Propagation of Gaussian Belief Functions

Liping LIU

School of Business, University of Kansas, Lawrence, KS 66045 Bitnet: LIU@UKANVM

Abstract. Gaussian belief functions are represented in both variable space and configuration space. Their combinations are defined in terms of the Dempster's rule, sweep operators, and restrictions in configuration space. The equivalence of the alternative definitions is proved. The computation of Gaussian belief functions is shown to follow the Shafer-Shenoy axioms.

INTRODUCTION

Dempster (1990a, b) has shown how the Kalman filter can be understood in terms of the theory of belief functions. As Dempster shows, the equations and Gaussian probability distributions that are combined in the Kalman filter can be regarded as belief functions, and the recursion involved in the filter can be regarded as a special case of the recursion involved in the computation of belief-function marginals in join trees.

Dempster sketches how join-tree computation works for belief functions in general. There is some work yet to be done, however, in justifying this computation rigorously in the case of Gaussian belief functions. The rigorous work in this area (Kong 1986; Shafer *et al.* 1987; Shenoy and Shafer 1990) applies to finite and to condensable belief functions, but not to Gaussian belief functions, which are usually continuous but not condensable. Presumably the justification for the finite case can be extended to a justification for the continuous case by a straightforward limiting argument, but this has not been done to date. Dempster's description of Gaussian belief functions in geometric terms suggests that we take a different tack. We should be able to justify the join-tree computation by showing directly from this geometric description that the operations of combination of marginalization satisfy the axioms of Shenoy and Shafer (1990).

Dempster (1990b) defined the notion of Gaussian belief functions and speculated the possibility of their local computations. Shafer (1992) defined the concept in more rigorous mathematical forms by elaborating on Dempster's idea and left many "open" but important questions about the equivalence of various definitions and the possibility of local computations. The intent of this article is to provide proofs for the Dempster's speculations that answer some of the Shafer's "open" questions. In particular, this article attempts to understand the relationships among various concepts of Gaussian belief functions and explore whether the local computation scheme works for Gaussian belief functions.

An outline of this paper is as follows. Gaussian belief functions are represented respectively in variable and configuration space in Sections 2 and 3. Combination is defined in terms of the Dempster-Shafer rule and equivalently represented in terms of sweep operators and restrictions in configuration space respectively in Sections 4 to 6. In Section 7 we prove the possibility of local computation by showing that the computation of Gaussian belief functions follows the Shafer-Shenoy axioms. We refer readers to Liu [1993] for notations and formulas for computing the combination of Gaussian belief functions and Proofs of all results in the paper.

REPRESENTATION IN VARIABLE SPACE

Suppose $\mathbf{U}$ is a random variable space (Dempster 1969) – a finite dimensional vector space whose elements are random variables. A Gaussian belief function (Shafer 1992) on $\mathbf{U}$ is a quintuplet ($\mathbf{C}$, $\mathbf{B}$, $\mathbf{L}$, π, E), where $\mathbf{C}$, $\mathbf{B}$, and $\mathbf{L}$ are nested subspaces of $\mathbf{U}$,

$$\mathbf{C} \subseteq \mathbf{B} \subseteq \mathbf{L} \subseteq \mathbf{U},$$

π is a wide sense inner product on $\mathbf{B}$ with $\mathbf{C}$ as its null space, and E is a linear functional on $\mathbf{B}$. We call $\mathbf{C}$ the certainty space, $\mathbf{B}$ the belief space, $\mathbf{L}$ the label space, π the covariance, and E the expectation.

For ease of understanding above terms, we use the coordinate representation of a Gaussian belief function. Assume $V_1, ..., V_n$ is a basis for $\mathbf{U}$ such that $V_1, ..., V_c$ is a basis for $\mathbf{C}$, $V_1, ..., V_c, V_{c+1}, ..., V_b$ is a basis for $\mathbf{B}$, and $V_1, ..., V_c, V_{c+1}, ..., V_b, V_{b+1}, ..., V_l$ is a basis for $\mathbf{L}$. Let μ_i denote the mean of V_i (i = 1, 2, ..., b). For any $V = (v_1, ..., v_b) \in \mathbf{B}$, define the mean of V as $E(V) = v_1\mu_1 + ... + v_b\mu_b$. Let $\Sigma_{i,j}$ denote the covariance between V_{c+i} and V_{c+j} (i, j = 1, 2, ..., b-c). For $V^1 = (v^1{}_1, ..., v^1{}_n)$, $V^2 = (v^2{}_1, ..., v^2{}_n)$ in $\mathbf{B}$, define their covariance as

$$\pi(V^1, V^2) = (v^1{}_{c+1}, ..., v^1{}_b)\ \Sigma\ (v^2{}_{c+1}, ..., v^2{}_b)^T$$

where $\Sigma = (\Sigma_{ij})_{(b-c)\times(b-c)}$ is a covariance matrix. Then $E(\cdot)$ is a linear functional on $\mathbf{B}$ and $\pi(\cdot,\cdot)$ is a wider sense inner product on $\mathbf{B}$ with $\mathbf{C}$ as its null space: $\pi(V^1, V^2) = 0$ if V^1 or $V^2 \in \mathbf{C}$.

The expectation E and the covariance π define a Gaussian distribution for the variables in $\mathbf{B}$ by giving their means and covariances. This Gaussian distribution is regarded as a full expression of our beliefs, based on a given body of evidence, about the variables in $\mathbf{L}$; this item of evidence justifies no beliefs about variables in $\mathbf{L}$ going beyond what is implied by the beliefs about the variables in $\mathbf{B}$. (The evidence might justify some further beliefs about variables in $\mathbf{U}$ that are not in $\mathbf{L}$, but these are outside the conversation so far as a belief function with label **L** is concerned.) The Gaussian distribution assigns zero variances to the variables in **C**; if V is in **C**, we are certain that V takes the value E(V) with certainty. It assigns non-zero variances to variables in **B** that are not in **C**. Actually, the covariance π is defined over a subspace **F** in **B**, where **F**$\oplus$**C** = **B**. We call **F** the uncertainty space. V_{c+1}, ..., V_b is a basis for **F**. Equivalently, a Gaussian belief function can be represented by a quintuplet (**C**, **F**, **L**, π, E). In this paper we will use (**C**, **B**, **L**, π, E) and (**C**, **F**, **L**, π, E) interchangeably to represent a Gaussian belief function.

We will sometimes choose a linear functional t on **U** that agrees with E on **C**. (This means that t(V)=E(V) for every variable V in **C**; t is allowed to disagree with E on variables in **B** that are not in **C**, and t must also assign values to variables in **U** that are not in **C**.) When such a linear functional t has been chosen, we say that the Gaussian belief function is marked, and we call t its mark. We write (**C**, **B**, **L**, π, E, t) or (**C**, **F**, **L**, π, E, t) for a marked Gaussian belief function.

REPRESENTATION IN CONFIGURATION SPACE

Let **S** denote the dual space for **U**--the space of all linear functionals on **U**, and S_1, ..., S_n be its basis dual to V_1, ..., V_n. A point X = $(x_1,...,x_n)$ in **S** can be considered both as a linear functional such that $X(V_i) = x_i$ (i = 1, 2,..., n) and a vector value taken by random vector $(V_1,...,V_n)$. In the latter sense, X is a point in the sample space of random variables V_1, ..., V_n, and therefore is called a configuration. A probability distribution function is usually defined in space **S**.

Let S_2 be the dual kernel of **C**, i.e., the subspace of **S** consisting of all configurations which map all the variables in **C** to the value zero. That is, $S_2 = \{(x_1,...,x_n) \mid x_1 = ... = x_c = 0\}$; let S_1 be the dual kernel of **B**, i.e., $S_1 = \{(x_1,...,x_n) \mid x_1 = ... = x_b = 0\}$; let S_0 be the dual kernel of **L**, i.e., $S_0 = \{(x_1,...,x_n) \mid x_1 = ... = x_l = 0\}$. Let **E** be a hyperplane in **S** consisting of all the linear functionals on **U** that agrees with E on **B**, i.e., in terms of coordinates,

$$\mathbf{E} = \{(x_1,...,x_n) \mid x_1 = \mu_1, ..., x_b = \mu_b\}.$$

Obviously, we have the nested relationships $S_2 \supseteq S_1 \supseteq S_0$, where S_1 is parallel to **E**.

Let ρ be the dual of π, a wide sense inner product on S_2 with null space S_1. For any

$$X^1 = (0,0,...,0,x^1_{c+1},x^1_{c+2},...,x^1_n),$$

$$X^2 = (0,0,...,0,x^2_{c+1},x^2_{c+2},...,x^2_n),$$

in S_2, its coordinate representation is as follows:

$$\rho(X^1,X^2) = (x^1_{c+1},x^1_{c+2},...,x^1_b)\,\Sigma^{-1}\,(x^2_{c+1},x^2_{c+2},...,x^2_b)^T$$

and

$$\rho(X^1,X^2)\begin{cases}\geq 0 & X^1 \text{ and } X^2 \in S_2\\ =0 & X^1 \text{ or } X^2 \in S_1\end{cases}$$

Recall that a mark t for (**C**, **B**, **L**, π, E) is any linear functional on **U** that agrees with E on **C**, and so is a configuration in space **S**. Let **C***, **B***, and **L*** denote the hyperplanes which contain t and are respectively parallel to the space S_2, S_1 and S_0. Using coordinates, we represent these hyperplanes as follows: $\mathbf{C}^* = \{(x_1,...,x_n) \mid x_1 = \mu_1, ..., x_b = \mu_c\}$, $\mathbf{B}^* = \{(x_1,...,x_n) \mid x_1 = \mu_1, ..., x_b = \mu_c, x_{c+1} = t(V_{c+1}), ..., x_b = t(V_b)\}$, $\mathbf{L}^* = \{(x_1, ..., x_n) \mid x_1 = \mu_1, ..., x_b = \mu_c, x_{c+1} = t(V_{c+1}), ..., x_l = t(V_l)\}$. It follows that $\mathbf{L}^* \subseteq \mathbf{B}^* \subseteq \mathbf{C}^*$, $\mathbf{E} \subseteq \mathbf{C}^*$ and **E** and **B*** are parallel in **C***.

Notice that the probability distribution of a Gaussian belief function is defined on **C*** while **C*** is not a subspace. It is necessary to introduce operations for an inner product on **C***. For any h, h_1, $h_2 \in$ **C***, and any real number α, define

$$h_1 \oplus h_2 = (h_1 - t) + (h_2 - t) + t \qquad (1)$$

$$\alpha \otimes h = \alpha(h - t) + t \qquad (2)$$

It is easy to verify that $\oplus$, $\otimes$ are closed operations on **C***.

For h_1, $h_2 \in$ **C***, $h_1 - t$, $h_2 - t \in S_2$, ρ is a wide sense inner product on S_2. Thus, it is reasonable to define

$$\pi^*(h_1, h_2) = \rho(h_1 - t, h_2 - t) \qquad (3)$$

Lemma 1. $\pi^*(h_1, h_2)$ is a wide sense inner product on **C*** with operations $\oplus$, $\otimes$, and the null hyperplane **B***.

Lemma 2. The functional $g(\cdot)$ is linear on **C*** with operations $\oplus$ and $\otimes$ and takes the zero-value on **B*** iff there exists a matrix A such that

$$g(h)=A(x_{c+1} - t(V_{c+1}), ..., x_b - t(V_b)) \quad (4)$$

where $h = (\mu_1, ..., \mu_c, x_{c+1}, ..., x_n) \in$ **C***.

For any given $h_0 \in$ **C***, $\pi^*(h_0, h)$ is a linear functional on **C*** with null hyperplane **B***. Define

$$\mathbf{H}(\pi^*(\cdot,\cdot)) = \{h_0 \in \mathbf{C}^* \mid \pi^*(h_0, h) \text{ is an invariant linear functional of h on } \mathbf{C}^*\} \qquad (5)$$

Then, Lemma 2 implies that $\mathbf{H}(\pi^*(\cdot,\cdot))$ contains those configurations h_0 of **C*** such that $(x_{c+1}, ..., x_b)$ is a constant if $h_0 = (\mu_1, ..., \mu_c, x_{c+1}, ..., x_n)$. Hence, $\mathbf{H}(\pi^*(\cdot,\cdot))$ is a hyperplane in **C***, which is parallel to **B*** and whose location is specified by the constant vector $(x_{c+1},...,x_b)$. In general, for any linear functional $g(\cdot)$ in the form (4),

$$\mathbf{H}(g(\cdot)) = \{h_0 = (\mu_1, ..., \mu_c, x_{c+1}, ..., x_n) \mid (x_{c+1} - t(V_{c+1}), ..., x_b - t(V_b)) = A\Sigma\} \quad (6)$$

is a hyperplane parallel to **B*** in **C***. Therefore, we have

Lemma 3. Any linear functional on **C*** with null hyperplane **B*** specifies a hyperplane parallel to **B*** in **C***, whose location is determined by (6).

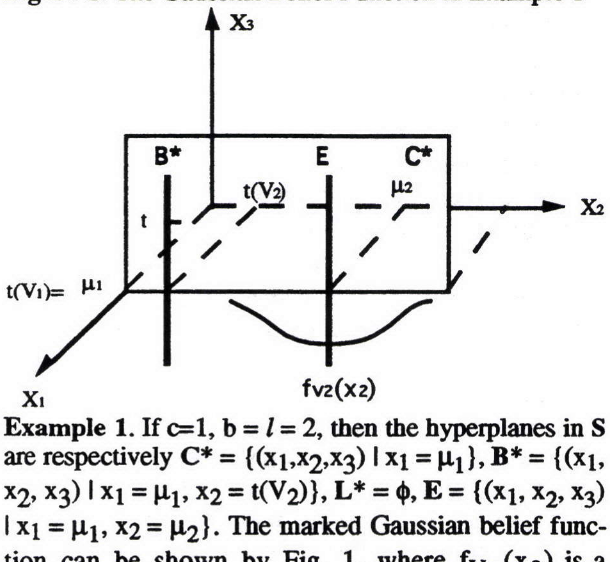
On the other hand, for any hyperplane **H** parallel to **B*** in **C***, $(x_{c+1},...,x_b)$ is a constant vector for any $h_0 = (\mu_1, ..., \mu_c, x_{c+1}, ..., x_n) \in$ **H**. Thus, for $h_0 \in$ **H**,

$\pi^*(h_0,h)$ is a linear functional on $\mathbf{C}^*$ with the null $\mathbf{B}^*$ and is denoted by $H^*(h)$. Since the covariance matrix Σ has a full rank, a hyperplane parallel to $\mathbf{B}^*$ and a linear functional in the form (4) are in a one-to-one correspondence. As a corollary, noticing that $\mathbf{E}$ is a hyperplane parallel to $\mathbf{B}^*$ in $\mathbf{C}^*$, we have

Theorem 1. Hyperplane $\mathbf{E}$ and the linear functional $E^*(h)=\pi^*(h_0,h)$ ($h_0 \in \mathbf{E}$, $h \in \mathbf{C}^*$) are mutually and uniquely determined.

So we arrive at the representation ($\mathbf{C}^*$, t, $\mathbf{B}^*$, $\mathbf{L}^*$, π^*, E^*) for a marked Gaussian belief function. We write t before $\mathbf{B}^*$, $\mathbf{L}^*$, π^*, and E^* because all these objects depend on the choice of t. Intuitively, ($\mathbf{C}^*$, t, $\mathbf{B}^*$, $\mathbf{L}^*$, π^*, E^*) expresses beliefs about which element of $\mathbf{S}$ is the true configuration of our variables. We are certain that the true configuration is in the hyperplane $\mathbf{C}^*$ (the certainty hyperplane). Within $\mathbf{C}^*$, our belief is distributed over ellipsoidal cylinders around a smaller dimensional hyperplane $\mathbf{E}$ (the expectation hyperplane) parallel to $\mathbf{B}^*$. The wide sense inner product π^* (the concentration inner product) specifies the shape, scale, and direction of the ellipsoidal cylinders, and the linear functional E^* (the location functional) specifies $\mathbf{E}$ by giving its inner product with every other hyperplane parallel to $\mathbf{B}^*$ within $\mathbf{C}^*$. We call $\mathbf{B}^*$ the no-opinion-expressed space, since the Gaussian belief function does not express any opinions about where the true configuration is along its coordinates. Similarly, we call $\mathbf{L}^*$ the no-opinion-allowed space, since the Gaussian belief function, so long as it has the label $\mathbf{L}$, is not allowed to express any opinions about where the true configuration is along its coordinates.

Figure 1. The Gaussian Belief Function in Example 1



Example 1. If c=1, b = l = 2, then the hyperplanes in $\mathbf{S}$ are respectively $\mathbf{C}^* = \{(x_1,x_2,x_3) \mid x_1 = \mu_1\}$, $\mathbf{B}^* = \{(x_1, x_2, x_3) \mid x_1 = \mu_1, x_2 = t(V_2)\}$, $\mathbf{L}^* = \phi$, $\mathbf{E} = \{(x_1, x_2, x_3) \mid x_1 = \mu_1, x_2 = \mu_2\}$. The marked Gaussian belief function can be shown by Fig. 1, where $f_{V_2}(x_2)$ is a Gaussian probability distribution function for V_2.

COMBINATION BY DEMPSTER'S RULE

Assume any two Gaussian belief functions Bel_1 and Bel_2 whose representations in variable space are

$\mathrm{Bel}_1 = (\mathbf{C}_1, \mathbf{F}_1, \mathbf{L}_1, \pi^1, E^1)$ (7)

where, $\mathbf{C}_1 = \{\xi_0, \xi_1, X_1\}$, $\mathbf{F}_1 = \{X_2, X_3, X_4\}$,

$E^1(\xi_0) = \lambda_0$, $E^1(\xi_1) = \lambda_1$, $E^1(X_1) = \mu_1$, $E^1(X_2, X_3, X_4) = (\mu_2{}^1, \mu_3{}^1, \mu_4{}^1)$, $\pi^1(\cdot,\cdot) = \cdot\Sigma_1 \cdot^T$,

$\Sigma_1 = (\Sigma_{ij}{}^1)$, $\Sigma_{ij}{}^1 = \mathrm{cov}(X_i, X_j)$ (i, j = 2, 3, 4).

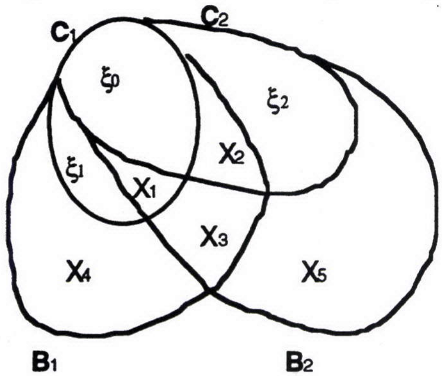
$\mathrm{Bel}_2 = (\mathbf{C}_2, \mathbf{F}_2, \mathbf{L}_2, \pi^2, E^2)$ (8)

where $\mathbf{C}_2 = \{\xi_0, \xi_2, X_2\}$, $\mathbf{F}_2 = \{X_1, X_3, X_5\}$,

$E^2(\xi_0) = \lambda_0$, $E^2(\xi_2) = \lambda_2$, $E^2(X_2) = \mu_2$, $E^2(X_1, X_3, X_5) = (\mu_1{}^2, \mu_3{}^2, \mu_5{}^2)$, $\pi^2(\cdot,\cdot) = \cdot\Sigma_2 \cdot^T$,

$\Sigma_2 = (\Sigma_{ij}{}^2)$, $\Sigma_{ij}{}^2 = \mathrm{cov}(X_i, X_j)$ (i, j = 1, 3, 5).

Figure 2. Combination of Bel_1 and Bel_2



Without loss of generality, we assume that Bel_1 and Bel_2 share a common certainty subspace ξ_0 and a common uncertainty subspace X_3. Bel_1 is certain but Bel_2 is uncertain about X_1; Bel_2 is certain but Bel_1 is uncertain about X_2. Bel_1 is certain about ξ_1 and is uncertain about X_4, but Bel_2 has no opinion about ξ_1 and X_4. Similarly, Bel_2 is certain about ξ_2 and is uncertain about X_5, but Bel_1 has no opinion about ξ_2 and X_5. The variables involved in these two belief functions is shown in Fig. 2. Notice that in (7), (8) and Fig. 2, we do not represent $\mathbf{L}_1$ and $\mathbf{L}_2$ explicitly. We also assume Bel_1 and Bel_2 have the same belief about the mean value of $\mathbf{C}_1 \cap \mathbf{C}_2$, i.e., ξ_0. The notations appearing in Fig. 2 represent either variable spaces or the basis vectors of the variable spaces.

According to the Dempster's rule (Shafer 1976), the combination of the two belief functions is to pool all information of Bel_1 and Bel_2 together through the following procedures to obtain a combined belief function Bel = ($\mathbf{C}$, $\mathbf{B}$, $\mathbf{L}$, π, E):

(1)Bel accepts all certain information from both Bel_1 and Bel_2: $\mathbf{C} = \mathbf{C}_1 \oplus \mathbf{C}_2$;

(2)Calculate conditional probabilities given $\mathbf{C}$;

(3)Calculate combined probabilities for common uncertain focal elements $F_1 \cap F_2$ using the Dempster-Shafer rule;

(4)Pool all above information including that accepted as certainty, changed conditionally, and combined by the Dempster's rule.

Procedure (4) implies that the belief space **B** of the combination is $B_1 \oplus B_2$, and that the label space **L** is $L_1 \oplus L_2$. The uncertainty space of $Bel_1 \oplus Bel_2$ is **F**, where $F \oplus C = B$. Notice that $F_1 \oplus F_2 \neq F$. In the following, we focus on the problem of how to get the pooled π and E for Bel.

For any positive definite nxn matrix Σ and n-dimensional vectors x and y, let

$$f(x, \Sigma, y)=\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}\exp\{-\frac{1}{2}(x-y)\Sigma^{-1}(x-y)^T\}.$$

Observe that $f(x, \Sigma, y)$ is actually a multivariate Gaussian distribution function with variable x, mean y and covariance matrix Σ.

In Gaussian belief function Bel_1, the belief on the uncertainty space F_1 is represented by the Gaussian probability function of (X_2, X_3, X_4) with mean $(\mu_2^1, \mu_3^1, \mu_4^1)$ and covariance Σ_1, i.e.,

$$P_{1,X_2,X_3,X_4}(x_2, x_3, x_4)$$
$$= f[(x_2, x_3, x_4), \Sigma_1, (\mu_2^1, \mu_3^1, \mu_4^1)]$$

Similarly, Bel_2 expresses the belief about F_2 by the Gaussian distribution of (X_1, X_3, X_5), which has the mean $(\mu_1^2, \mu_3^2, \mu_5^2)$ and the covariance matrix Σ_2, i.e.,

$$P_{2,X_1,X_3,X_5}(x_1, x_3, x_5)$$
$$= f[(x_1, x_3, x_5), \Sigma_2, (\mu_1^2, \mu_3^2, \mu_5^2)]$$

Since X_2 and X_1 are both in the combined certainty space C, we are certain about both X_2 and X_1. According to procedure (1), it follows that

$$P(X_1 = \mu_1) = P(X_2 = \mu_2) = 1.$$

Given $X_1 = \mu_1$ and $X_2 = \mu_2$, we can compute conditional beliefs $P_{1,X_3|X_2=\mu_2}(x_3)$ and $P_{2,X_3|X_1=\mu_1}(x_3)$ about X_3 respectively in Bel_1 and Bel_2. Since X_3 is the common uncertain focal element for Bel_1 and Bel_2, we then use the Dempster-Shafer rule to combine

$$P_{1,X_3|X_2=\mu_2}(x_3) \text{ and } P_{2,X_3|X_1=\mu_1}(x_3)$$

to get the pooled belief about X_3:

$$P_{X_3}(x_3)=P_{1,X_3|X_2=\mu_2}(x_3)\otimes P_{2,X_3|X_1=\mu_1}(x_3)$$
$$=\frac{P_{1.X_3|X_2=\mu_2}(x_3)P_{2.X_3|X_1=\mu_1}(x_3)}{\int_\Omega P_{1.X_3|X_2=\mu_2}(x_3)P_{2.X_3|X_1=\mu_1}(x_3)dx_3} \qquad (9)$$

The rule $\otimes$ in (9) is the extension of the Dempster-Shafer rule (Shafer 1976) for continuous belief functions. Note that in this section we use P_1 and P_2 to denote a distribution respectively in Bel_1 and Bel_2 and P to denote a distribution in the combined belief function Bel.

Lemma 4. The pooled belief about X_3 is a multivariate Gaussian distribution. Assume its mean vector is a_3 and covariance matrix is σ_3. Then,

$$P_{X_3}(x_3) = f(x_3, \sigma_3, a_3).$$

Given $X_2 = \mu_2$ and $X_3 = x_3$, the conditional distribution of X_4 is Gaussian. Assume its covariance matrix is σ_4 and its regression coefficients against x_3 are a_4 and b_4. Then we have

$$P_{1,X_4|X_2=\mu_2,X_3=x_3}(x_4) = f(x_4, \sigma_4, a_4+x_3b_4^T). \qquad (10)$$

Similarly, the conditional distribution of X_5 given $X_1 = \mu_1$ and $X_3 = x_3$ is Gaussian. Assume its covariance matrix is σ_5 and its regression coefficients against x_3 are a_5 and b_5. The

$$P_{2,X_5|X_1=\mu_1,X_3=x_3}(x_5) = f(x_5, \sigma_5, a_5+x_3b_5^T). \qquad (11)$$

Now we come to the pooling of all pieces of belief. Since X_4 and X_5 are conditionally independent given X_3, the pooled belief about X_3, X_4 and X_5 can be shown as follows:

$$P_{X_3,X_4,X_5}(x_3,x_4,x_5) = P_{X_3}(x_3)$$
$$P_{1,X_4|X_2=\mu_2,X_3=x_3}(x_4)\, P_{2,X_5|X_1=\mu_1,X_3=x_3}(x_5).$$

Let Ω denote the covariance matrix of X_3, X_4 and X_5. Then

$$\Omega=\begin{bmatrix} \sigma_3 & \sigma_3 b_4^T & \sigma_3 b_5^T \\ b_4\sigma_3 & \sigma_4+b_4\sigma_3 b_4^T & b_4\sigma_3 b_5^T \\ b_5\sigma_3 & b_5\sigma_3 b_4^T & \sigma_5+b_5\sigma_3 b_5^T \end{bmatrix}.$$

Lemma 5. $P_{X_3,X_4,X_5}(x_3, x_4, x_5) = f[(x_3, x_4, x_5), \Omega, (a_3, a_4 + a_3b_4^T, a_5 + a_3b_5^T)]$.

Note that the formulas for computing a_3, a_4, b_4, a_5, b_5, σ_3, σ_4, and σ_5 in Lemmas 4 and 5 and the equations (10) and (11) can be found in Liu [1993].

So we arrive at the conclusion about the combination of Bel_1 and Bel_2: $Bel = Bel_1 \oplus Bel_2$:

Theorem 2. The combination of the two Gaussian belief functions (7) and (8) is as follows:

$$Bel = Bel_1 \oplus Bel_2=(C,B,L,\pi,E) \qquad (12)$$

where, $C=C_1\oplus C_2$, $B=B_1\oplus B_2$, $F=\{X_3,X_4,X_5\}$, $L = L_1 \oplus L_2$, $E(\xi_0) = \lambda_0$, $E(\xi_1) = \lambda_1$, $E(\xi_2) = \lambda_2$, $E(X_1) = \mu_1$, $E(X_2) = \mu_2$, $E(X_3) = a_3$, $E(X_4) = a_4 + a_3b_4^T$, $E(X_5) = a_5 + a_3b_5^T$, $\pi(\cdot,\cdot) = \cdot\Omega\cdot^T$

COMBINATION IN TERMS OF SWEEP OPERATORS

The representation of Bel in Theorem 2 is very complicated. The relationships among E, π, π^1, E^1, π^2, and E^2 have no closed forms. However, Theorem 2 will serve as a reference for more refined definitions. In this section, we will use sweep operators to represent combination in variable space. Sweep operations were initially defined for positive definite matrices (Dempster 1969). Dempster (1990b) used the same operations to obtain the *potential form* (Lauritzen and Spiegelhalter 1988) of the representation matrix for a Gaussian distribution, which is an extended matrix consisting of a mean vector and a covariance matrix. Dempster (1990b) suggested a combination rule for Gaussian belief functions using sweep operators. Unfortunately, I am not able to verify his combination rule. The sweep operations are slightly extended in this paper. We will see shortly that the concept of sweeping and reverse sweeping in Dempster [1990b] is a special case of the concept here when the sweeping or reverse sweeping occurs at the origin of a configuration space.

Let $X_1, X_2, \dots, X_n$ be Gaussian random vectors such that $E(X_i) = \mu_i$, $Cov(X_i, X_j) = \Sigma_{ij}$, $i, j = 1, 2, \dots, n$. Set matrix

$$Q = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_n \\ \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2n} \\ \Sigma_{n1} & \Sigma_{n2} & \cdots & \Sigma_{nn} \end{bmatrix}.$$

Q is called the representation matrix of the probability function of $X_1, X_2, \dots, X_n$, or is said to describe $X_1, X_2, \dots, X_n$. We introduce the following four operations on representation matrices:

(1)**Marginalization:** The marginalization of Q on Z, denoted by $Q^{\downarrow Z}$, is the submatrix produced by retaining the mean vector row, covariance rows and columns corresponding to Z while deleting the rest. For example, $Q^{\downarrow X_1} = \begin{bmatrix} \mu_1 \\ \Sigma_{11} \end{bmatrix}$;

(2)**Direct sum:** For any two representation matrices Q_1 and Q_2, assuming their commonly described variables are X, rearrange Q_1 and Q_2 such that

$$Q_1 = \begin{bmatrix} \mu_X{}^1 & \mu_Y{}^1 \\ \Sigma_{XX}{}^1 & \Sigma_{XY}{}^1 \\ \Sigma_{YX}{}^1 & \Sigma_{YY}{}^1 \end{bmatrix}, Q_2 = \begin{bmatrix} \mu_X{}^2 & \mu_Z{}^2 \\ \Sigma_{XX}{}^2 & \Sigma_{XZ}{}^2 \\ \Sigma_{ZX}{}^2 & \Sigma_{ZZ}{}^2 \end{bmatrix}.$$

$$\text{Then } Q_1 \oplus Q_2 = \begin{bmatrix} \mu_X{}^1+\mu_X{}^2 & \mu_Y{}^1 & \mu_Z{}^2 \\ \Sigma_{XX}{}^1+\Sigma_{XX}{}^2 & \Sigma_{XY}{}^1 & \Sigma_{XZ}{}^2 \\ \Sigma_{YX}{}^1 & \Sigma_{YY}{}^1 & 0 \\ \Sigma_{ZX}{}^2 & 0 & \Sigma_{ZZ}{}^2 \end{bmatrix}.$$

(3)**Sweep operation:** The matrix Q will be said to be swept at configuration point x_i if Q is replaced by

$$Swp(X_i|_{X_i=x_i})Q = \begin{bmatrix} \mu_{1.i} & \mu_{2.i} & \cdots & \mu_{n.i} \\ \Sigma_{11.i} & \Sigma_{12.i} & \cdots & \Sigma_{1n.i} \\ \Sigma_{21.i} & \Sigma_{22.i} & \cdots & \Sigma_{2n.i} \\ \Sigma_{n1.i} & \Sigma_{n2.i} & \cdots & \Sigma_{nn.i} \end{bmatrix},$$

$$\text{where } \mu_{j.i} = \begin{cases} \mu_j-(\mu_i-x_i)(\Sigma_{ii})^{-1}\Sigma_{ij} & i\neq j \\ \mu_i(\Sigma_{ii})^{-1} & i=j \end{cases}$$

$$\Sigma_{kj.i} = \begin{cases} -(\Sigma_{ii})^{-1} & k=j=i \\ \Sigma_{kj}(\Sigma_{ii})^{-1} & j=i \text{ and } k\neq i \\ (\Sigma_{ii})^{-1}\Sigma_{kj} & k=i \text{ and } j\neq i \\ \Sigma_{kj}-\Sigma_{ki}(\Sigma_{ii})^{-1}\Sigma_{ij} & \text{otherwise} \end{cases}$$

It is easy to see that

$$(Swp(X_i|_{X_i=x_i})Q)^{\downarrow\{X_1,X_2,\dots,X_n\}\setminus\{X_i\}}$$

is the representation matrix of conditional probability function given $X_i = x_i$;

(4)**Reverse sweep operation:** A reverse sweep operation at the configuration point $X_i=x_i$, denoted by $Rwp(X_i|_{X_i=x_i})Q$, is an operation on Q such that

$$Rwp(X_i|_{X_i=x_i})Q = \begin{bmatrix} \mu'_{1.i} & \mu'_{2.i} & \cdots & \mu'_{n.i} \\ \Sigma'_{11.i} & \Sigma'_{12.i} & \cdots & \Sigma'_{1n.i} \\ \Sigma'_{21.i} & \Sigma'_{22.i} & \cdots & \Sigma'_{2n.i} \\ \Sigma'_{n1.i} & \Sigma'_{n2.i} & \cdots & \Sigma'_{nn.i} \end{bmatrix},$$

$$\text{where } \mu'_{j.i} = \begin{cases} \mu_j-(\mu_i+x_i\Sigma_{ii})(\Sigma_{ii})^{-1}\Sigma_{ij} & i\neq j \\ -\mu_i(\Sigma_{ii})^{-1} & i=j \end{cases}$$

$$\Sigma'_{kj.i} = \begin{cases} -(\Sigma_{ii})^{-1} & k=j=i \\ -\Sigma_{kj}(\Sigma_{ii})^{-1} & j=i \text{ and } k\neq i \\ -(\Sigma_{ii})^{-1}\Sigma_{kj} & k=i \text{ and } j\neq i \\ \Sigma_{kj}-\Sigma_{ki}(\Sigma_{ii})^{-1}\Sigma_{ij} & \text{otherwise} \end{cases}$$

Dempster (1990b) defines the sweeping and reverse sweeping operations at the configuration point $X_i=0$. According to his definition, the representations of $\mu_{j.i}$ and $\mu'_{j.i}$ above are the same in both sweeping and reverse sweeping when $i \neq j$. Shortly, in Lemma 10, we will see that the above generalized definitions of sweeping and reverse sweeping are necessary in representing the combination of Gaussian belief functions in variable space.

Lemma 6. If $Z \supset X$, then

$$Swp(X|_{X=x})Q^{\downarrow Z} = (Swp(X|_{X=x})Q)^{\downarrow Z},$$
$$Rwp(X|_{X=x})Q^{\downarrow Z} = (Rwp(X|_{X=x})Q)^{\downarrow Z}.$$

Lemma 7. For any two matrices A and B, $(A\oplus B)^{\downarrow Z} = A^{\downarrow Z} \oplus B^{\downarrow Z}$, $(A^{\downarrow Z})^{\downarrow Y} = A^{\downarrow Z\cap Y}$.

Lemma 8. $Swp(X_i|_{X_i=x_i})Rwp(X_i|_{X_i=x_i})Q = Rwp(X_i|_{X_i=x_i})Swp(X_i|_{X_i=x_i})Q = Q$.

As we know, $Swp(X_i|_{X_i=x_i})Q$ represents the conditional probability function given $X_i = x_i$. Lemma 8 implies that $Rwp(X_i|_{X_i=x_i})Swp(X_i|_{X_i=x_i})Q$ represents the joint $P_{X_i}(x_i)P_{X_1,X_2,\dots,X_n|X_i=x_i}(x_1, x_2, \dots, x_n)$. Currently, we do not know what $Rwp(X_i|_{X_i=x_i})Q$ represents. But it is interesting to note the commutativity of both sweep and reverse sweep operations at the same configuration point.

Lemma 9.

(i)$(Swp(X_i|_{X_i=x_i})\ Swp(X_j|_{X_j=x_j})Q)^{\downarrow\{X_1,..,X_n\}\backslash\{X_i,X_j\}}$

$=(Swp(X_j|_{X_j=x_j})Swp(X_i|_{X_i=x_i})Q)^{\downarrow\{X_1,..,X_n\}\backslash\{X_i,X_j\}}$

$=(Swp(X_i,\ X_j|_{X_i=x_i,X_j=x_j})Q)^{\downarrow\{X_1,...,X_n\}\backslash\{X_i,X_j\}}$

(ii)$Swp(X_i|_{X_i=0})\ Swp(X_j|_{X_j=0})\ Q = Swp(X_j|_{X_j=0})$
$Swp(X_i|_{X_i=0})\ Q = Swp(X_i,X_j|_{X_i=0,X_j=0})\ Q$.

Now we turn to the representation of Bel = $Bel_1 \oplus Bel_2$ = (**C**, **B**, **L**, π, E), where Bel_1 and Bel_2 are defined as in Section 4. Let Q, Q_1 and Q_2 respectively denote the representation matrices of Bel, Bel_1, and Bel_2.

Lemma 10.

$Q = Rwp(X_3|_{X_3=x_3})\ \{[Swp(X_3|_{X_3=x_3})$
$Swp(X_2|_{X_2=\mu_2})\ Q_1]^{\downarrow\{X_3,X_4\}} \oplus [Swp(X_3|_{X_3=x_3})$
$Swp(X_1|_{X_1=\mu_1})\ Q_2]^{\downarrow\{X_3,X_5\}}\}$.

In coordinate free terms, $X_3=\mathbf{F}_1\cap\mathbf{F}_2$, $X_2=\mathbf{F}_1\cap C_2$, $X_1=\mathbf{F}_2\cap C_1$. In Theorem 3, we use $Swp(\mathbf{F}_1\cap\mathbf{F}_2)$, $Rwp(\mathbf{F}_1\cap\mathbf{F}_2)$, $Swp(\mathbf{F}_1\cap C_2)$, and $Swp(C_1\cap\mathbf{F}_2)$ to replace $Swp(X_3|_{X_3=x_3})$, $Rwp(X_3|_{X_3=x_3})$, $Swp(X_2|_{X_2=\mu_2})$, and $Swp(X_1|_{X_1=\mu_1})$ respectively.

Theorem 3. For any two Gaussian belief functions $Bel_1 = (C_1, \mathbf{B}_1, L_1, Q_1)$ and $Bel_2 = (C_2, \mathbf{B}_2, L_2, Q_2)$, their combination is Bel = $Bel_1 \oplus Bel_2$ = (C, **B**, L, Q), where, $C = C_1\oplus C_2$, $\mathbf{B} = \mathbf{B}_1\oplus\mathbf{B}_2$, $L = L_1\oplus L_2$, $Q = Rwp(\mathbf{F}_1\cap\mathbf{F}_2)\ \{[Swp(\mathbf{F}_1\cap\mathbf{F}_2)\ Swp(\mathbf{F}_1\cap C_2)Q_1]^{\downarrow\mathbf{F}_1\cap\mathbf{F}}$ $\oplus[Swp(\mathbf{F}_1\cap\mathbf{F}_2)\ Swp(C_1\cap\mathbf{F}_2)Q_2]^{\downarrow\mathbf{F}_2\cap\mathbf{F}}\}$,
where **F**, $\mathbf{F}_1$, and $\mathbf{F}_2$ are respectively the uncertainty spaces for Bel, Bel_1 and Bel_2.

In Theorem 3, we notice that the means and covariances in Gaussian belief functions are replaced by their corresponding representation matrices. Theorem 3 gives an equivalent definition of combination in variable space. However, the relationship between Q and Q_1 and Q_2 in Theorem 3 is independent of coordinates. The closed form of combination in variable space will be very useful in proving the distributivity law of the Shafer-Shenoy axioms in Section 7.

COMBINATION AS RESTRICTIONS

In this section, we represent the combination in the configuration space. The representation of combination in configuration space is useful for proving the commutativity and associativelity law of combination in the Shafer-Shenoy axioms. Using the same notions as in Section 3, we denote the dual space of **U** by **S**. The dual representations for marked Bel_1 and Bel_2 are respectively

$Bel_1 = (\mathbf{C}^{*1}, t, \mathbf{B}^{*1}, \mathbf{L}^{*1}, \pi^{*1}, E^{*1})$,

$Bel_2 = (\mathbf{C}^{*2}, t, \mathbf{B}^{*2}, \mathbf{L}^{*2}, \pi^{*2}, E^{*2})$,

where t is a common mark for both Bel_1 and Bel_2; $\mathbf{C}^{*i}$, $\mathbf{B}^{*i}$, and $\mathbf{L}^{*i}$ are the hyperplanes containing t and parallel respectively to the dual kernel of $\mathbf{C}^i$, $\mathbf{B}^i$, and $\mathbf{L}^i$ (i = 1, 2), π^{*i} and E^{*i} are defined by Theorem 1 (i =1, 2).

Theorem 4. The combination of Bel_1 and Bel_2 is

$$\begin{aligned}Bel &= Bel_1\oplus Bel_2\\ &= \{\mathbf{C}^{*2}\cap\mathbf{C}^{*1}, t, \mathbf{B}^{*2}\cap\mathbf{B}^{*1}, \mathbf{L}^{*2}\cap\mathbf{L}^{*1},\\ &\pi^{*2}|_{\mathbf{C}^{*2}\cap\mathbf{C}^{*1}} + \pi^{*1}|_{\mathbf{C}^{*2}\cap\mathbf{C}^{*1}},\\ &E^{*2}|_{\mathbf{C}^{*2}\cap\mathbf{C}^{*1}} + E^{*1}|_{\mathbf{C}^{*2}\cap\mathbf{C}^{*1}}\},\end{aligned}$$

where $\pi^{*i}|_{\mathbf{C}^{*2}\cap\mathbf{C}^{*1}}$ and $E^{*i}|_{\mathbf{C}^{*2}\cap\mathbf{C}^{*1}}$ are respectively the restriction of π^{*i} and E^{*i} (i = 1, 2) on the intersection $\mathbf{C}^{*2}\cap\mathbf{C}^{*1}$.

LOCAL COMPUTATION OF GAUSSIAN BELIEF FUNCTIONS

In this section, we will verify that the computation of Gaussian belief functions follows the Shafer-Shenoy axioms (Shenoy and Shafer 1990). First we briefly describe the operation of marginalization of Gaussian belief functions.

Marginalization of Gaussian belief functions can be most naturally described in variable space. Suppose (C, **B**, **L**, π, E) is a Gaussian belief function, and **M** is a subspace of **L**. Then the marginal of (C, **B**, **L**, π, E) on **M**, denoted by $(C, \mathbf{B}, \mathbf{L}, \pi, E)^{\downarrow\mathbf{M}}$, is a Gaussian belief function with label **M**. We obtain this marginal by intersecting the certainty and belief spaces with **M** and restricting the covariance and the expectation to the new belief space. In other words, the marginal is ($C\cap\mathbf{M}$, $\mathbf{B}\cap\mathbf{M}$, **M**, $\pi|\mathbf{B}\cap\mathbf{M}$, $E|\mathbf{B}\cap\mathbf{M}$). Notice that $\pi|\mathbf{B}\cap\mathbf{M}$ and $E|\mathbf{B}\cap\mathbf{M}$ can be described as the marginalization of the representation matrix of Bel defined in Section 5.

The Shafer-Shenoy axioms are conditions under which exact local computation of marginals is possible (Shenoy and Shafer 1990). Therefore, Theorem 5 justifies that the join-tree computation works for Gaussian belief functions.

Theorem 5. Combination and marginalization of Gaussian belief functions satisfy the following conditions:

(1)(Commutativity and associativity of combination): Suppose Bel_1, Bel_2, and Bel_3 are three Gaussian belief functions. Then $Bel_1 \oplus Bel_2 = Bel_2 \oplus Bel_1$,
$Bel_1 \oplus (Bel_2 \oplus Bel_3) = (Bel_1 \oplus Bel_2)\ Bel_3$;

(2)(Consonance of marginalization): Suppose Bel = (C, **B**, **L**, π, E), and $\mathbf{L}\supset\mathbf{M}\supset\mathbf{K}$. Then

$$(Bel^{\downarrow\mathbf{M}})^{\downarrow\mathbf{K}} = Bel^{\downarrow\mathbf{K}};$$

(3)(Distributivity of marginalization over combination): Suppose $Bel_1 = (C_1, \mathbf{B}_1, L_1, \pi^1, E^1)$ and Bel_2 =

$(\mathbf{C}_2, \mathbf{B}_2, \mathbf{L}_2, \pi^2, \mathrm{E}^2)$ are two Gaussian belief functions. Then $(\mathrm{Bel}_1 \oplus \mathrm{Bel}_2)^{\downarrow \mathbf{L}_1} = \mathrm{Bel}_1 \oplus (\mathrm{Bel}_2)^{\downarrow \mathbf{L}_1 \cap \mathbf{L}_2}$.

Proof: Since interaction and addition are commutative and associative, Theorem 4 implies that the combination of Gaussian belief functions is commutative and associative. Margonalization is obviously transitive according to the definition. Thus we only need to verify the distributivity of marginalization over combination. Let $\mathrm{Bel} = \mathrm{Bel}_1 \oplus \mathrm{Bel}_2 = (\mathbf{C}, \mathbf{B}, \mathbf{L}, \pi, \mathrm{E})$ with the uncertainty space $\mathbf{F}$. Then we want to show that $(\mathbf{C} \cap \mathbf{L}_1, \mathbf{B} \cap \mathbf{L}_1, \mathbf{L}_1, \pi|\mathbf{B} \cap \mathbf{L}_1, \mathrm{E}|\mathbf{B} \cap \mathbf{L}_1) = (\mathbf{C}_1, \mathbf{B}_1, \mathbf{L}_1, \pi^1, \mathrm{E}^1) \oplus (\mathbf{C}_2 \cap \mathbf{L}_1, \mathbf{B}_2 \cap \mathbf{L}_1, \mathbf{L}_2 \cap \mathbf{L}_1, \pi^2|\mathbf{B}_2 \cap \mathbf{L}_1, \mathrm{E}^2|\mathbf{B}_2 \cap \mathbf{L}_1)$, where $\mathbf{C} = \mathbf{C}_1 \oplus \mathbf{C}_2$, $\mathbf{B} = \mathbf{B}_1 \oplus \mathbf{B}_2$, $\mathbf{L} = \mathbf{L}_1 \oplus \mathbf{L}_2$, and $\mathbf{B} = \mathbf{C} \oplus \mathbf{F}$. The identity of components in the above equation is verified as follows: $\mathbf{C} \cap \mathbf{L}_1 = (\mathbf{C}_1 \oplus \mathbf{C}_2) \cap \mathbf{L}_1 = (\mathbf{C}_1 \cap \mathbf{L}_1) \oplus (\mathbf{C}_2 \cap \mathbf{L}_1) = \mathbf{C}_1 \oplus (\mathbf{C}_2 \cap \mathbf{L}_1)$. Similarly we can examine that $\mathbf{B} \cap \mathbf{L}_1 = \mathbf{B}_1 \oplus (\mathbf{B}_2 \cap \mathbf{L}_1)$, $\mathbf{L}_1 = \mathbf{L}_1 \oplus (\mathbf{L}_2 \cap \mathbf{L}_1)$, $\mathbf{B} \cap \mathbf{L}_1 = (\mathbf{C} \cap \mathbf{L}_1) \oplus (\mathbf{F} \cap \mathbf{L}_1)$. The last equation above represents that $\mathbf{F} \cap \mathbf{L}_1$ is the uncertainty space of $(\mathbf{C}_1, \mathbf{B}_1, \mathbf{L}_1, \pi^1, \mathrm{E}^1) \oplus (\mathbf{C}_2 \cap \mathbf{L}_1, \mathbf{B}_2 \cap \mathbf{L}_1, \mathbf{L}_2 \cap \mathbf{L}_1, \pi^2|\mathbf{B}_2 \cap \mathbf{L}_1, \mathrm{E}^2|\mathbf{B}_2 \cap \mathbf{L}_1)$. Suppose the representation matrices of Bel, Bel_1, and Bel_2 are respectively Q, Q_1, and Q_2. According to Lemmas 6, 7, and 9 and Theorem 4, the representation matrix of $(\mathbf{C} \cap \mathbf{L}_1, \mathbf{B} \cap \mathbf{L}_1, \mathbf{L}_1, \pi|\mathbf{B} \cap \mathbf{L}_1, \mathrm{E}|\mathbf{B} \cap \mathbf{L}_1)$ is

$$
\begin{aligned}
Q^{\downarrow \mathbf{L}_1} = {} & \{\mathrm{Rwp}(\mathbf{F}_1 \cap \mathbf{F}_2)\,\{[\mathrm{Swp}(\mathbf{F}_1 \cap \mathbf{F}_2) \\
& \mathrm{Swp}(\mathbf{F}_1 \cap \mathbf{C}_2)Q_1]^{\downarrow \mathbf{F}_1 \cap \mathbf{F}} \oplus [\mathrm{Swp}(\mathbf{F}_1 \cap \mathbf{F}_2) \\
& \mathrm{Swp}(\mathbf{C}_1 \cap \mathbf{F}_2)Q_2]^{\downarrow \mathbf{F}_2 \cap \mathbf{F}}\}\}^{\downarrow \mathbf{L}_1} \\
= {} & \mathrm{Rwp}(\mathbf{F}_1 \cap \mathbf{F}_2)\,\{[\mathrm{Swp}(\mathbf{F}_1 \cap \mathbf{F}_2) \\
& \mathrm{Swp}(\mathbf{F}_1 \cap \mathbf{C}_2)Q_1]^{\downarrow \mathbf{F}_1 \cap \mathbf{F}} \oplus [\mathrm{Swp}(\mathbf{F}_1 \cap \mathbf{F}_2) \\
& \mathrm{Swp}(\mathbf{C}_1 \cap \mathbf{F}_2)Q_2]^{\downarrow \mathbf{F}_2 \cap \mathbf{F}}\}^{\downarrow \mathbf{L}_1} \\
= {} & \mathrm{Rwp}(\mathbf{F}_1 \cap \mathbf{F}_2)\,\{[\mathrm{Swp}(\mathbf{F}_1 \cap \mathbf{F}_2) \\
& \mathrm{Swp}(\mathbf{F}_1 \cap \mathbf{C}_2)Q_1^{\downarrow \mathbf{L}_1}]^{\downarrow \mathbf{F}_1 \cap \mathbf{F} \cap \mathbf{L}_1} \oplus \\
& [\mathrm{Swp}(\mathbf{F}_1 \cap \mathbf{F}_2)\mathrm{Swp}(\mathbf{C}_1 \cap \mathbf{F}_2)Q_2^{\downarrow \mathbf{L}_1}]^{\downarrow \mathbf{F}_2 \cap \mathbf{F} \cap \mathbf{L}_1}\} \\
= {} & \mathrm{Rwp}(\mathbf{F}_1 \cap (\mathbf{F}_2 \cap \mathbf{L}_1))\{[\mathrm{Swp}(\mathbf{F}_1 \cap (\mathbf{F}_2 \cap \mathbf{L}_1)) \\
& \mathrm{Swp}(\mathbf{F}_1 \cap \mathbf{C}_2)Q_1]^{\downarrow \mathbf{F}_1 \cap (\mathbf{F} \cap \mathbf{L}_1)} \oplus [\mathrm{Swp}(\mathbf{F}_1 \cap (\mathbf{F}_2 \cap \mathbf{L}_1)) \\
& \mathrm{Swp}((\mathbf{F}_2 \cap \mathbf{L}_1) \cap \mathbf{C}_1)Q_2^{\downarrow \mathbf{L}_1}]^{\downarrow (\mathbf{F}_2 \cap \mathbf{L}_1) \cap (\mathbf{F} \cap \mathbf{L}_1)}\} \\
= {} & \mathrm{Rwp}(\mathbf{F}_1 \cap (\mathbf{F}_2 \cap \mathbf{L}_1))\,\{[\mathrm{Swp}(\mathbf{F}_1 \cap (\mathbf{F}_2 \cap \mathbf{L}_1)) \\
& \mathrm{Swp}(\mathbf{F}_1 \cap (\mathbf{C}_2 \cap \mathbf{L}_1))Q_1]^{\downarrow \mathbf{F}_1 \cap (\mathbf{F} \cap \mathbf{L}_1)} \oplus \\
& [\mathrm{Swp}(\mathbf{F}_1 \cap (\mathbf{F}_2 \cap \mathbf{L}_1))\mathrm{Swp}((\mathbf{F}_2 \cap \mathbf{L}_1) \cap \mathbf{C}_1) \\
& Q_2^{\downarrow \mathbf{L}_1}]^{\downarrow (\mathbf{F}_2 \cap \mathbf{L}_1) \cap (\mathbf{F} \cap \mathbf{L}_1)}\}.
\end{aligned}
$$

The last term is just the representation matrix of combination $(\mathrm{C}_1, \mathrm{B}_1, \mathrm{L}_1, \Pi^1, \mathrm{E}^1) \oplus (\mathrm{C}_2 \cap \mathrm{L}_1, \mathrm{B}_2 \cap \mathrm{L}_1, \mathrm{L}_2 \cap \mathrm{L}_1, \Pi^2|\mathrm{B}_2 \cap \mathrm{L}_1, \mathrm{E}^2|\mathrm{B}_2 \cap \mathrm{L}_1)$.

ACKNOWLEDGMENT

I am indebted to Professor Shafer, under whom it has been my privilege to study Dempster-Shafer theory of belief functions. This paper benefited from many discussions with Professor Shafer and the foundational work done by Professors Dempster (Dempster 1990) and Shafer (Shafer 1992). The research was supported by a research assistantship from the Harper Fund.

REFERENCES

DEMPSTER, A. P. (1990a) Construction and local computation aspects of network belief functions. In Oliver R. M. and Smith J. Q. (eds.) Influence Diagrams, Belief Nets, and Decision Analysis, Chichester: John Wiley and Sons.

DEMPSTER, A. P. (1990b) Normal Belief Functions and the Kalmam Filter. Research Report, Department of Statistics, Harvard University.

DEMPSTER, A. P. (1969) *Elements of Continuous Multivariate Analysis.* Massachusetts: Addison-Wesley.

KONG, A. (1986) *Multivariate Belief Functions and Graphical Models.* Ph.D. Thesis, Department of Statistics, Harvard University.

SHAFER, G. (1976) *A Mathematical Theory of Evidence.* Princeton: Princeton University Press.

SHAFER, G. (1992) *A Note on Dempster's Gaussian Belief Functions.* Working paper, University of Kansas.

SHAFER, G., SHENOY, P. P., AND MELLOULI, K. (1987) Propagating belief functions in qualitative Markov Trees. *International Journal of Approximate Reasoning*, 3, 383-411.

SHENOY, P. P. AND SHAFER, G. (1990) Axioms for Probability and Belief-Function Propagation. *Uncertainty in Artificial Intelligence*, 4, 169-198.

LAURITZEN, S. L. AND SPIEGELHALTER, D. J. (1988) Local computation with probabilities on graphical structures and their application to expert systems (with discussion). Journal of the Royal Statistical Society, Series B, 50, 157–224.

LIU, L. P. (1993) Local Compution of Gaussian Belief Functions. Working Paper No. 255, School of Business, University of Kansas.