $\mathcal{X}_i^{(t)}$ based on $\mathcal{X}_i^{(t-1)}$ through Eq. (2).
(5): Construct parameter vector $\mathbf{p}_i^{(0)}$ of node i at iteration $l = 0$ calculated through Eq. (10) based on $\mathcal{X}_i^{(t)}$.
(6): **end parallel**
(7): **for** $l = 1$ **to** N_{iter} **iterate**
(8): **for** $\forall\, i \in H^{(t)}$ **in parallel**
(9): Broadcast $\mathbf{p}_i^{(l-1)}$ associated with $\widehat{p}\big(\mathbf{x}_i^{(t)}|\mathbf{z}_i^{(1:t)}\big)^{(l-1)}$.
(10): Receive $\mathbf{p}_j^{(l-1)}$ and create $\widehat{p}\big(\mathbf{x}_j^{(t)}|\mathbf{z}_j^{(1:t)}\big)^{(l-1)}$ through Eq. (4), $\forall j \in H^{(t)}(i)$.
(11): Calculate $w_{i,t}^{[m]}$ for each particle $\mathbf{x}_{i,t}^{[m]}$ through Eq. (3).
(12): $\mathcal{X}_i^{(t)} := \text{LowVarianceSampler}\Big(\mathbf{w}_{i,t}, \mathcal{X}_i^{(t)}\Big)$ [9].
(13): **if** $(l < N_{\text{iter}})$ calculate $\mathbf{p}_i^{(l-1)}$ from $\mathcal{X}_i^{(t)}$ through Eq. (10).
(14): **else** $(l = N_{\text{iter}})$ $\widehat{p}\big(\mathbf{x}_i^{(t)}|\mathbf{z}_i^{(1:t)}\big)^{(l)} = \mathcal{H}_{\text{ist}}\{\mathbf{1}_M/M, \mathcal{X}_i^{(t)}\}$.
(15): **end parallel**
(16): **end for**
(17): **end for** (time index)

where $\boldsymbol{\pi}_i = \{\pi_{ik}\}$ denotes the scalar mixture weights, $\boldsymbol{\mu}_i = \{\boldsymbol{\mu}_{ik}\}$ denotes the component means and $\boldsymbol{\Lambda}_i = \{\boldsymbol{\Lambda}_{ik}\}$ denotes the precision (inverse covariance) matrices. We define the compound vector of the parameters corresponding to $\widehat{p}\big(\mathbf{x}_i^{(t)}|\mathbf{z}_i^{(1:t)}\big)^{(l)}$ as $\mathbf{p}_i^{(l)} \triangleq \{\boldsymbol{\pi}_i, \boldsymbol{\mu}_i, \boldsymbol{\Lambda}_i\}$. The number of mixture components K is assumed a-priori unknown and the estimation of all necessary parameters *relies only* on the resampled particle set $\mathcal{X}_i^{(t)}$.

The most common technique to estimate the parameters $\mathbf{p}_i^{(l)}$ in Eq. (4) is through expectation-maximization (EM). However, in cases where the number of components K is unknown, EM is not a suitable option. This work proposes a variational inference Bayesian GMM estimation approach that can overpass such problem and the number of components can be estimated during the variational inference procedure [10, pp. 481]. The variational GMM inference algorithm can be initialized by a relatively large number of mixture components (K) and as iterative steps proceed, such redundancy is eliminated.

The variational approach utilizes a latent variable $\boldsymbol{\omega}_{i,t}^{[m]} \in \{0,1\}^K$ for each particle $\mathbf{x}_{i,t}^{[m]} \in \mathcal{X}_i^{(t)}$ representing the mixture component responsible for generating the particle. The K elements of $\boldsymbol{\omega}_{i,t}^{[m]}$ are such that $[\boldsymbol{\omega}_{i,t}^{[m]}]_k = 1$ when $\mathbf{x}_{i,t}^{[m]}$ has been generated by the k-th component and zero otherwise. The set of latent variables is denoted as $\boldsymbol{\Omega}_i^{(t)} = \{\boldsymbol{\omega}_{i,t}^{[m]}; 1 \leq m \leq M\}$. Within variational GMM framework the parameters $\boldsymbol{\theta}_i \triangleq \{\boldsymbol{\Omega}_i^{(t)}, \boldsymbol{\pi}_i, \boldsymbol{\mu}_i, \boldsymbol{\Lambda}_i\}$ are treated as stochastic, whose priors belong to *specific parametric distributions* that enjoy special properties. Specifically, the joint p.d.f. of $\mathcal{X}_i^{(t)}$ and $\boldsymbol{\theta}_i$ is assumed to factorize as:

$$p\Big(\mathcal{X}_i^{(t)}, \boldsymbol{\Omega}_i^{(t)}, \boldsymbol{\pi}_i, \boldsymbol{\mu}_i, \boldsymbol{\Lambda}_i\Big) = p\left(\mathcal{X}_i^{(t)}|\boldsymbol{\Omega}_i^{(t)}, \boldsymbol{\mu}_i, \boldsymbol{\Lambda}_i\right) p\left(\boldsymbol{\pi}_i\right) \cdot \cdot\, p\Big(\boldsymbol{\Omega}_i^{(t)}|\boldsymbol{\pi}_i\Big)\, p\left(\boldsymbol{\mu}_i|\boldsymbol{\Lambda}_i\right) p\left(\boldsymbol{\Lambda}_i\right) \quad (5)$$

where,

$$p(\mathcal{X}_i^{(t)}|\boldsymbol{\Omega}_i^{(t)}, \boldsymbol{\mu}_i, \boldsymbol{\Lambda}_i) = \prod_{m=1}^{M}\prod_{k=1}^{K} \mathcal{N}(\mathbf{x}_{i,t}^{[m]}; \boldsymbol{\mu}_{ik}, \boldsymbol{\Lambda}_{ik})^{[\boldsymbol{\omega}_{i,t}^{[m]}]_k}, \quad (6)$$

while the rest priors (associated with parameters $\boldsymbol{\theta}_i$) are factorized as $p(\boldsymbol{\theta}_i) = p\Big(\boldsymbol{\Omega}_i^{(t)}|\boldsymbol{\pi}_i\Big)\, p(\boldsymbol{\pi}_i) p(\boldsymbol{\mu}_i|\boldsymbol{\Lambda}_i) p(\boldsymbol{\Lambda}_i)$,

$$p(\boldsymbol{\theta}_i) = \left(\prod_{m=1}^{M}\prod_{k=1}^{K} \pi_{ik}^{[\boldsymbol{\omega}_{i,t}^{[m]}]_k}\right)\Big(\mathsf{Dir}(\boldsymbol{\pi}_i; \boldsymbol{\alpha}_0)\Big)\cdot$$
$$\cdot\left(\prod_{k=1}^{K}\mathcal{N}\big(\boldsymbol{\mu}_{ik}; \mathbf{0}, (\beta_0\boldsymbol{\Lambda}_{ik})^{-1}\big)\right)\left(\prod_{k=1}^{K}\mathcal{W}\big(\boldsymbol{\Lambda}_{ik}; \nu_0, \mathbf{W}_0\big)\right). \quad (7)$$

Symbol $\mathcal{W}(\cdot; \nu_0, \mathbf{W}_0)$ denotes the Wishart distribution parametrized by variables ν_0 and $\mathbf{W}_0$, while $\mathsf{Dir}(\cdot; \boldsymbol{\alpha}_0)$ denotes the Dirichlet distribution parametrized by vector $\boldsymbol{\alpha}_0$ [10, Appendix B]. The initial values of distribution parameters $\{\beta_0, \nu_0, \mathbf{W}_0, \boldsymbol{\alpha}_0\}$ are omitted here due to space constraints, and can be found in [10, pp. 474–475]. The priors of Eq. (7) are also called conjugate priors.

The standard variational estimation approach aims to maximize the lower bound $\mathcal{L}(q)$ of the logarithmic marginal likelihood $\ln\Big(p\Big(\mathcal{X}_i^{(t)}\Big)\Big) = \mathcal{L}(q) + \mathsf{KL}(q||p)$,

$$\mathcal{L}(q) = \int q(\boldsymbol{\theta}_i) \ln\left(\frac{p\Big(\mathcal{X}_i^{(t)}, \boldsymbol{\theta}_i\Big)}{q(\boldsymbol{\theta}_i)}\right) d\boldsymbol{\theta}_i. \quad (8)$$

The above quantity is a lower bound on $\ln\Big(p\Big(\mathcal{X}_i^{(t)}\Big)\Big)$ for any possible distribution on $\boldsymbol{\theta}_i$ (denoted by $q(\boldsymbol{\theta}_i)$). This observation stems from fact that Kullback–Leibler divergence between $q(\boldsymbol{\theta}_i)$ and $p\Big(\mathcal{X}_i^{(t)}\Big)$, given by $\mathsf{KL}(q||p) = -\int q(\boldsymbol{\theta}_i) \ln\left(\frac{p\left(\boldsymbol{\theta}_i|\mathcal{X}_i^{(t)}\right)}{q(\boldsymbol{\theta}_i)}\right) d\boldsymbol{\theta}_i$, is always non-negative for any possible $q(\boldsymbol{\theta}_i)$. For mathematical convenience, variational approach assumes that the maximization of $\mathcal{L}(q)$ is carried out over all possible distributions $q(\boldsymbol{\theta}_i)$ that adhere to the following factorization: $q(\boldsymbol{\theta}_i) = q(\boldsymbol{\Omega}_i^{(t)}) q(\boldsymbol{\pi}_i) q(\boldsymbol{\mu}_i|\boldsymbol{\Lambda}_i) q(\boldsymbol{\Lambda}_i)$.

It turns out that if the priors are initialized through Eq. (7), then, the optimal distribution $q^*(\boldsymbol{\theta}_i)$ that maximizes $\mathcal{L}(q)$ is given by

$$q^*(\boldsymbol{\theta}_i) = q^*\Big(\boldsymbol{\Omega}_i^{(t)}\Big)\, q^*(\boldsymbol{\pi}_i)\, q^*(\boldsymbol{\Lambda}_i)\, q^*(\boldsymbol{\mu}_i|\boldsymbol{\Lambda}_i)$$
$$= \left(\prod_{m=1}^{M}\prod_{k=1}^{K} r_{km}^{[\boldsymbol{\omega}_{i,t}^{[m]}]_k}\right)\Big(\mathsf{Dir}(\boldsymbol{\pi}_i; \boldsymbol{\alpha})\Big)\left(\prod_{k=1}^{K}\mathcal{W}\big(\boldsymbol{\Lambda}_{ik}; \nu_k, \mathbf{W}_k\big)\right)\cdot$$
$$\cdot\left(\prod_{k=1}^{K}\mathcal{N}\big(\boldsymbol{\mu}_{ik}; \mathbf{m}_k, (\beta_k\boldsymbol{\Lambda}_{ik})^{-1}\big)\right). \quad (9)$$

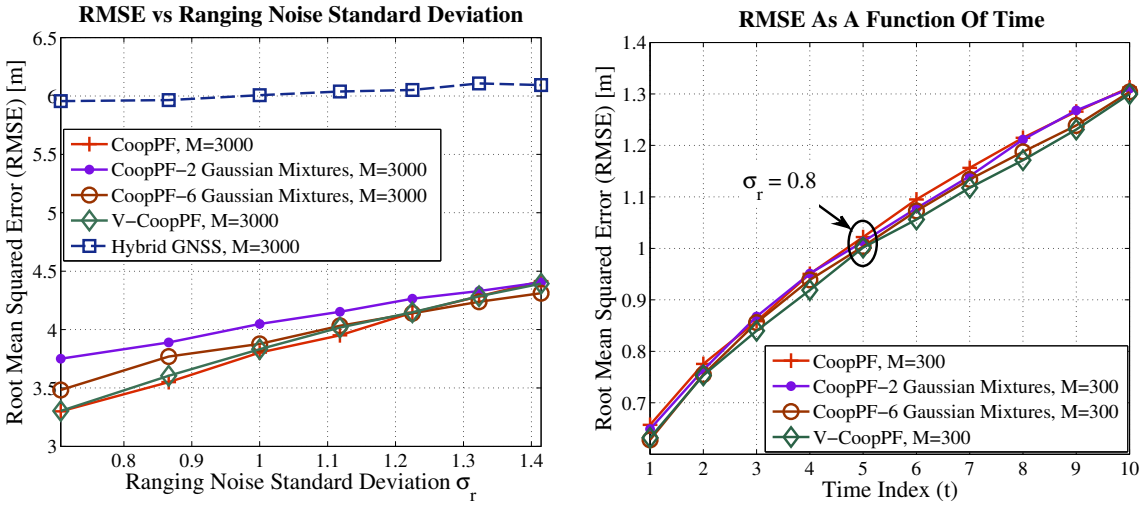

Fig. 1. Variational CoopPF offers accuracy and communication gains. Left: RMSE as a function of ranging noise standard deviation σ_r for a motionless network. Right: RMSE as a function of time for a mobile network and fixed value of σ_r.

The intuition here is that the initialization of the parameters $\boldsymbol{\theta}_i$ through Eq. (7), leads to optimal distribution $q^*(\boldsymbol{\theta}_i)$ that has simple factorization form (of Eq. (9)), which in turn has the same form with the product of conjugate priors of Eq. (7). The mathematical machinery in this section is introduced to find tractable update expressions for density parameters $\{r_{km}\}_{m,k=1}^{M,K}, \{\alpha_k, \mathbf{m}_k, \beta_k, \nu_k, \mathbf{W}_k\}_{k=1}^{K}$ in Eq. (9), omitted here due to space constraints (and can be found in [10, pp. 476–479]. The variational update equations are coupled and must be solved iteratively. The update rules on density parameters are iteratively applied until a change less than a predetermined threshold ϵ in the value of $\mathcal{L}(q^*)$ between successive updates is found. The final values of the estimated mixture parameters of Eq. (4), $\{\widehat{\pi}_{ik}, \widehat{\boldsymbol{\mu}}_{ik}, \widehat{\boldsymbol{\Lambda}}_{ik}\}_{k=1}^{K}$, are finally given by ($k = 1, 2, \ldots, K$):

$$\widehat{\pi}_{ik} = \frac{\sum_{m=1}^{M} r_{km}}{M}, \quad \widehat{\boldsymbol{\mu}}_{ik} = \mathbf{m}_k, \quad \widehat{\boldsymbol{\Lambda}}_{ik} = \nu_k \mathbf{W}_k. \tag{10}$$

4. NUMERICAL RESULTS

Numerical results are provided for 2D localization for both static and mobile scenarios. The 30×30 topology of [8, Fig. 1] with 20 agents and 4 anchors is considered with grid resolution $\delta = 1$m. The reported root mean squared error (RMSE) has been calculated after 500 experiments per reported noise standard deviation. For static/mobile scenarios the number of particles used is $M = \{3000, 300\}$, respectively.

Fig. 1-left offers the RMSE calculated across all agents, as a function of the ranging noise standard deviation σ_r. It is observed that utilizing 2 Gaussian mixtures (GMs) with EM degrades performance compared to V-CoopPF, verifying that parametric approximations need to be designed more carefully. On the other hand, it is shown that the proposed algorithm offers RMSE similar to CoopPF, even though V-CoopPF utilizes exchanged messages of significantly reduced size. It is remarked that CoopPF schemes (solid lines) offer smaller RMSE compared to the technique in [5] (dashed line), since the latter was designed for networks with dense anchor deployment.

Table 1. Communication requirements and RMSE for static scenario and $\sigma_r = 0.8660$

Algorithms	RMSE $[m]$	Average real numbers
CoopPF	3.5504	~ 40362
V-CoopPF	3.6041	~ 1576
2-GMM CoopPF	3.8897	~ 532
6-GMM CoopPF	3.7687	~ 1596
Hybrid GNSS	5.9648	~ 240

Table 1 offers the size of the exchanged messages among all nodes and the corresponding RMSE across all agents for $\sigma_r = 0.8660$. The size of the messages is expressed as the total real numbers broadcasted in the network. V-CoopPF reduces the total numbers of exchanged information 1 order of magnitude compared to CoopPF.

Finally, a 2D tracking scenario is considered, in which agents begin with perfect prior location. The latter are mobile and during each time step travel a known distance $d_i^{(t)} \sim \mathcal{N}(0, 1)$, in an unknown direction $\theta_i^{(t)} \sim \mathcal{U}[0, 2\pi), \forall i$, with $\sigma_r = 0.8$. Fig. 1-right offers the RMSE as the state evolves over $T = 10$ time steps. V-CoopPF as well as CoopPF with 2 and 6 GMs achieve slightly improved RMSE performance compared to the CoopPF because only 300 particles are utilized for the calculation of particle weights. As expected, V-CoopPF offers lower RMSE compared to CoopPF with 2 and 6 GMs due to more careful mixture modeling.

5. CONCLUSION

This work designed a novel variational inference GMM-based cooperative particle filter for cooperative node localization. Total communication overhead was reduced, with accuracy comparable to state-of-the-art PF cooperative localization.

6. REFERENCES

[1] M. Coates, "Distributed particle filters for sensor networks," in *Proc. ACM IPSN*, Berkeley, California, USA, Apr. 2004, pp. 99–107.

[2] O. Hlinka, F. Hlawatsch, and P. M. Djuric, "Distributed particle filtering in agent networks: A survey, classification, and comparison." *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 61–81, Jan 2013.

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

[4] X. Sheng, Y.-H. Hu, and P. Ramanathan, "Distributed particle filter with GMM approximation for multiple targets localization and tracking in wireless sensor network," in *Proc. ACM IPSN*, Los Angeles, California, USA, Apr. 2005, pp. 181–188.

[5] F. Sottile, H. Wymeersch, M. A. Caceres, and M. A. Spirito, "Hybrid gnss-terrestrial cooperative positioning based on particle filter," in *Proc. IEEE Global Commun. Conf. (Globecom)*, Houston, Texas, USA, 2011, pp. 1–5.

[6] M. A. Caceres, F. Penna, H. Wymeersch, and R. Garello, "Hybrid cooperative positioning based on distributed belief propagation," *IEEE J. Select. Areas Commun.*, vol. 29, pp. 1948–1958, Dec. 2011.

[7] P. Alevizos, N. Fasarakis-Hilliard, and A. Bletsas, "Cooperative localization in wireless networks under bandwidth constraints," in *Proc. Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov. 2012, pp. 904–908.

[8] N. Fasarakis-Hilliard, P. N. Alevizos, and A. Bletsas, "Narrowband cooperative network localization," in *Proc. IEEE Global Commun. Conf. (Globecom)*, Austin, TX, USA, Dec. 2014.

[9] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, MA: The MIT Press, 2006.

[10] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY: Springer, 2006.