PROBABILITY DENSITY UPDATE FOR A DISTRIBUTED SYSTEM BASED ON UNNORMALIZED LOCAL DENSITIES

Claire L. McCullough
ECSE Department
University of Alabama in Huntsville
Huntsville, Alabama 35899

Abstract

Results have been published [1] which give the conditional probability density update for a distributed nonlinear stochastic system based on the conditional density updates for the local estimators. However, the local densities are required to be solutions to Kushner's equation, which is unsolvable in the general case. In this paper, a formulation is given which uses unnormalized local densities, which are solutions of the more tractable Zakai equation [2], and yet produces the same normalized density for the global system as the previous method. This new result is applicable to a much wider variety of problems than the former due to the large body of problems for which the Zakai equation is solvable and Kushner's equation is not.

Motivation and Problem Formulation

The problem referred to as "filtering" or "estimation" consists of taking all state measurements up to the present time and producing an estimate of the present state of the system. In the classic problem, all measurements are available to a central processor, which combines the information to produce the desired state estimate.

In many realistic problems, it is sometimes more feasible to perform the measurement processing on a local level and transmit some statistic, or statistics, of the local estimate to a processor which combines them to form an estimate of the global system state [2]. This is known as distributed estimation. The applications of distributed estimation can be divided into three major areas:

1. systems with spatial constraints, requiring communication between widely spaced subsystems;
2. systems with reliability (or survivability) constraints; and
3. systems with the need for parallel processing of data.

The basic form of the generic distributed estimation problem is shown on the following page in Figure 1. The N local measurement stations take measurements of the global process. Each local estimator, which has no knowledge (or limited knowledge) of the global system, produces an estimate of the local state conditioned on the local measurement history. This estimate, which in the general case is a conditional probability density, is communicated to a central processor. The central processor then combines the N local estimates to form an estimate of the global state conditioned on the measurement histories of all N measurement stations. The central processor does have knowledge of the global system.

The problem addressed here is the continuous case in which the system state evolves according to a continuous stochastic differential equation, and measurements are taken continuously. Mathematically, this problem may be formulated as follows: the global system is defined by

\[ dx_t = f(x_t,t)dt + g(x_t,t)dB_t \]  \hspace{1cm} (1.1)

where \( B_t \) is a Brownian motion process with variance parameter

\[ E[db_t^Tdb_t] = Q(t)dt \quad Q(t) \geq 0 . \]  \hspace{1cm} (1.2)

It is assumed that N measurement stations exist. The measurement \( dy_{il} \) is the incremental data taken by the \( i \)th measurement station at time \( t \), and satisfies the equation

\[ dy_{il} = h(x_t,t)dt + \sigma(x_t,t)dW_t \]  \hspace{1cm} (1.3)

where \( W_t \) is a Wiener process with variance parameter

\[ E[dW_t^Tdw_t] = R(t)dt \quad R(t) \geq 0 . \]  \hspace{1cm} (1.4)

The central processor combines the local estimates to form an estimate of the global state conditioned on the measurement histories of all N measurement stations. This estimate, which in the general case is a conditional probability density, is communicated to a central processor. The central processor then combines the N local estimates to form an estimate of the global state conditioned on the measurement histories of all N measurement stations. The central processor does have knowledge of the global system.
\[ dy_{i,t} = H_i(x_i,t)dt + dw_{i,t} \]  

(1.3)

The measurement noise \( w_{i,t} \) is also Brownian with variance parameter \( R_i(t)dt \), where \( R_i(t) > 0 \). We follow Alouani [1] in assuming that \( w_{i,t} \) and \( w_{j,t} \) are independent for \( i \neq j \). The functions \( H_i \) are vector functions bounded in \( t \) and continuous in both \( t \) and \( x_t \). The local model of processor \( i \) is given as

\[ dx_{i,t} = f_i(x_{i,t},t)dt + g_i(x_{i,t})dw_{i,t} \]  

(1.4)

with local measurement equation

\[ dy_{i,t} = h_i(x_{i,t},t)dt + dw_{i,t} \]  

(1.5)

Both \( w_{i,t} \) and \( w_{j,t} \) are Brownian processes. The variance parameter of \( dw_{i,t} \) is \( Q(t)dt \), with \( Q(t) < 0 \). The processes \( w_{i,t} \) and \( w_{j,t} \) are assumed to be independent. The measurement noise \( w_{i,t} \) in equation (1.5) is assumed to be the same as \( w_{i,t} \) in equation (1.3) above (noise in the global measurement case). The physical measurements modeled by equations (1.3) and (1.5) are assumed to be the same also: the measurement is actually being generated by the local system (equations (1.1) and (1.3)), but the local estimator believes it to be produced by the local model (equations (1.4) and (1.5)). Also assume that local models and the global system produce valid states (see [5] for sufficient conditions).

For the measurement models, both local and global, it is required that the vector functions \( H_i \) and \( h_i \) are bounded for all \( t \), are continuous in both \( x \) (or \( x_i \)) and \( t \), and that for each sample path \( x \) of equation (1.1) there exists a sample path \( x_i \) of equation (1.4) such that for some \( \beta_i \)

\[ H_i(x_i,t) = h_i(x_{i,t},t). \]  

(1.6)

The estimates produced by the local filters (in the general case, the estimate produced by each local filter is in the form of the conditional probability density of the state of the local model conditioned on the measurements taken at that measurement station) are communicated to a central processor, which combines them to produce a probability density for the state of the global system conditioned on the measurements taken at all \( N \) measurement stations.

**Results**

Using an information fusion approach, Alouani [1] generates the probability density update function for the global state conditioned on the distributed measurements as follows:

\[ p_x[Y|X_1|T_{0}^{+dt}] = \frac{\prod_{i=1}^{N} p_x[Y_i|X_i|T_{0}^{+dt}]C(x_i)}{\int_{X_1} p_x[Y_i|X_i|T_{0}^{+dt}]C(x_i)dx_i} \]  

(2.1)

where

\[ C(x_i) = \frac{p_x[Y_i|X_i|T_{i}^{+dt}]}{\prod_{i=1}^{N} p_x[Y_i|X_i|T_{i}^{+dt}]C(x_i)} \]  

(2.2)

and \( p_x[Y_i|X_i|T_{i}^{+dt}] \) is the conditional density of the local state conditioned on the local measurements, and evaluated at \( x_i = T_i(x) \). Here use Alouani’s convention of allowing \( M_{i+dt} \) to represent the set of all measurements taken at time \( t + dt \) at all measurement stations, \( Y_i \) to be the set of all measurements taken at station \( i \) up to and including those at time \( t \), \( Y_t = (Y_i, i = 1, N) \), and \( y_i \) to be the measurement taken at station \( i \) at time \( t \).

The difficulty in calculating Alouani’s conditional density is that the normalized local conditional densities of which it is composed evolve according to Kushner’s equation [3]

\[ dp_i = \mathcal{L}(p_i)dt + \frac{1}{2} \left( h_i(t)R_i(t)dy_i,t \right) \]  

(2.3)

for \( t > t_0 \) where

\[ h_i(t) = E[h_i(x_t,0)|Y_i] \]  

(2.4)

and \( \mathcal{L} \) is the diffusion operator defined by

\[ \mathcal{L}(p,f,g,Q) = -\sum_{j=1}^{n} \frac{\partial}{\partial x_j} f_j + \frac{1}{2} \sum_{j,k=1}^{n} \frac{\partial^2}{\partial x_j \partial x_k} Q_{jk}. \]  

(2.5)

Kushner’s equation is extremely difficult to solve, as it is nonlinear in the conditional density \( p_i \). The calculation of the problem could be simplified considerably if Alouani’s density could be expressed in terms of the unnormalized versions of the local densities, which would be generated by the Zakai equation as follows [4]:

\[ dq_i = \mathcal{L}(q_i)dt + \frac{1}{2} \left( h_i(t)R_i^{-1}(t)dy_i,t \right) \]  

(2.6)

with the operator \( \mathcal{L} \) defined as above. Though still a difficult equation to solve, this is not as unmanageable as the problem of solving Kushner’s equation, as it lacks the nonlinearity caused by the presence of the \( h_i(t) \) term in equation (2.3). Exact solutions are known, in fact, for a fairly broad class of nonlinear systems [5,6,7].

The solutions to the Zakai equation and Kushner’s equation are related by the equation
It is implied by Alouani [1] that taking the numerator of the density in equation (2.1) would give an unnormalized density which would evolve according to the Zakai equation. This is false but can be demonstrated by taking the formal derivative of this numerator and comparing it to the Zakai equation. This calculation is straightforward, but is omitted because of its length and complexity. We can, however, use the Zakai unnormalized densities for the local states in the determination of a conditional density for the global state. We can, in fact, get an equation for the normalized conditional density update of the global state constructed out of the unnormalized (Zakai equation) densities for the local states. This density is as follows:

**Theorem**

\[ P_x|Y(x_t|Y_{1:t}) = \frac{q_{x_t|Y(x_t|Y_{1:t})}}{\int X q_{x_t|Y(x_t|Y_{1:t})} dx} \quad (2.7) \]

where

\[ \hat{C}(x_t) = \frac{q_{x_t|Y(x_t|Y_{1:t})}}{\int X q_{x_t|Y(x_t|Y_{1:t})} dx} \quad (2.9) \]

and \( q_{x_t|Y(x_t|Y_{1:t})} \) is the unnormalized conditional density of the local state conditioned on the local measurements, and evaluated at \( x_t = T(x) \).

**Proof of Result**

This proof is similar to Alouani's proof of the normalized density update result in the discrete time case [1]. Allowing \( M_{t+dt} \) to represent the set of all measurements taken at time \( t+dt \) at all measurement stations, \( Y_t^i \) to be the set of all measurements taken at station \( i \) up to and including those at time \( t \) if \( i = 1, N \), and \( y_t^i \) to be the measurement taken at station \( i \) at time \( t \), we can write

\[ P_x|Y(x_t|Y_{1:t}) = P_x|M_{t+dt}|Y_{1:t}|. \quad (3.1) \]

Using Bayes' rule, this becomes

\[ P_x|Y(x_t|Y_{1:t}) = \frac{P_{x|M_{t+dt}}|Y_{1:t}|P_{y|M_{t+dt}}|Y_{1:t}|}{P_{M|M_{t+dt}}|Y_{1:t}|}. \quad (3.2) \]

By definition,

\[ P_{M|M_{t+dt}}|X_t = P_{y|x|^t}_{t+dt}= \cdot \cdot \cdot y_N^N|x_{t+dt}|. \quad (3.3) \]

The probability of the measurements given the state is just the probability of the noise in the measurements. Using the assumption that measurement noise at the various stations is independent, we can write

\[ P_M|M_{t+dt}|X_t = \prod_{i=1}^N P_{y|x|^i}_{t+dt}|x_i| \quad (3.4) \]

Equation (3.2) can be written as

\[ P_x|Y(x_t|Y_{1:t}) = \frac{P_{M}|M_{t+dt}|x_t|I_{x_t}|y_{x_t}|Y_t}{P_{M}|Y|M_{t+dt}|Y_t|X_{t}|x_t|y_{x_t}|Y_t|dx} \quad (3.5) \]

Letting \( \int x q_{x_t|Y_{1:t}} dx \equiv \bar{N} \) and using equation (3.4)

\[ P_x|Y_{1:t} = \frac{\int X q_{x_t|Y_{1:t}} dx}{\bar{N}} \quad (3.6) \]

From [1]

\[ P_{y|x|X_{1:t}} = P_{y|x}|T(x_{1:t}) \quad (3.7) \]

Hereafter written as \( p_{y|x|Y_{1:t}} = T(x_{1:t}) \). By Bayes' rule

\[ P_{y|x|X_{1:t}} = P_{x|Y_{1:t}}|x_{1:t}|p_{y|x|Y_{1:t}}|x_{1:t}| \quad (3.8) \]

Using the relationship between Zakai's unnormalized density and Kushner's normalized one

\[ P_{y|x|X_{1:t}} = \frac{P_{x|Y_{1:t}}|x_{1:t}|p_{y|x|Y_{1:t}}|x_{1:t}|}{P_{y|x|Y_{1:t}}|x_{1:t}|} \quad (3.9) \]

Call the bracketed quantity in equation (3.9) \( N_y \) and substitute equations (3.7) and (3.9) into equation (3.6), giving
From the Chapman–Kolmogorov equation [3]

\[ p_{x|y}[M_{t+dt}|Y_t] = \int_{\mathcal{X}_t} p_{y|M_{t+dt}|X_t} \left[ \prod_{i=1}^{N} q_{x_{i}|y}[Y_{t+dt}|x_{i}] \prod_{i=1}^{N} p_{y|[x_{i}]}[Y_{t+dt}|x_{i}] \right] dx_t \]

Substituting equations (3.7) and (3.9) into equation (3.13), and using the definitions of \( \mathcal{N} \) and \( \mathcal{N}_t \) gives

\[ p_{x|y}[M_{t+dt}|Y_t] = \mathcal{N} \prod_{i=1}^{N} \prod_{i=1}^{N} q_{x_{i}|y}[Y_{t+dt}|x_{i}] \left[ \int_{\mathcal{X}_t} p_{y|[x_{i}]}[Y_{t+dt}|x_{i}] \right] dx_t \]

Substituting equations (3.7) and (3.9) into equation (3.13), and using the definitions of \( \mathcal{N} \) and \( \mathcal{N}_t \) gives

\[ p_{x|y}[M_{t+dt}|Y_t] = \mathcal{N} \prod_{i=1}^{N} \prod_{i=1}^{N} q_{x_{i}|y}[Y_{t+dt}|x_{i}] \left[ \int_{\mathcal{X}_t} p_{y|[x_{i}]}[Y_{t+dt}|x_{i}] \right] dx_t \]

Substituting equations (3.7) and (3.9) into equation (3.13), and using the definitions of \( \mathcal{N} \) and \( \mathcal{N}_t \) gives

\[ p_{x|y}[M_{t+dt}|Y_t] = \mathcal{N} \prod_{i=1}^{N} \prod_{i=1}^{N} q_{x_{i}|y}[Y_{t+dt}|x_{i}] \left[ \int_{\mathcal{X}_t} p_{y|[x_{i}]}[Y_{t+dt}|x_{i}] \right] dx_t \]

Therefore, simplifying and defining

\[ \tilde{C}(x_t) = \frac{q_{x_{i}|y}[Y_{t+dt}|x_{i}]}{p_{x_{i}|y}[Y_{t+dt}|x_{i}]} \]

gives

\[ p_{x|y}[M_{t+dt}|Y_t] = \frac{\mathcal{N} \prod_{i=1}^{N} q_{x_{i}|y}[Y_{t+dt}|x_{i}] \left[ \int_{\mathcal{X}_t} p_{y|[x_{i}]}[Y_{t+dt}|x_{i}] \right] \tilde{C}(x_t)}{\int_{\mathcal{X}_t} \prod_{i=1}^{N} q_{x_{i}|y}[Y_{t+dt}|x_{i}] \left[ \int_{\mathcal{X}_t} p_{y|[x_{i}]}[Y_{t+dt}|x_{i}] \right] \tilde{C}(x_t) dx_t} \]

The density \( p_{y|[Y_{t+dt}|Y_t} \) can be written as

\[ p_{y|[Y_{t+dt}|Y_t} = \frac{\int_{\mathcal{X}_t} \prod_{i=1}^{N} q_{x_{i}|y}[Y_{t+dt}|x_{i}] \left[ \int_{\mathcal{X}_t} p_{y|[x_{i}]}[Y_{t+dt}|x_{i}] \right] \tilde{C}(x_t) dx_t}{\int_{\mathcal{X}_t} \prod_{i=1}^{N} q_{x_{i}|y}[Y_{t+dt}|x_{i}] \left[ \int_{\mathcal{X}_t} p_{y|[x_{i}]}[Y_{t+dt}|x_{i}] \right] \tilde{C}(x_t) dx_t} \]