

# Distributed and Centralized Observers for Sensor Fusion with Non-Gaussian Noises

Duarte Silva

*Dept. Electrical and Computer Engineering*

*Instituto Superior Tecnico*

duarte.leao@tecnico.ulisboa.pt

**Abstract**—This work addresses the challenges of State Estimation (SE) for linear systems under diverse noise conditions using both single and multi-sensor frameworks. Traditional SE methods often assume Gaussian noise, limiting their effectiveness in non-Gaussian scenarios. To overcome these limitations, this work explores advanced filtering techniques leveraging Characteristic Functions (CFs) for handling non-Gaussian noise. Additionally, we investigate the Wasserstein Barycenter (WB) methodology to reliably integrate sensor data with diverse noise patterns. Two hybrid SE approaches are proposed, combining CF filtering with particle-based methods to address limitations of traditional Gaussian assumptions. The first uses a Dirac-based approximation for CF discretization, while the second employs Gaussian mixtures. Both methods are evaluated across Gaussian, Gaussian mixture, and exponential noise scenarios for single-sensor configurations. These hybrid methods are benchmarked against classical techniques such as the Kalman Filter (KF) and Particle Filter (PF). Results indicate that hybrid CF-based filters demonstrate superior robustness and accuracy in non-Gaussian environments. The second part of this work examines the use of WB for sensor fusion in multi-sensor systems. The WB approach enhances SE accuracy by integrating data from sensors with diverse noise characteristics, offering a robust framework for minimizing the impact of inconsistent noise. The effectiveness of the WB-based approach is validated through simulations, highlighting its advantages in scenarios with heterogeneous sensor noise.

**Index Terms**—State Estimation; Non-Gaussian Noise; Bayes Filter; Wasserstein Barycenter; Multi-Agent Systems; Kalman Filter; Characteristic Function Filter

## I. INTRODUCTION

State Estimation (SE) is a fundamental concept with widespread applications in engineering. In control systems, SE supports feedback mechanisms for stability and performance. In robotics, SE is critical for localization, mapping, and navigation, while in signal processing, it aids in filtering noise and reconstructing signals. The increasing complexity of systems, along with advancements in big data and the Internet of Things (IoT), has driven demand for robust and efficient SE algorithms, especially with the growth of sensor networks requiring real-time operation.

Conventional SE techniques, such as the Kalman Filter (KF) [1], assume Gaussian noise, which often fails to capture real-world conditions. The Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) [2], [3] address non-linearities but still assume Gaussian noise, limiting their effectiveness under non-Gaussian disturbances. Particle Filters (PFs) [4]

handle non-Gaussian and nonlinear systems well but are computationally intensive, particularly for real-time use.

Traditional SE methods are mainly designed for centralized, single-sensor architectures. As systems move towards multi-sensor, distributed, and decentralized configurations, new challenges in data fusion and consensus arise. These architectures are more scalable and robust but require advanced SE solutions for reliable performance, especially under non-Gaussian noise.

This thesis has two main objectives: to explore the use of Characteristic Functions (CFs) to enhance SE performance under non-Gaussian noise, and to investigate the Wasserstein Barycenter (WB) for multi-sensor data fusion to improve accuracy and robustness in distributed SE scenarios.

## II. SINGLE SENSOR STATE ESTIMATION

### A. Related Work

This section provides an overview of the current state of research on SE methods for single-sensor systems, focusing on both Gaussian and non-Gaussian noise assumptions, as well as the role of CFs in the filtering process.

1) *Bayes Filter*: At the foundation of SE for linear systems is the Bayes Filter, which offers a probabilistic framework for estimating system states from noisy observations. The Bayes Filter is applicable in both single-agent (SA) and multi-agent (MA) architectures and is effective under both Gaussian and non-Gaussian noise conditions. Bayes filtering is a recursive process composed of two fundamental steps: prediction and update, enabling continuous refinement of the state estimate as new data becomes available.

- **Prediction Step**: This step propagates the prior belief about the system's state forward in time by incorporating system dynamics and inherent uncertainties. Mathematically, it is represented as:

$$p(x_t|z_{1:t-1}) = \int p(x_t|x_{t-1}) \cdot p(x_{t-1}|z_{1:t-1}) dx_{t-1} \quad (1)$$

where  $x_t$  represents the state at time  $t$ ,  $z_{1:t-1}$  denotes the observation sequence up to time  $t-1$ ,  $p(x_t|x_{t-1})$  is the transition model characterizing the system dynamics, and  $p(x_{t-1}|z_{1:t-1})$  is the prior state estimate at time  $t-1$ .

- **Update Step**: This step refines the predicted state by incorporating the new observation at time  $t$ . It adjusts

the probability of the state based on the likelihood of the new observation given the predicted state:

$$p(x_t|z_{1:t}) = \frac{p(z_t|x_t) \cdot p(x_t|z_{1:t-1})}{p(z_t|z_{1:t-1})} \quad (2)$$

where  $z_t$  represents the new observation at time  $t$ ,  $p(z_t|x_t)$  is the likelihood of the observation given the state, and  $p(x_t|z_{1:t-1})$  is a normalization factor ensuring that the posterior distribution is a valid probability distribution.

2) *Kalman Filter*: The KF [1], introduced by R.E. Kalman in 1960, provides an optimal solution for SE in linear systems with Gaussian noise. Its recursive nature and matrix-based formulation make it computationally efficient for many real-time applications. However, the KF is heavily reliant on the Gaussian noise assumption, and its performance degrades significantly in environments characterized by non-Gaussian noise. To address system nonlinearities, extensions such as the EKF and UKF [2], [3] have been developed, yet these approaches continue to assume Gaussian noise, limiting their applicability in more complex noise scenarios.

3) *Particle Filter*: The PF [4] offers a powerful approach for SE in systems with nonlinear dynamics and non-Gaussian noise. It approximates the posterior probability density function (PDF) using a set of weighted samples, commonly referred to as particles. Although the PF is effective at handling non-Gaussian noise, the computational burden associated with managing a large number of particles is substantial, particularly in high-dimensional state spaces, which presents a significant challenge for real-time applications.

4) *Gaussian Mixture Filter*: The GMF [5] leverages Gaussian Mixture Models (GMMs) to approximate the system's state distribution. By employing Gaussian mixture reduction techniques [6], the GMF mitigates the exponential growth of mixture components during the estimation process, providing a computationally efficient solution under non-Gaussian noise conditions.

5) *Characteristic Function Filter*: Recent research has investigated the use of CFs for SE in non-Gaussian systems. CFs represent the Fourier transform of probability density functions, offering an alternative means to efficiently perform operations such as convolution in the frequency domain. This approach significantly simplifies the update step by transforming complex operations into multiplications [7], [8]. The CF Filter follows a process similar to the Bayes Filter, but with the convolution operation occurring in the update step rather than the prediction step. Techniques such as the use of the inverse fast Fourier transform (IFFT) [9] enable the CF Filter to function effectively in real-time applications, making it a promising approach for SE in complex, non-Gaussian environments.

## B. Single Sensor Problem

The single sensor problem is formulated to estimate the state of a dynamic system from the perspective of a single sensor. Specifically, SE in this context involves accurately determining

the system's state, denoted as  $x_t$ , at each time step  $t$ , using a sequence of observations  $z_1, z_2, \dots, z_t$  from the sensor. The goal is to derive an estimate,  $\hat{x}_t$ , that closely approximates the true state of the system by processing noisy and possibly incomplete observations.

To mathematically represent the single sensor problem, a state-space model is employed, consisting of two main equations: the state equation and the observation equation. The formulation for these equations is as follows:

- **State Equation**: This equation characterizes how the system state evolves over time:

$$x_t = Ax_{t-1} + Bu_t + d_t \quad (3)$$

Here,  $x_t \in \mathbb{R}^n$  represents the state vector, while  $A \in \mathbb{R}^{n \times n}$  is the state transition matrix that defines the evolution dynamics.  $B \in \mathbb{R}^{n \times d}$  maps the control input  $u_t \in \mathbb{R}^d$  to the state, and  $d_t$  is the process noise, modeled by a distribution  $\mathcal{D}(\alpha)$ .

- **Observation Equation**: This equation links the current state of the system to the sensor measurements:

$$z_t = Hx_t + v_t \quad (4)$$

Here,  $z_t \in \mathbb{R}^p$  represents the observations, while  $H \in \mathbb{R}^{p \times n}$  is the observation matrix that maps the state space to the observation space.  $v_t$  is the measurement noise, modeled by  $\mathcal{D}(\theta)$ .

The model relies on several key assumptions:

- The state and observation equations are linear, indicating linear relationships among the variables.
- The system is time-invariant, as indicated by the constant matrices  $A$ ,  $B$ , and  $H$ .
- The matrices are known, and the pair  $(A, H)$  is observable.
- The noise terms  $d_t$  and  $v_t$  are independent and have known distributions.

## C. Hybrid Filter Approach

The Hybrid Filter is formulated to address the computational complexities inherent in Bayesian filtering by integrating both Bayesian and CF frameworks. In Bayesian filtering, the prediction step requires the evaluation of a convolution integral, which can be computationally expensive, especially for high-dimensional problems. By leveraging the CF framework, this convolution can be transformed into a simple multiplication operation in the frequency domain, thereby significantly enhancing computational efficiency.

Nevertheless, while the CF framework simplifies the prediction step, it introduces additional complexity during the update step, necessitating convolution operations at this stage. To mitigate these challenges, the Hybrid Filter alternates between time and frequency domains, reducing the frequency of repeated Fourier transformations. This hybrid approach thus provides a balanced and computationally efficient solution, particularly when confronted with non-Gaussian noise distributions.

The Hybrid Filter operates via a two-step process:

- **Prediction Step:** Utilizes the CF Filter to simplify the convolution operations within the frequency domain.
- **Update Step:** Implements the PF update step to more effectively handle non-Gaussian noise.

The key innovation of the Hybrid Filter lies in its ability to transition between frequency and time domains by approximating the real CF with a known CF. For instance, an arbitrary CF can be approximated using a Gaussian CF. Given that the Gaussian CF possesses a simple and well-characterized form, this reduces the problem to one of optimization—namely, finding the parameters of the Gaussian distribution that best approximate the original CF.

Once these parameters are determined, both the CF in the frequency domain and the corresponding PDF in the time domain can be derived. The objective of this approximation can be formulated as follows:

$$\begin{aligned} \min_{\theta} \quad & \int |\phi(\boldsymbol{\nu}) - \phi_{\text{approx.}}(\boldsymbol{\nu}, \theta)|^2 d\boldsymbol{\nu} \\ \text{s.t.} \quad & \text{constraints dependent on known CF} \end{aligned} \quad (5)$$

Here,  $\phi$  and  $\phi_{\text{approx.}}$  denote the real and approximated characteristic functions, respectively,  $\boldsymbol{\nu}$  is a vector in  $\mathbb{R}^n$ , and  $\theta$  represents the vector of parameters to be optimized. The specific constraints are dictated by the properties of the CF being used for approximation.

This optimization problem can be adapted depending on the properties of the known CF, particularly with respect to the constraints it imposes.

1) *Hybrid Filter via Dirac Approximation:* The Hybrid Filter via Dirac Approximation builds upon the Hybrid Filter by incorporating a discrete approximation of PDFs derived from their CFs. The CF of a probability distribution can be expressed as an expected value over kernel functions of the form  $e^{i\boldsymbol{\nu}^\top \mathbf{x}}$ . This formulation permits the CF to be discretized as:

$$\phi(\boldsymbol{\nu}) \approx \sum_{k=1}^N p_k e^{i\boldsymbol{\nu}^\top \mathbf{x}_k} \quad (6)$$

where  $N$  represents the number of discrete points  $\mathbf{x}_k$ , and  $p_k$  are the weights indicating the probabilities at these points. This approximation transforms the CF integral into a finite sum, thereby facilitating numerical computation. Using this representation, the original PDF can be approximated by a weighted sum of Dirac distributions.

To derive the discretized PDF, a minimization problem is formulated to fit the discrete approximation of the CF to the actual CF. This is accomplished by minimizing the squared  $L^2$ -norm of the difference between the actual CF and its approximation:

$$\begin{aligned} \min_{p_k} \quad & \int \left| \phi(\boldsymbol{\nu}) - \sum_{k=1}^N p_k e^{i\boldsymbol{\nu}^\top \mathbf{x}_k} \right|^2 d\boldsymbol{\nu} \\ \text{s.t.} \quad & \sum_{k=1}^N p_k = 1 \\ & p_k \geq 0 \quad \text{for all } k = 1, \dots, N. \end{aligned} \quad (7)$$

This formulation ensures that the weights  $p_k$  form a valid probability distribution. In computational terms, this optimization problem can be effectively solved by discretizing the problem over a grid of points  $\boldsymbol{\nu}$ , thereby transforming the continuous integral into a sum over  $M$  grid points. This reduces the problem to a constrained linear least squares problem, where the objective is to find the weights  $p_k$  that best align the discrete CF approximation with the actual CF values at these grid points. The constraints are crucial for maintaining the probabilistic interpretation of the PDF.

$$\begin{aligned} \min_{p_k} \quad & \sum_{l=1}^M \left| \phi(\boldsymbol{\nu}_l) - \sum_{k=1}^N p_k e^{i\boldsymbol{\nu}_l^\top \mathbf{x}_k} \right|^2 \\ \text{s.t.} \quad & \sum_{k=1}^N p_k = 1 \\ & p_k \geq 0 \quad \text{for all } k = 1, \dots, N. \end{aligned} \quad (8)$$

Upon solving this optimization problem, a discretized representation of the PDF is obtained, which corresponds to particles in the PF framework. The PF update can subsequently be applied to these particles, and the updated particles and their respective weights are then used to reconstruct the CF, thereby completing one iteration of the Hybrid Filter via Dirac Approximation.

2) *Hybrid Filter via Gaussian Mixture Approximation:* The Hybrid Filter via Gaussian Mixture Approximation extends the core Hybrid Filter approach by employing Gaussian kernels instead of Dirac kernels to approximate the PDF. In this formulation, the PDF is represented as a mixture of Gaussian components, corresponding to the CF of a GMM.

The CF of a Gaussian distribution centered at  $\boldsymbol{\mu}$  with covariance  $\boldsymbol{\Sigma}$  is given by:

$$\phi_{\text{Gaussian}}(\boldsymbol{\nu}) = e^{i\boldsymbol{\nu}^\top \boldsymbol{\mu}} e^{-\frac{1}{2}\boldsymbol{\nu}^\top \boldsymbol{\Sigma} \boldsymbol{\nu}}. \quad (9)$$

Thus, the CF of a GMM can be represented as a weighted sum of the CFs of its individual Gaussian components. Since any arbitrary distribution can be effectively approximated by a GMM with a sufficient number of Gaussian components, the approximation of the original CF can be expressed as follows:

$$\phi(\boldsymbol{\nu}) \approx \sum_{k=1}^N p_k e^{i\boldsymbol{\nu}^\top \boldsymbol{\mu}_k} e^{-\frac{1}{2}\boldsymbol{\nu}^\top \boldsymbol{\Sigma}_k \boldsymbol{\nu}}, \quad (10)$$

where  $N$  is the number of Gaussian components,  $\boldsymbol{\mu}_k$  and  $\boldsymbol{\Sigma}_k$  represent the mean vector and covariance matrix of the  $k$ -th Gaussian, respectively, and  $p_k$  are the corresponding weights.

To approximate the PDF, we formulate a minimization problem similar to the Dirac approximation, aiming to minimize the squared  $L^2$ -norm of the difference between the actual CF and the GMM approximation:

$$\begin{aligned} \min_{p_k, \mu_k, \Sigma_k} \quad & \int \left| \phi(\nu) - \sum_{k=1}^N p_k e^{i\nu^\top \mu_k} e^{-\frac{1}{2}\nu^\top \Sigma_k \nu} \right|^2 d\nu \\ \text{s.t.} \quad & \sum_{k=1}^N p_k = 1 \\ & p_k \geq 0 \quad \text{for all } k = 1, \dots, N, \\ & \Sigma_k \succ 0 \quad \text{for all } k = 1, \dots, N. \end{aligned} \quad (11)$$

Here, the constraints ensure that the weights  $p_k$  form a valid GMM and that the covariance matrices  $\Sigma_k$  are positive definite, preserving the integrity of the Gaussian components.

In practical implementation, this minimization problem involves discretizing the domain  $\nu$  into grid points, transforming the continuous integral into a finite sum over these points. This process results in a constrained non-linear optimization problem, which can be solved using numerical techniques such as gradient-based optimization methods.

Once the GMM approximation is established, sampling from the GMM provides a discretized representation of the PDF. This representation is subsequently employed in the PF update step. After the update step is completed, the process transitions back to the frequency domain to reconstruct the CF, thereby completing one iteration of the Hybrid Filter via Gaussian Mixture Approximation.

#### D. Experimental Results

This section presents the results of simulations conducted to address the single-sensor problem, with a particular focus on the application of the Hybrid Filter. To evaluate the filter's performance, a simulation was performed to estimate the position of a vehicle undergoing a random walk, modeled using discrete double integrator dynamics with a sampling time of  $T_s = 0.1s$ . These simulations were designed to investigate the effectiveness of the Hybrid Filter under various assumptions regarding the input and observation noise distributions.

In all simulations, the state-space model used is analogous to the one defined in the problem formulation (equations 3 and 4), except that no external disturbance  $d_t$  is considered. Instead, at each time step, the control input  $u_t$  is sampled from either a Gaussian distribution or a GMM with 2 or 3 peaks. The observation model remains consistent across all simulations, with the observation matrix  $H$  defined as:

The observation noise  $v_t$  is drawn from one of three distributions: a Gaussian distribution, a Gaussian mixture, or an exponential distribution. These variations in both the process and observation noise provide a comprehensive testbed for evaluating the Hybrid Filter under different noise conditions.

The objective of each simulation is to estimate the position of the vehicle as it undergoes a random walk. Since the exact initial state of the vehicle is unknown, it is modeled

as a zero-mean Gaussian distribution. Both the Dirac-based and Gaussian Mixture approaches of the Hybrid Filter were tested under all combinations of random walks (Gaussian or GMM) and observation noise types (Gaussian, GMM, and exponential). Thus, nine distinct scenarios were evaluated:

- Gaussian random walk with Gaussian noise.
- Gaussian random walk with exponential noise.
- Gaussian random walk with GMM noise.
- GMM (2 peaks) random walk with Gaussian noise.
- GMM (2 peaks) random walk with exponential noise.
- GMM (2 peaks) random walk with GMM noise.
- GMM (3 peaks) random walk with Gaussian noise.
- GMM (3 peaks) random walk with exponential noise.
- GMM (3 peaks) random walk with GMM noise.

For each combination, the Dirac approach was tested with varying numbers of particles, while the Gaussian Mixture approach varied the number of Gaussian components in the mixture, keeping the number of particles fixed at 256.

For benchmarking purposes, both the KF and PF were used as comparative methods. The KF serves as an optimal solution under Gaussian noise assumptions, while the PF was configured with the same number of particles as the Hybrid Filter to ensure a fair comparison.

The filters were evaluated over 50 independent runs, each consisting of 150 time steps. In every run, the vehicle's initial velocity was set to  $[0 \ 0]^\top$ , while its initial position was uniformly sampled within the square defined by the vertices at  $(-8, -8)$ ,  $(-8, -10)$ ,  $(-10, -10)$ , and  $(-10, -8)$ . During each run, the type of random walk and observation noise remained fixed, allowing for a controlled assessment of filter performance under specific noise conditions.

The performance of each filter was measured using two key metrics:

- Cumulative RMSE over the 150 time steps to evaluate estimation accuracy.
- Covariance matrix trace to assess the uncertainty associated with the estimates at each time step.

These simulations provide a rigorous framework for evaluating the Hybrid Filter's capabilities across a wide range of noise conditions. By comparing the results of the Hybrid Filter with those of the KF and PF, the robustness of the proposed approach in both Gaussian and non-Gaussian noise environments is thoroughly examined.

1) *Hybrid Filter via Dirac Approximation:* For the Dirac approach, the nine distinct scenarios were evaluated with the number of particles varying between 256, 625, and 1296, resulting in a total of 27 simulations. Here, we present and discuss one of the most significant cases.

The performance of the CF Filter tested with a Gaussian mixture walk and exponential noise using 1296 particles is illustrated in Figure 1. The results show a distinct shift in performance compared to the Gaussian noise setting. The first three plots depict the cumulative RMSE of each of the filters across all 50 runs (represented by gray lines). These plots also include the upper and lower bounds of the filters' performance,

represented by the Max and Min lines, respectively. The 75<sup>th</sup> percentile of the RMSE is depicted in green, offering further insights into the overall performance distribution. Lastly, to facilitate direct comparison, Figure 1(d) presents the average cumulative RMSE at each time step, averaged over the 50 runs.

The CF Filter demonstrates a marked improvement in performance, surpassing the KF, which no longer operates optimally due to the introduction of non-Gaussian, exponential noise. The adaptability of the CF Filter becomes particularly evident in this scenario, as its ability to manage non-Gaussian noise allows it to outperform the KF. The tight error bounds and reduced variability across the runs (as shown in Figure 1(a)) underscore the robustness of the CF Filter, which maintains consistent accuracy in the presence of a complex noise distribution.

In contrast, the KF, which is tailored for Gaussian noise, exhibits a decline in relative performance. This outcome aligns with theoretical expectations, given that the KF’s assumptions regarding noise characteristics are no longer valid in non-Gaussian settings.

The PF, while theoretically well-suited to handle non-Gaussian noise, displays weaker performance in this case. This may be attributed to the number of particles (1296), which appears insufficient to accurately capture the underlying noise characteristics and system dynamics. The greater variance and broader error bounds (as depicted in Figure 1(c)) suggest that the PF struggles to match the consistency and accuracy achieved by the CF Filter.

The overall comparison of the average cumulative RMSE in Figure 1(d) further highlights the superiority of the CF Filter in this setting. Its ability to outperform both the KF and PF demonstrates the efficacy of its Dirac approximation in efficiently handling non-Gaussian noise distributions. While the KF remains a strong performer under simpler noise conditions, it proves less competitive in this more challenging scenario, whereas the PF is hindered by the limitations inherent to particle count.

2) *Hybrid Filter via Gaussian Mixture Approximation:* The Gaussian mixture approach was evaluated across nine distinct scenarios, with the number of Gaussian components used in the approximation varying between 3 and 4, resulting in a total of 18 simulations. Here, we focus on presenting and analyzing the most significant cases.

The performance of the CF Filter using the Gaussian mixture approximation with 3 Gaussian components, tested under conditions involving both a Gaussian mixture walk and Gaussian mixture noise, is depicted in Figures 2(a), 2(b), 2(c), and 2(d). The objective of this analysis is to understand the relative performance of the CF Filter, compared to the KF and PF, over 50 simulation runs.

As illustrated in Figure 2(a), the CF Filter using the Gaussian mixture approximation with 3 components exhibits more erratic behavior compared to the Dirac approach discussed earlier. The cumulative RMSE reveals fluctuations across different runs, which is reflected in the larger spread between

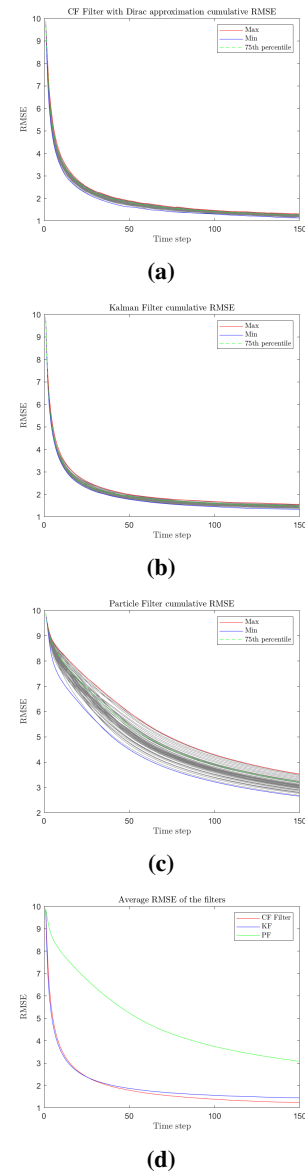


Fig. 1. Cumulative RMSE of the position of Gaussian Mixture walk with Exponential noise: a) CF Filter with 1296 particles; b) Kalman Filter; c) Particle Filter with 1296 particles; d) Average of the cumulative RMSE over the 50 runs.

maximum and minimum error bounds. Such variability can be attributed to challenges inherent in the Gaussian mixture approximation, particularly the non-convex nature of the optimization problem, which may converge to local minima, thus impeding optimal state representation. Furthermore, the number of Gaussian components used might be insufficient to fully capture the complexity of the state dynamics.

Nevertheless, despite these fluctuations, the CF Filter maintains strong average performance, as highlighted in Figure 2(d). The average cumulative RMSE of the CF Filter remains competitive, particularly when compared to the PF. This suggests that while individual runs might display some instability due to the Gaussian mixture approximation, the

filter remains generally effective in estimating the state over the 50 simulations.

In contrast, the KF, depicted in Figure 2(b), demonstrates a more stable performance. This result is expected, as the Gaussian mixture noise used in these simulations can be reasonably approximated by a zero-mean Gaussian distribution. Under these circumstances, the KF excels because its underlying assumptions about the noise model are largely satisfied, resulting in smoother convergence and lower overall RMSE compared to the CF Filter. This underscores the KF's inherent advantage in scenarios with Gaussian-like noise, even when the noise model presents some additional complexity, as seen in the Gaussian mixture case.

The PF, on the other hand, continues to underperform relative to both the CF Filter and KF. As shown in Figure 2(c), the PF exhibits slower convergence and a higher cumulative RMSE. The limited number of particles (256) appears to be a key factor limiting its performance, leading to a broader spread of error bounds and highlighting difficulties in effectively capturing the system dynamics under these conditions.

When comparing the average cumulative RMSE across all filters, as seen in Figure 2(d), the CF Filter, despite its fluctuations, still demonstrates competitiveness relative to the KF. This suggests that, while the Gaussian mixture approximation introduces additional challenges, the CF Filter retains the capability to provide effective state estimates on average.

Increasing the number of Gaussian components from 3 to 4, as depicted in Figures 3(a) and 3(b), leads to a notable improvement in the stability of the CF Filter. The spread between the maximum and minimum error bounds narrows, and the overall convergence becomes smoother compared to the 3-Gaussian scenario. This indicates that increasing the number of Gaussian components enhances the filter's ability to capture the underlying system and noise complexities, resulting in more consistent state estimates.

In terms of average cumulative RMSE, the CF Filter with 4 Gaussian components demonstrates a slight reduction in error compared to the 3-Gaussian case, as shown in Figure 3(b). This improvement narrows the performance gap between the CF Filter and the KF, highlighting the advantages of using a more refined Gaussian mixture approximation to better capture the underlying noise characteristics.

3) *Approaches Comparison:* The Dirac-based approach consistently demonstrates stable and reliable performance across all tested scenarios, exhibiting robustness and adaptability. It performs comparably to the KF in scenarios involving Gaussian noise and surpasses the KF in non-Gaussian conditions, such as those involving exponential noise. In contrast, the Gaussian mixture approach struggles with maintaining consistent performance in more complex noise environments.

The Dirac approach also benefits from computational efficiency due to the relative simplicity of the optimization problem involved in the prediction step. Specifically, the prediction step approximates the CF using a finite sum of Dirac distributions, thereby reducing the problem to a constrained linear least squares problem. This is significantly

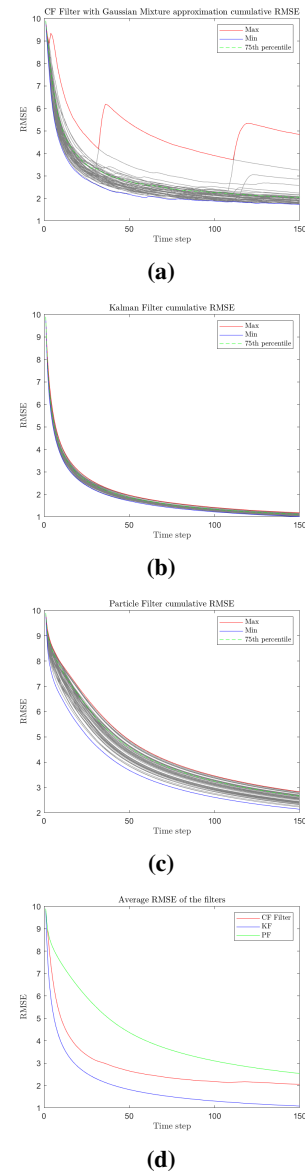


Fig. 2. Cumulative RMSE of the position of Gaussian Mixture w/ 2 peaks walk with Gaussian Mixture noise: a) CF Filter with 3 Gaussians; b) Kalman Filter; c) Particle Filter with 256 particles; d) Average of the cumulative RMSE over the 50 runs.

simpler than the Gaussian mixture approach, which involves optimization over both the weights and the parameters (means and covariances) of Gaussian components. The non-convex nature of this optimization adds computational complexity and increases the risk of convergence to local minima. As the number of Gaussian components increases, the computational burden grows correspondingly, making the Dirac approach more suitable in resource-constrained environments.

In summary, the Dirac approach provides consistent performance across diverse noise scenarios at a lower computational cost, making it a practical choice for real-world applications that require both accuracy and efficiency. Although the Gaussian mixture approach offers greater theoretical flexibility, it

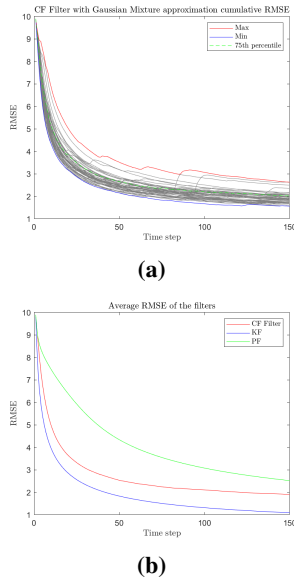


Fig. 3. Cumulative RMSE of the position of Gaussian Mixture w/ 2 peaks walk with Gaussian Mixture noise: a) CF Filter with 4 Gaussians; b) Average of the cumulative RMSE over the 50 runs.

may require enhanced optimization techniques or additional components to achieve the level of reliability observed with the Dirac method.

### III. MULTI-SENSOR STATE ESTIMATION

#### A. Related Work

This section provides a comprehensive overview of key methodologies and challenges associated with multi-sensor SE, emphasizing both Gaussian and non-Gaussian noise considerations. The discussion focuses on consensus-based sensor fusion and distributed Kalman filtering approaches, examining the significant contributions and advancements in the field.

1) *Distributed Sensor Fusion Based on Average Consensus*: The concept of average consensus is fundamental to many distributed sensor fusion algorithms, as it allows a network of sensors to iteratively converge to an average value of their local measurements through localized interactions. The scheme proposed in [10] leverages this principle to compute the maximum-likelihood estimate of parameters under Gaussian noise, thereby providing a robust solution suitable for dynamic networks with intermittent communication links. This method enables each sensor to iteratively compute its local estimate, which eventually converges to the global solution, enhancing the scalability and resilience of the distributed system.

2) *Distributed Kalman Filtering*: The Distributed Kalman Filter (DKF) integrates consensus mechanisms with the traditional KF to ensure consistent state estimates across a distributed sensor network. The Kalman-Consensus Filter, introduced in [11], combines the state estimation capabilities of the KF with a high-pass consensus filter to achieve robust distributed state estimation. This hybrid approach enhances the cohesiveness among individual sensors' state estimates, resulting in accurate and reliable SE even in the presence

of dynamic environments, communication constraints, and inconsistencies among sensor readings.

3) *Distributed State Estimation Based on Kullback-Leibler Average*: The method presented in [12] addresses the complexities introduced by non-Gaussian noise within sensor networks by employing a Kullback-Leibler (KL) divergence-based consensus algorithm. This approach allows each sensor node to reach a consensus on the posterior probability density function (PDF) by minimizing the cumulative KL divergences across the network. The resulting KL average PDF offers a robust framework for distributed SE in non-Gaussian environments, ensuring that each sensor's estimate converges to a consistent representation of the system's state, despite the non-Gaussian nature of the disturbances.

#### B. Multi-Sensor Problem

The multi-sensor problem extends the SE challenge to encompass a network of sensors, each contributing observations to collaboratively estimate the state of a dynamic system. The primary objective is to utilize collective data from multiple sensors to derive a more robust and accurate estimate of the system's state compared to using a single sensor.

The problem is formulated as an interconnected network of sensors, modeled by a graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where  $\mathcal{V} = \{v_1, v_2, \dots, v_m\}$  represents the set of  $m$  sensors, and  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  represents the communication links between these sensors.

- **State Equation**: The state equation remains identical to equation 3 of the single-sensor problem, describing the evolution of the system state over time.
- **Observation Equation for Sensor  $i$** : For sensor  $i$ , the observation equation is given below, where  $z_t^{(i)} \in \mathbb{R}^{p_i}$  denotes the observation vector from sensor  $i$  at time  $t$ ,  $H^{(i)} \in \mathbb{R}^{p_i \times n}$  is the observation matrix, and  $v_t^{(i)} \sim \mathcal{D}(\theta^{(i)})$  represents the measurement noise vector for sensor  $i$ .

$$z_t^{(i)} = H^{(i)} x_t + v_t^{(i)} \quad (12)$$

The multi-sensor problem builds on assumptions similar to those in the single-sensor case, including system linearity, time invariance of system matrices, known parameters, and independent noise vectors. However, the multi-sensor context also introduces considerations regarding the network topology, where the sensor network is represented as a connected graph  $\mathcal{G}$ , ensuring the existence of communication pathways between all sensor pairs.

Two key observability assumptions are considered in this multi-sensor context:

- **Individual Sensor Observability**: Each sensor's matrix pair  $(A, H^{(i)})$  must satisfy the condition of observability, ensuring that the individual sensor can independently contribute to state estimation.
- **Network Collective Observability**: The network as a whole should function as a collective observer, meaning that the entire sensor network, when considered collectively, is capable of observing the system's state. This

condition is less restrictive than individual sensor observability and is often preferred in practical applications.

### C. Wasserstein Barycenter for Sensor Fusion

The WB is a mathematical concept derived from optimal transport theory, providing an advanced framework for aggregating multiple probability distributions into a single, representative distribution. Analogous to the classical mean in Euclidean space, the WB can be regarded as the 'mean' of probability distributions in the space of probability measures, accounting for the underlying geometric relationships among these distributions.

Formally, the WB aims to identify a central distribution that minimizes divergence from a given set of distributions in the Wasserstein sense. This objective is akin to computing an average, but instead of averaging scalar values, the WB averages entire distributions, explicitly considering the 'distance' between them.

Consider a set of probability distributions  $\mu_i$ ,  $i=1, \dots, N$ . The goal is to determine a barycenter distribution  $\nu$  that minimizes the sum of Wasserstein distances to each  $\mu_i$ , expressed as:

$$\nu^* = \arg \min_{\nu \in \mathcal{P}(\mathbb{R}^n)} \sum_{i=1}^N \lambda_i \Phi_2^2(\mu_i, \nu), \quad (13)$$

where  $\mathcal{P}(\mathbb{R}^n)$  represents the space of probability measures on  $\mathbb{R}^n$ ,  $\Phi_2(\mu_i, \nu)$  denotes the Wasserstein-2 distance between distributions  $\mu_i$  and  $\nu$ , and  $\lambda_i$  are the weights associated with each distribution, satisfying  $\sum_{i=1}^N \lambda_i = 1$  and  $\lambda_i \geq 0$ . The Wasserstein-2 distance is defined as:

$$\Phi_2^2(\mu_i, \nu) = \inf_{\gamma \in \Gamma(\mu_i, \nu)} \int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|_2^2 d\gamma(x, y), \quad (14)$$

where  $\Gamma(\mu_i, \nu)$  denotes the set of all probability measures on  $\mathbb{R}^n \times \mathbb{R}^n$  with marginals  $\mu_i$  and  $\nu$ , and  $\|x - y\|_2$  is the Euclidean norm in  $\mathbb{R}^n$ .

The need for the WB in sensor fusion is motivated by the inherent complexity involved in fusing data from multiple sensors, each characterized by distinct probability distributions. Traditional fusion methods often assume homogeneity among these distributions, leading to suboptimal performance, particularly when the sensor data follows diverse, non-Gaussian distributions.

The WB addresses this challenge by providing a principled mechanism to merge disparate distributions into a single, representative distribution. This approach is particularly advantageous for sensor data exhibiting varied noise characteristics, as it inherently captures the underlying geometry of the distributions rather than relying on simple averaging.

The core idea behind employing the WB in sensor fusion is to combine information from multiple sensors in a way that minimizes the overall divergence between the fused result and the individual distributions. Instead of averaging scalar values, the WB framework averages full distributions, ensuring that the final estimate respects the geometric differences intrinsic to the sensor data.

By leveraging the WB for sensor fusion, data from multiple sources can be effectively combined while preserving the essential characteristics of each distribution. This ensures that the fused result is not merely an aggregate, but rather a well-balanced representation that minimizes discrepancies among the original distributions. The optimization involved in computing the WB allows for flexibility in weighting each sensor's contribution, ensuring that the final distribution appropriately reflects the relative importance of each sensor's data.

To illustrate the application of the WB in sensor fusion, consider a scenario involving  $N$  state estimates, each generated by a distinct sensor. These sensors may be located at different positions or subjected to varying types of noise, leading to heterogeneous probability distributions for their respective state estimates. In such cases, directly combining these estimates using traditional methods may result in inaccuracies, as they typically fail to account for the unique characteristics of each distribution.

Instead, each sensor can communicate its local probability distribution or state estimate to a central computational unit. This unit, equipped with the WB framework, can fuse the received distributions by calculating a central distribution that best represents the combined information from all sensors. The resulting PDF, derived through the WB, minimizes the Wasserstein distance between the central estimate and each individual sensor distribution, ensuring that the fused result balances contributions from all sensors.

After obtaining the fused PDF, the central unit can distribute this updated distribution back to each sensor, enabling each sensor to refine its own state estimate based on the collective information. In this way, every sensor benefits from the combined data, leading to improved accuracy and coherence across the network.

Beyond centralized implementation, the WB can also be computed in a distributed manner [13], enhancing scalability and robustness, particularly in large sensor networks. While a fully centralized framework can be effective, it may become a bottleneck in large-scale systems due to the reliance on a single processing unit for data aggregation and dissemination.

Conversely, distributed computation of the WB allows sensors, or groups of sensors, to locally compute partial updates to the WB, which are subsequently shared with neighboring sensors or other segments of the network. This reduces communication overhead and distributes the computational burden across the entire network. Such a decentralized approach not only scales efficiently as the number of sensors increases but also enhances the resilience of the network to single points of failure.

In a distributed setting, sensors can collaborate by exchanging intermediate estimates, progressively refining the fused PDF without relying on a central authority. Over several iterations, the network collectively converges to a global WB that encapsulates the aggregate information from all sensors. This characteristic makes the WB framework particularly well-suited for real-time applications in dynamic environments, where scalability, robustness, and computational efficiency are

of critical importance.

#### D. Experimental Results

This section presents the results of a simulation designed to address the multi-sensor problem using the WB framework. The simulation focuses on a scenario where different sensors are influenced by distinct types of noise:

- Sensor 1 assumes GMM noise and employs a Gaussian mixture filter.
- Sensor 2 assumes Gaussian noise and runs a KF.
- Sensor 3 assumes uniform noise and utilizes an interval observer.

The simulation involves a network of three sensors with mutual communication. Each sensor runs its local filtering process and transmits outputs to a centralized unit, which then fuses all information using the WB approach, i.e., the algorithm operates in a centralized manner.

The dynamics of the system were modeled using a discrete double integrator with a sampling time of  $T_s = 0.1$  s. The model equations are analogous to equations 3 and 12, with the exception that the disturbance term  $d_t$  is not considered. The observation matrices,  $H^{(i)}$ , are assumed to be identical across all sensors:

$$H^{(i)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \forall i \in \mathcal{V} \quad (15)$$

The task is to estimate the position of a vehicle following a figure-eight trajectory. The initial position is assumed to be  $\begin{bmatrix} 0 & 0 \end{bmatrix}^\top$  with some uncertainty, allowing for the estimation of a PDF for the initial state.

The WB was employed to fuse the local filter outputs for a final estimate. For comparison purposes, three additional algorithms were implemented:

- A centralized KF.
- Average consensus as discussed in Subsection III-A1.
- Averaging the position estimates from each sensor.

The scenario was simulated 30 times, each run consisting of 250 iterations. Initial noise distributions were randomized for each run. Performance was evaluated based on the following metrics:

- RMSE, averaged over 30 runs.
- The  $\ell^2$ -norm of the position error vector,  $e_t = x_t - \hat{x}_t$ , for a randomly selected run.
- The trace of the covariance matrix, indicating the estimation uncertainty for a randomly selected run.

The results of the simulation are presented in Figure 4, with Figures 4(a) and 4(b) showing the y-axis in log scale. The performance of the WB fusion method highlights its adaptability to complex environments. The heterogeneous noise conditions across the sensor network introduce challenges that traditional methods, such as the centralized KF, are less capable of managing effectively. This is evident across all performance metrics, where WB fusion consistently demonstrates superior results.

In Figure 4(a), WB fusion shows a clear advantage in terms of position error, particularly during the initial iterations. Although the centralized KF maintains reasonable accuracy, it struggles with noise variability across sensors, resulting in higher errors. In contrast, the WB effectively aggregates the sensor data in the Wasserstein sense, yielding a consistently lower error profile.

The average method of combining filter outputs achieves competitive results in terms of position error; however, it has a significant limitation: it only provides a single estimate vector representing an average of the individual sensor estimates. It does not convey information on the uncertainty or the full PDF of the estimate, rendering it less practical in scenarios where understanding uncertainty is critical.

As shown in Figure 4(b), the covariance matrix trace for the WB fusion approach is higher than that of the centralized KF, indicating greater uncertainty. The WB algorithm operates by aggregating probability distributions from multiple sensors, and while it provides an optimal central distribution in the Wasserstein sense, it does not explicitly minimize estimation uncertainty as the KF does. The KF is specifically designed to iteratively reduce uncertainty by incorporating measurements and updating estimates using an optimal gain (Kalman Gain), which results in a systematic reduction in the covariance matrix trace as confidence in the state estimates improves over time.

The results in Table I and Figure 4(c) underscore the robustness and superior performance of the WB fusion approach. The WB consistently achieves lower RMSE compared to the centralized KF, consensus algorithm, and the averaging method. This lower RMSE and reduced variability across simulations indicate that the WB method is better equipped to handle diverse noise profiles.

The reduced variability of WB fusion across multiple simulations, as evidenced by the standard deviation in Table I, aligns with the stability observed in Figure 4(c). The WB method maintains strong performance despite different noise realizations, whereas both the consensus algorithm and the centralized KF exhibit greater fluctuations, highlighting their sensitivity to noise type.

Although the average method delivers competitive RMSE results, its inability to provide insights into uncertainty remains a substantial drawback.

In summary, WB fusion proves to be highly effective in managing both non-Gaussian and variable noise conditions without necessitating specific assumptions about noise distributions. This flexibility makes it a versatile and powerful tool for sensor fusion in complex, heterogeneous environments.

Algorithm	Mean RMSE [m]	RMSE Standard Deviation [m]
WB	1.1473	0.2802
Filters average	1.3462	0.4304
Centralized KF	1.9733	0.8658
Consensus average	2.0569	0.8505

TABLE I  
RMSE.

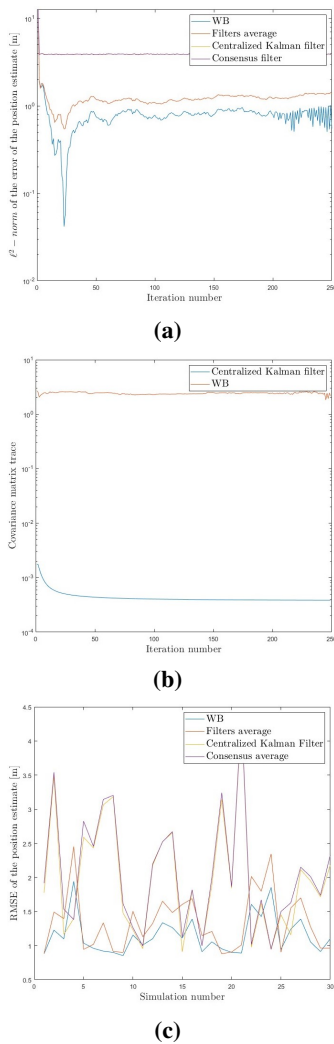


Fig. 4. Different noises simulation: a)  $\ell^2$ -norm of the position error; b) Covariance matrix trace of the position uncertainty; c) RMSE over multiple simulations.

#### IV. CONCLUSION

##### A. Summary

This work focused on SE in linear systems for single-sensor and multi-sensor scenarios under Gaussian and non-Gaussian noise. The primary goal was to enhance estimation accuracy where traditional methods like the KF fall short. Advanced filtering techniques, including CF-based methods, were introduced to address complex noise distributions.

For single-sensor systems, the limitations of the KF and PF under non-Gaussian noise were evaluated. A hybrid filtering approach combining Bayesian and CF-based methods was developed, demonstrating robustness in challenging noise conditions.

In multi-sensor networks, SE techniques using the WB for sensor fusion were proposed. These methods improved both accuracy and robustness, ensuring reliable performance even with diverse sensors and non-Gaussian noise.

##### B. Future Work

Despite the progress made, several challenges remain:

- **extbfHybrid Filter Refinement:** Future research should refine the Gaussian Mixture Approximation in the hybrid filter to enhance accuracy and computational efficiency, focusing on reducing Gaussian component complexity.
- **extbfDistributed WB Fusion Algorithm:** The WB-based fusion algorithm should be tested in distributed settings to improve scalability and robustness, especially in geographically dispersed networks, addressing issues like consensus and communication delays.

These directions aim to advance state estimation techniques, addressing challenges with non-Gaussian noise, sensor diversity, and distributed settings.

#### REFERENCES

- [1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME - Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.
- [2] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," in *Proc. SPIE 3068, Signal Processing, Sensor Fusion, and Target Recognition VI*, Orlando, FL, USA, Jul. 1997, p. 182. [Online]. Available: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.280797>. doi: 10.1117/12.280797.
- [3] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium*, Oct. 2000, pp. 153–158. doi: 10.1109/ASSPCC.2000.882463.
- [4] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002. doi: 10.1109/78.978374.
- [5] A. G. Wills, J. Hendriks, C. Renton, and B. Ninness, "A Bayesian filtering algorithm for Gaussian mixture models," *arXiv preprint arXiv:1705.05495*, Jun. 2023. [Online]. Available: <http://arxiv.org/abs/1705.05495>.
- [6] A. R. Runnalls, "Kullback-Leibler approach to Gaussian mixture reduction," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 989–999, Jul. 2007. doi: 10.1109/TAES.2007.4383588.
- [7] M. Idan and J. L. Speyer, "An estimation approach for linear stochastic systems based on characteristic functions," *Automatica*, vol. 78, pp. 153–162, Apr. 2017. doi: 10.1016/j.automatica.2016.12.038.
- [8] A. P. Vinod, B. Homchaudhuri, and M. M. K. Oishi, "Forward stochastic reachability analysis for uncontrolled linear systems using Fourier transforms," in *Proceedings of the 20th International Conference on Hybrid Systems: Computation and Control*, Apr. 2017, pp. 35–44. doi: 10.1145/3049797.3049818.
- [9] H. Zhang, "Non-linear Bayesian filtering by convolution method using fast Fourier transform," *Fusion*, vol. 1, no. 1, pp. 1–6, Jul. 2011.
- [10] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks*, Apr. 2005, pp. 63–70. doi: 10.1109/IPSN.2005.1440896.
- [11] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proceedings of the 46th IEEE Conference on Decision and Control*, Dec. 2007, pp. 5492–5498. doi: 10.1109/CDC.2007.4434303.
- [12] G. Battistelli and L. Chisci, "Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability," *Automatica*, vol. 50, no. 3, pp. 707–718, Mar. 2014. doi: 10.1016/j.automatica.2013.11.042.
- [13] C. A. Uribe, D. Dvinskikh, P. Dvurechensky, A. Gasnikov, and A. Nedić, "Distributed computation of Wasserstein barycenters over networks," *arXiv preprint arXiv:1803.02933*, Sep. 2018. [Online]. Available: <http://arxiv.org/abs/1803.02933>.