A Hyperparameter-Based Approach for Consensus Under Uncertainties

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
A Hyperparameter-Based Approach for Consensus Under Uncertainties

Cameron S. R. Fraser, Luca F. Bertuccelli, Han-Lim Choi, and Jonathan P. How

Abstract—This paper addresses the problem of information consensus in a team of networked agents with uncertain local estimates described by parameterized distributions in the exponential family. In particular, the method utilizes the concepts of pseudo-measurements and conjugacy of probability distributions to achieve a steady-state estimate consistent with a Bayesian fusion of each agent’s local knowledge, without requiring complex channel filters or being limited to normally distributed uncertainties. It is shown that this algorithm, termed hyperparameter consensus, is adaptable to any local uncertainty distribution in the exponential family, and will converge to a Bayesian fusion of local estimates over arbitrary communication networks so long as they are known and strongly connected.

I. INTRODUCTION

Consensus algorithms over communication networks have received much attention due to their many applications in distributed decision making for multi-agent flocking to rendezvous [1]–[9]. These consensus algorithms are typically computationally inexpensive even over large, complex networks, and guarantee convergence of the situational awareness over many different dynamic network topologies [6], [8]–[11]. However, most of these approaches (e.g., Refs. [1], [3], [4]) have assumed that (a) there is no uncertainty in each agent’s local estimate of the quantity of interest, and (b) there is no true value to which the agents are attempting to agree (i.e., the agents having agreed is the only necessary result).

In contrast, Kalman consensus approaches [12], [13] allow for uncertainty that can be modeled with a Gaussian distribution. These algorithms give rise to a consensus result that is influenced more heavily by agents with smaller covariance (therefore higher certainty) in their estimates. Unfortunately, many stochastic systems do not satisfy the assumption of a linear system with Gaussian noises, and applying these Kalman filter-based consensus methods to other distributions can produce a steady-state estimate that is biased away from the Bayesian estimate obtained by using the true form of the distribution [14].

While most consensus algorithms are designed to agree on subjective parameters, Bayesian decentralized data and sensor fusion methods are designed to determine the best combined Bayesian parameter estimate given a set of observations [15]–[18]. Unfortunately, as noted in [16], [17], the channel filters required to handle common or repeated information in messages between neighboring nodes are not easily defined for network topologies other than the simplest fully connected and tree networks, limiting traditional decentralized data fusion methods from being easily applied over arbitrary, large-scale networks, even if the topology were known.

Recent results have shown that a combination of traditional consensus-based communication protocols with data fusion information updates can achieve scalable, representative information fusion results without requiring complex channel filters or specific network topologies [14], [19]–[21]. In particular, [19], [20] utilized dynamic-average consensus filters to achieve an approximate distributed Kalman filter, while [21] implemented a linear consensus protocol on the parameters of the information form of the Kalman filter. Both methods permitted the agents to execute a Bayesian fusion of normally-distributed random variables. However, since the methods in [19]–[21] are derived specifically for normally-distributed uncertainties, they can produce biased results if the local distributions are non-Gaussian.

Ref. [14] derived a method to permit a consensus on the Bayesian aggregation of multiple Dirichlet distributions, and demonstrated its use in distributed estimation of transition probabilities in a multi-agent learning problem. This paper utilizes the same concepts of pseudo-measurement updates and conjugacy of probability distributions [22] to extend the Bayesian fusion results obtained in [14] to a more general class of distributions.

We start by introducing some required background theory in Section II. The main contribution of this paper, termed hyperparameter consensus, is introduced in Section III. Here, we show it will permit convergence to a Bayesian fusion of local distributions in the exponential family over known, strongly connected but otherwise arbitrarily large and complex networks and without the need for channel filters. Section IV then presents an example demonstrating the application of the method to the case of local gamma distributions.

II. BACKGROUND

A. Probability Theory

This section introduces some of the terminology that will be used in this paper (see [22] for more information). The conditional distribution of a random variable $X$ parameterized by $Y$ is given by $f_{X|Y}(X = x|Y = y)$, and is often denoted simply as $f_{X|Y}(x|y)$ or $f_{X|Y}$. The random variable $\Theta$ denotes the parameter of interest that the agents are attempting to agree upon. Parameter estimates

C. S. R. Fraser has his Masters degree from the Department of Aeronautics and Astronautics, MIT fraserc@mit.edu

L. F. Bertuccelli is a Postdoctoral Associate in the Humans and Automation Lab in the Department of Aeronautics and Astronautics, MIT lucab@mit.edu

H.-L. Choi is an Assistant Professor in the Division of Aerospace Engineering, KAIST hanlimc@kaist.ac.kr

J. P. How is the Richard Cockburn Maclaurin Professor of Aeronautics and Astronautics, MIT, jhow@mit.edu

978-1-4244-7427-1/10/$26.00 ©2010 AACC 3192
are realizations of this random variable and are denoted by \( \theta \). \( Z \) is the random variable of which each measurement, \( z \), is a realization. It is defined by a stationary stochastic process1 governed by the true value of the parameter of interest. Finally, \( \Omega \) represents the set of hyperparameters, \( \omega \), a (possibly multivariate) parameterization of a prior (or posterior) distribution on \( \Theta \).

The Bayesian measurement update of a prior distribution on \( \theta \) is shown in (1), where \( f_{Z|\Theta}(\theta) \) is the measurement model, also known as the likelihood function on \( \theta \) for a given \( z \); \( f_{\Theta|\Omega}(\Theta) \) is the prior distribution on \( \Theta \); \( f_{\Theta|Z,\Omega}(\theta|z,\omega) \) is the posterior distribution; and the denominator is a constant that ensures the posterior integrates to unity.

\[
f_{\Theta|Z,\Omega}(\theta|z,\omega) = \frac{f_{Z|\Theta}(z|\theta) f_{\Theta|\Omega}(\theta|\omega)}{\int f_{Z,\Theta|\Omega}(z,\theta|\omega) d\theta}
\]  
(1)

Conjugate priors are well defined distributions parameterized by a set of hyperparameters \( \omega \), and are often used so that the Bayesian update in (1) can be solved in closed form, thereby avoiding difficulties in the numerical integration of the denominator. These priors are considered conjugate to the likelihood function, and produce posterior distributions of the same family as the prior distribution such that the required normalizing constant is inherently known. Using a conjugate prior, the Bayesian update in (1) can be replaced with a simple, closed-form update of the hyperparameters,

\[
\omega \leftarrow h(z, f_{Z|\Theta}) + \omega,
\]
(2)

where \( z \) is realized from an underlying distribution that we are modeling using the likelihood function in (1); and \( h(\cdot) \) is a possibly nonlinear function of the observed measurement and the measurement model and results in a positive (or positive-definite) change in hyperparameters. For example, the hyperparameters of the arrival rate, \( \lambda \), for a Poisson likelihood (with known period, \( T \)) are the parameters of the conjugate prior gamma distribution, \( \alpha \) and \( \beta \). These are updated by adding the number of events observed, \( x \), and the period of time in which they were observed, \( T \), to \( \alpha \) and \( \beta \), respectively, as in (19), to achieve the new hyperparameters that define the posterior gamma distribution on \( \lambda \). In some cases, traditional non-additive hyperparameter updates, such as those used in the Kalman filter, can be transformed into additive updates by using, for example, the inverse covariance and information state parameters of an equivalent information filter.

This conjugacy property is required for the hyperparameter consensus method, and, therefore, the method is applicable to any parameterized distributions which are conjugate to a likelihood function, such as those in the exponential family.

B. Linear Consensus

This paper utilizes linear consensus results as outlined in Refs. [8], [9], and the reader is referred there for more information. The linear consensus protocol used in this paper is assumed to operate over a network defined by a strongly connected directed graph \( G \), where each agent \( i \in \{1, 2, \ldots, N\} \) is represented by a node, there is a directed edge from agent \( i \) to \( j \) if and only if agent \( i \) can transmit to \( j \), and each agent is assumed to be able to talk to itself. Let \( \mathcal{N}_i \) denote the set of agents that can transmit to \( i \), and let the adjacency matrix, \( A \), for the graph be defined by entries \( a_{ij} \geq 0 \), which may be non-zero if there exists an edge from \( j \) to \( i \) and are zero otherwise.

Though the main results of this paper can be achieved with a range of consensus protocols, we will limit discussion to the discrete time consensus protocol shown in (3) as applied to a scalar parameter, \( \xi \), at consensus iteration \( k \).

\[
\begin{align*}
\xi_i[k+1] &= \xi_i[k] + \epsilon \sum_{j \in \mathcal{N}_i} (\xi_j[k] - \xi_i[k]) = \sum_{j=1}^{N} a_{ij} \xi_j[k] \\
\xi[k+1] &= A \xi[k] = A^{k+1} \xi[0],
\end{align*}
\]  
(3)

For (3) to converge, the weighting constant, \( \epsilon \), is subject to \( \epsilon \in (0, \frac{1}{\max_{i,j} |A_{ij}|}) \) [9]. The result of this linear consensus protocol is found in the limit as

\[
\lim_{k \to \infty} \xi_i[k+1] = \nu^T \xi[0] = \sum_{j=1}^{N} \nu_j \xi_j[0],
\]

where \( \nu \) is the consensus eigenvector, the normalized left-eigenvector of the adjacency matrix corresponding to its 1-eigenvalue, solving \( \lim_{k \to \infty} A^k = \nu^T A = \nu^T \) [8].

In a balanced network where \( \sum_{j=1}^{N} a_{ij} = \sum_{j=1}^{N} a_{ji}, \forall i \in \{1, 2, \ldots, N\} \), all \( \nu_i = 1/N \) and the resulting consensus value is an arithmetic average of the initial conditions.

Note that if the underlying connectivity graph is undirected then the network is inherently balanced, regardless of its specific topology, and the average consensus result is always achieved. However, if the network is unbalanced then the weights are not all equal, and the agents must weight their initial conditions as \( \xi_i[0] = \frac{\xi_i[0]}{\nu_i} \), at consensus iteration \( k \) [8].

Of primary importance is that a sum-consensus is achieved when each agent divides its initial condition only by the agent’s associated consensus weight, \( \xi_i[0] = \frac{\xi_i[0]}{\nu_i} \), leading to

\[
\lim_{k \to \infty} \xi_i[k+1] = \nu_i \sum_{j=1}^{N} \xi_j[0] = \sum_{j=1}^{N} \xi_j[0].
\]  
(4)

Finally, the sum-consensus result can also be achieved with a dynamic network topology if each graph and switching time is known, all graphs are strongly connected, and there exist a finite number of switches. This can be achieved if each agent multiplies their local hyperparameters by \( \nu_{new}/\nu_{old} \) when a network switch occurs, where \( \nu_{old} \) and \( \nu_{new} \) are the consensus eigenvectors of the communication graphs before and after the switch, respectively. The proof can be found in Proposition 2.3.1 in [23].
III. HYPERPARAMETER CONSENSUS METHOD

This section derives the main result of the paper, the hyperparameter consensus method. First, the desired result of the consensus protocol is derived as a Bayesian fusion of each agent’s local estimate. The hyperparameter consensus method is then introduced as a means to achieve this Bayesian result, and some key results are highlighted.

A. Bayesian Fused Estimate

Consider $N$ agents attempting to combine their uncertain estimates of $\Theta$, where each agent maintains a local distribution on $\Theta$ of the same form (i.e. all Dirichlet distributions) but parameterized by local hyperparameters, $\omega_i$. We address shared prior information first then proceed to the derivation:

1) Common Information: Globally shared information is allowed so long as the information is common to all agents, such as a common initial prior on $\Theta$. The hyperparameters are updated in a linear fashion and completely define each agent’s knowledge measurement. Since the hyperparameters are updated in a shared prior information first then proceed to the derivation:

The basic concept of the proof is to note that the additive hyperparameter updates imply that the information unique to each agent can be described simply by $\Delta \omega_i = \omega_i - \omega^-$, and that the fused distribution can therefore be rewritten as:

$$f_{\Theta|\Omega}((\theta_1, \ldots, \theta_N)|\omega^-) = f_{\Theta|\Omega}((\theta_1, \ldots, \theta_N)|\omega^-) = \frac{1}{\prod_{i=1}^N f_{\Theta|\Omega}(\Delta \omega_i | \omega^-) f_{\Theta|\Omega}(\theta|\omega^-) J.}$$

The $f_{\Theta|\Omega}$ terms in the product represent what we call pseudo-likelihood functions and can be shown to be conjugate to the distribution on $\theta$, $f_{\Theta|\Omega}$, implying that the fused distribution is of the same form as the local distributions and that it is uniquely defined by a set of hyperparameters. The pseudo-likelihoods are functionally equivalent to the likelihood in (1), but lead to an update of the prior hyperparameters with each agent’s unique hyperparameters:

$$\omega_{\text{posterior}} = \Delta \omega_i + \omega_{\text{prior}}. \quad (7)$$

Thus, the last line in (6) can be considered akin to $N$ of these pseudo-measurement updates, and the resulting distribution can therefore be defined by the set of fused hyperparameters found through iteratively applying (7):

$$\omega_{\text{fused}} = \omega^- + \sum_{i=1}^N \Delta \omega_i \quad (8)$$

Using these fused hyperparameters, the fused distribution in (5) follows immediately.

In addition to this pure agreement result, we can also describe the fused distribution that arises when measurements are made while the agents are running a consensus protocol. In particular, each agent is assumed to take an independent measurement drawn from the same underlying stationary stochastic process, $f_{Z|\Theta}$ governed by the true parameter.

**Proposition 3.2 (Fused Distribution with Measurements):** Consider the case when each agent, $i$, takes a measurement at each consensus iteration, $k$, defined as $z_i[k]$, for $k \leq K < \infty$. This set of independent measurements, $z_i[0], \ldots, z_i[N], \ldots, z_i[K], \ldots, z_N[K]$, gives rise to the fused distribution at iteration $K$ defined by hyperparameters $\omega_{\text{fused}}[K]$, given by:

$$\omega_{\text{fused}}[K] = \omega_{\text{fused}}[0] + \sum_{i=1}^N \sum_{j=1}^K h(z_i[j], f_{Z|\Theta}), \quad (9)$$

where $\omega_{\text{fused}}[0]$ is the fused result ignoring any measurements and is equivalent to (8).

**Proof:** Proposition 3.1 states that the fused estimate without measurements is of the same form as the local distributions and is uniquely defined by a set of hyperparameters, $\omega_{\text{fused}}$. When a new measurement, $z_i$, is taken, this fused result should be updated as if the measurement were applied to it directly. Since the fused distribution is of the same form as the local distribution, and therefore conjugate to the measurement model, an equivalent hyperparameter update can be used:

$$\omega_{\text{fused}} \leftarrow \omega_{\text{fused}} + h(z, f_{Z|\Theta}) \quad (10)$$

With repeated independent measurements by each agent, $z_i[0], \ldots, z_i[N], \ldots, z_i[K], \ldots, z_N[K]$, the update in (10) becomes recursive, leading to the fused hyperparameter result:

$$\omega_{\text{fused}}[K] \leftarrow \omega_{\text{fused}}[0] + \sum_{i=1}^N \sum_{j=1}^K h(z_i[j], f_{Z|\Theta}).$$
This, in turn, completely describes the fused distribution and the proof is complete.

Remark: These results show that the the fused Bayesian estimate is of the form of the local uncertainties and is defined by hyperparameters equivalent to the sum of each agent’s independent local hyperparameters (including those corresponding to local measurements) and the shared hyperparameters. If there is no shared information, then the shared hyperparameters are set to zero. If a given measurement is shared across all agents, such that $z_i[\kappa] = z_j[\kappa] = z^*[\kappa] \forall i, j$ at some iteration $\kappa$, then

$$\omega_{fused}[K] \leftarrow \omega_{fused}[0] + \sum_{i=1}^{N} \sum_{j=1, j \neq \kappa}^{K} h(z_i[j], f_{Z[\theta]}) + h(z^*[\kappa], f_{Z[\theta]}).$$

The case of multiple shared measurements and multiple measurements per iteration per agent follow by extension.

B. Hyperparameter Consensus Method

Having defined the fused distribution that we wish to agree upon, we now turn to the derivation of the hyperparameter consensus method as a means to achieve it. The basic premise behind the method is to use existing linear consensus results as highlighted in Section II-B to achieve a sum-consensus to the hyperparameters representative of the Bayesian fused distribution. In general, the hyperparameter consensus method consists of three main actions:

- **Initialization:** At the beginning of a consensus, each agent must initialize its hyperparameters according to (11) so that common information is accounted for and the proper consensus result is achieved.

$$\omega_i[0] \equiv \omega^- + \frac{\omega_i - \omega^-}{\nu_i} = \omega^- + \frac{\Delta \omega_i}{\nu_i} \forall i,$$  

(11)

The resulting $\omega_i[0]$ serves as the initial information used in the consensus protocol.

- **Consensus Protocol:** After initialization, the agents all run the linear consensus protocol in (12) using the initialized values at iteration $k = 0$.

$$\omega_i[k+1] = \omega_i[k] + \epsilon \sum_{j \in N_i} (\omega_j[k] - \omega_i[k]) = \sum_{j=1}^{N} a_{ij} \omega_j[k],$$

(12)

- **Measurement Update:** If, at some iteration $\kappa$, agent $i$ takes an independent measurement, $z_i[\kappa]$, of the underlying process $f_{Z[\theta]}$, it must then update its local hyperparameters according to (13).

$$\omega_i[\kappa] \leftarrow \omega_i[\kappa] + \frac{h(z_i[\kappa], f_{Z[\theta]})}{\nu_i}.$$  

(13)

If the measurement is shared among all agents, such that $z_i[\kappa] = z^*[\kappa] \forall i$, then each agent must update their hyperparameters as in (14).

$$\omega_i[\kappa] \leftarrow \omega_i[\kappa] + h(z^*[\kappa], f_{Z[\theta]}).$$

(14)

Together with some mild assumptions on the network and consistency of representation, these three actions guarantee convergence to the Bayesian fused hyperparameters, the proof of which is given in Theorem 3.3. From these hyperparameters, each agent then re-constructs its probability distribution on $\Theta$ and uses an agreed-upon loss function (such as the Minimum Mean-Squared Error, MMSE, or others - see [22]) to evaluate the posterior distribution and select an appropriate realization of $\Theta$. Thus, the agents are able to implicitly come to agreement on a best estimate of $\theta$ based on the Bayesian fusion of their local uncertainty distributions.

**Theorem 3.3 (Convergence with Measurements):** A group of $N$ agents will come to an asymptotic agreement on the Bayesian fused distribution on an uncertain parameter, $\theta$, in the presence of global common information and a finite number of concurrent measurements, by running the consensus algorithm defined by (11) through (14), with the following properties holding true:

1. The connectivity graph $G$ is time-invariant, strongly connected and $\nu$ is known$^2$.
2. All agents use the consensus protocol as defined in (12) and initialized in (11), with $\epsilon \in (0, \frac{1}{\max_{i,N}[\nu_i]})$.
3. The agents’ maintain uncertain local estimates of $\theta$ through $f_{\theta[i]}$, the conjugate distribution to $f_{Z[\theta]}$.
4. There may exist some shared information among all agents in the network, denoted by $\omega^-$, conditioned upon which each agent’s initial hyperparameters are independent of all other agent’s hyperparameters.
5. The agents have decided a priori upon the form of the distribution to use and upon what loss function to implement to determine the best estimate of $\theta$ and
6. Each agent takes up to one measurement per consensus iteration, for each iteration up to some finite iteration $K < \infty$, and each measurement is independent of the measurements of other agents$^4$.

For a finite number of independent measurements occurring up to time $K < \infty$, the fused hyperparameters at iteration $k = K$ are given by (15),

$$\omega_{fused}[k] = \sum_{i=1}^{N} \left( \Delta \omega_i + \sum_{j=1}^{K} h(z_i[j], f_{Z[\theta]}) \right) + \omega^-,$$  

(15)

**Proof:** For the sake of brevity, the majority of this proof is omitted but can be reconstructed from Theorems 3.1.1 and 3.1.2 in [23]. Properties 1 to 6 ensure the existence of $\nu$ and that the fused hyperparameters are given by (15). If all the agents follow (11) through (13) and make independent measurements at time $\kappa$, letting $\omega = [\omega_1, \ldots, \omega_N]$ and $z = \omega^-$.

$^2$Time invariance is included to simplify the proof, but not explicitly required (see Proposition 2.3.1 in [23]). Knowledge of $\nu$ is often implicitly achieved by assuming a balanced network; analysis of limiting consensus values on an unbalanced network requires $\nu$.

$^3$This is not as restrictive as it might seem since this is always the case with any Kalman-derived method, but is confined to the normal distribution.

$^4$This can be extended to the case of globally shared measurements or multiple measurements per iteration without difficulty, but is omitted here for simplicity of presentation.
\[ [z_1, \ldots, z_N]^T, \text{ then the resulting one-step consensus update} \]
\[ \omega_{\kappa + 1} = A \omega_{\kappa} + \text{diag}(\nu^T) h(z_{\kappa}, f_{Z|\Theta(\omega)}), \]  
where \((\cdot)^T\) denotes the element-wise inverse; \text{diag}(x) is a diagonal matrix with diagonal entries given by the elements of \(x\); and \(h(\cdot)\) acts element-wise on, and is the same size as, \(z_{\kappa}\). Expanding (16) recursively back to the initial conditions, taking the limit as \(\kappa \to \infty\) (with no measurements for \(k > K\)) and utilizing \(\lim_{k \to \infty} A^k = \nu^T, \nu^T A^m = \nu^T \forall m\) and \(\nu^T \text{diag}(\nu^T) = 1^T\), it can be shown that the limiting value of (16) is equivalent to (15).

The above proof can be extended to the case where multiple measurements are made by each agent between consensus iterations by recursively applying (13) for each measurement, as well as to the case of globally shared measurements by applying the measurement update in (14).

IV. EXAMPLE

This section elucidates the preceding derivation of the hyperparameter consensus method through the estimation of a Poisson arrival rate, \(\lambda\), by exploiting conjugacy with the gamma prior. The gamma and Poisson distributions are shown in (17) and (18), respectively.

\[ f_{\lambda|A,B}(\lambda|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta \lambda} \]  
\[ f_{X|\lambda,T}(x|\lambda, T) = \frac{(\lambda T)^x e^{-\lambda T}}{x!} \quad x = 0, 1, 2, \ldots \]  
The hyperparameters in this situation are the scalars \(\alpha\) and \(\beta\), which can be shown to be updated after a measurement as

\[ \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \leftarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} x \\ T \end{bmatrix} h(z, f_{Z|\Theta(\omega)}) \]  
Using the MMSE metric, the best estimate of \(\lambda\) is the mean of the gamma distribution, \(\lambda_{\text{MMSE}} = \frac{\alpha}{\beta}\), and the fused MMSE estimate is found as \(\lambda_{\text{fused}} = \frac{\sum_{i=1}^N \alpha_i}{\sum_{i=1}^N \beta_i}\).

The following discussion concerns five agents agreeing over either a) a known, balanced network or b) a known, unbalanced network with a consensus eigenvector of \(\nu = [1 \ 1 \ 1 \ 1 \ 1]^T\). Consensus is run simultaneously on both \(\alpha\) and \(\beta\) using the initial conditions for each agent shown in Table I, and the weighting factor \(\epsilon\) in (12) set to \(\frac{1}{N}\). The consensus is assumed to have no shared prior information, such that \(\alpha_0 = \beta_0 = 0\).  

<table>
<thead>
<tr>
<th>Table I</th>
<th>Initial Conditions for (\lambda) Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent:</td>
<td>1</td>
</tr>
<tr>
<td>(\alpha_i)</td>
<td>5</td>
</tr>
<tr>
<td>(\beta_i)</td>
<td>5</td>
</tr>
<tr>
<td>(\lambda_{\text{MMSE}})</td>
<td>1</td>
</tr>
</tbody>
</table>

A. Results

Known Network Without Measurements: Figure 1 shows the local estimates of \(\lambda\) for each of the five agents during consensus, where the dotted red line represents \(\lambda_{\text{fused}}\), and the difference between the “Initial” and “Weighted” data points is the initialization weighting of the hyperparameters in (11). Since the MMSE estimate is a linear ratio of the hyperparameters, the initialization has no effect on the local parameter estimate. The algorithm converges on both the balanced and unbalanced networks to the fused MMSE estimate, \(\lambda = 4\), with the only difference being the rate of convergence which is slower in the unbalanced network because it contains fewer edges than the balanced graph.

Known Network With Measurements: In order to demonstrate the adaptability of the method to simultaneous measurements, Figure 2 shows the result of one agent making a measurement corresponding to \(x = 80\) and \(T = 10\) at either iteration \(k = 5\) or iteration \(k = 20\) of the consensus. The figures show that, in both cases, the agents immediately begin to track the new Bayesian parameter estimate once a measurement has been made, regardless of whether the initial consensus has converged or not. This holds true regardless of the number of measurements or sensing agents, so long as there exists finite some iteration \(K\) after which no measurements are made.

Sensitivity to Network Topology: The hyperparameter consensus method works best on known networks so that the agents can weight their initial conditions according to the value of \(\nu\). However, even when the network is unknown and the initial conditions cannot be weighted properly, the hyperparameter consensus method with \(\eta_i = 1/N\) is shown to perform better than pure linear consensus on the parameter itself (ie. running (3) on \(\lambda_{\text{MMSE}} = \alpha_i/\beta_i\)). A comparison is made between the expected percent error in the parameter...
estimate for the hyperparameter and parameter consensus methods, where the errors are defined as
\[
\hat{e} = \left| \frac{\hat{\lambda} - \lambda_{\text{fused}}}{\lambda_{\text{fused}}} \right| \times 100\% \quad \text{and} \quad \bar{e} = \left| \frac{\bar{\lambda} - \lambda_{\text{fused}}}{\lambda_{\text{fused}}} \right| \times 100\%
\]
respectively, and where \(\hat{\lambda}\) is the steady-state hyperparameter consensus estimate and \(\bar{\lambda}\) is the parameter consensus estimate. Figure 3 shows the amount by which the expected parameter consensus error is greater than the expected hyperparameter consensus error for different network sizes as a function of the degree of biasedness of the network [23], where a network has a bias of 0 if it is balanced and a bias of 1 if \(\nu = [0 \ 0 \ldots \ 0 \ 1]^T\). The plotted results are obtained by sampling 200,000 Monte Carlo simulations for each data point with initial conditions created to represent each agent having a small, random number of observations. It shows that, for all values of network bias, it is expected that the hyperparameter consensus method will achieve a MMSE parameter estimate that is closer to the proper fused estimate than linear consensus on the parameter, even when the network is unknown.

V. CONCLUSIONS

This paper has derived a generic hyperparameter consensus method that allows a network of agents to come to an unbiased distributed agreement to the Bayesian fusion of their local parameter estimates defined by possibly non-Gaussian uncertainties. The primary innovation of this method is the use of the hyperparameters of local distributions as the consensus variables, and we show that the Bayesian fusion problem with measurements can be solved if there exists some knowledge of how each agent’s information contributes to the final consensus value (such as the consensus eigenvector) and the measurement is then weighted appropriately.

Future work is looking at extending hyperparameter consensus to the case of continuous-time consensus. In addition, we are investigating the role of the hyperparameter consensus problem in robust planning.

ACKNOWLEDGMENTS

This work is funded by AFOSR # FA9550-08-1-0086.

REFERENCES