



Dynamic Estimation of Linear Systems Constrained by Bounds

André Monin

► To cite this version:

André Monin. Dynamic Estimation of Linear Systems Constrained by Bounds. IEEE Transactions on Signal Processing, 2009, 57 (10), pp.4095 - 4099. 10.1109/TSP.2009.2021697 . hal-02915725

HAL Id: hal-02915725

<https://laas.hal.science/hal-02915725>

Submitted on 17 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dynamic Estimation of Linear Systems Constrained by Bounds

André Monin

LAAS-CNRS; Université de Toulouse; 7 avenue du Colonel Roche - 31077 Toulouse Cedex France.

EDICS: SSP-FILT

Abstract—This article deals with the minimum variance estimation of a Gaussian process constrained by bounds. A special truncated Gaussian probability is shown to be fairly well adapted to this filtering scheme as its set is linearly closed with respect to convolution and multiplication operations.

LIST OF FIGURES

1	Example of the distortion induced by the truncation.	2
2	Range estimation	4
3	Speed estimation	4

Index Terms—Truncated Gaussian pdf.

I. INTRODUCTION.

In some dynamic system estimation issues, the state, or some component of the state, may be assumed to be constrained by bounds. In this article, we address the problem of state estimation when one component of the state is constrained to be positive. This work follows the previous exact results obtained in the static case, which is in the parameter estimation case [8].

Some previous works in the dynamic case are discussed here. In [11], [4] and [6], the authors use first an unconstrained Kalman filter (or extended Kalman filter) and then, at each step, if the unconstrained estimator is outside the feasible region, it is projected to the boundary of the feasible region without any change in the covariance matrix of the Kalman filter (except in [6]). In [14], the author suggests introducing a nonlinearity in the system to represent the constraint. Note that if the indicator function is chosen to stand for the nonlinear function, this approach is close to the one used in [11]. In [7], a smoothing approach is used to prune the negative values. It may be seen as an extension of the approach

used in [11]. In [10], the author extends the pseudo-measurement approach, used in the equality constraint case [13], to the inequality case. Recall that the pseudo-measurement approach may lead to singularity of the filtering covariance matrix. Often, this singularity has to be artificially regularized in the nonlinear case [3]. In [15], the authors compute a maximum likelihood estimator under constraints using the Lagrangian multiplier tool. Note that it seems that the authors did not observe that this approach is close to the pseudo-measurement approach [13]. Finally, in [9], the authors employ online optimization using a moving horizon approximation and in [2], the author uses classical normalized truncated Gaussian probability density functions (pdf) which are approximated by random sampling. The main drawback of these two last approaches lies in their heavy computational costs.

Here, we suggest the use of a new pdf named Mirrored Gaussian pdf (MGP). It is shown that the MGP set is linearly closed with respect to convolution and multiplication operations. The Gaussian pdf properties are then exploited to derive a filtering algorithm based on Kalman filtering or extended Kalman filtering (EKF) nearby the so-called Gaussian sum filter (GSF) [12]. As it is the case for GSF, one observes an increase in the number of MGP devoted to represent the optimal solution. This exponential growth can be reduced using the same kind of heuristics as those used in [12] and [1].

Let us consider a linear dynamic system $x_t \in \mathbb{R}^n$ whose evolution is involved in the classical linear equations:

$$x_t = Fx_{t-1} + w_t \quad (1a)$$

$$y_t = Hx_t + v_t \quad (1b)$$

where w_t and v_t are independent white noises and where y_t stands for the output. If the noise w_t pdf is Gaussian with variance $Q = E[w_t w_t^T]$, the transition pdf of the state can be written

$$p(x_t | x_{t-1}) = \Gamma(x_t - Fx_{t-1}, Q) \quad (2)$$

where Γ stands for the Gaussian pdf

$$\Gamma(x, Q) \triangleq \frac{1}{\sqrt{(2\pi)^n |Q|}} \exp\left(-\frac{1}{2}x^T Q^{-1}x\right)$$

Suppose now that one wants to make sure that the first component of x_t , say $[x_t]_1$, remains positive for all x_{t-1} such that $[x_{t-1}]_1 > 0$. The first idea would be to consider the truncated Gaussian pdf defined by

$$p(x_t|x_{t-1}) = \alpha(x_{t-1}) \mathbb{I}_{[x_t]_1 \geq 0} \Gamma(x_t - Fx_{t-1}, Q)$$

where $\mathbb{I}_{[x>0]}$ stands for the indicator function of the set $\{x > 0\}$ and where $\alpha(x_{t-1})$ is a scaling factor, depending obviously on the Gaussian pdf mean Fx_{t-1} . Unfortunately, the expression of $\alpha(x_{t-1})$ is a little complicated and does not allow simple filtering computation, except for the static case (parameter estimation [8]). A better choice is accepting a minor distortion of the Gaussian pdf near zero (see Figure 1). This may be achieved by defining the MGP as follows:

$$\tilde{\Gamma}(x, \bar{x}, P) \triangleq \mathbb{I}_{[x]_1 \geq 0} (\Gamma(x - \bar{x}, P) + \Gamma(x + \bar{x}, P)) \quad (3)$$

It is easy to show that this pdf is really a normalized pdf

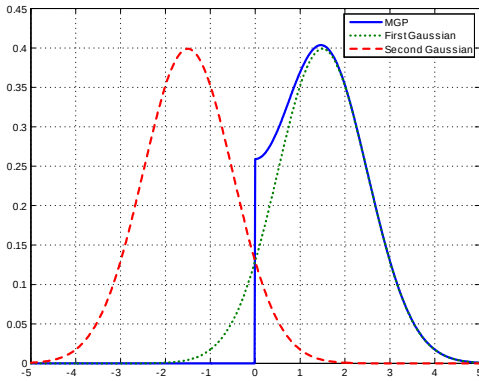


Fig. 1. Example of the distortion induced by the truncation.

(see Lemma 3 in the Appendix). Thus, let us replace the transition pdf, Expression (2), with the following MGP

$$p(x_t|x_{t-1}) = \tilde{\Gamma}(x_t, Fx_{t-1}, Q)$$

In the same manner, assume that the first component of the output y_t is positive. One may define again the output pdf as a MGP as follows:

$$p(y_t|x_t) = \tilde{\Gamma}(y_t, Hx_t, R) \quad (4)$$

II. OPTIMAL FILTERING

A. The linear case

The MGP introduced in this paper has important properties. Indeed, the MGP set is closed with convolution and linearly closed with multiplication (see Appendix I). Recall that the optimal filtering goal is to compute the so-called *a posteriori* pdf of the state $p(x_t|y_{0:t})$ where $y_{0:t} \triangleq \{y_0, \dots, y_t\}$. Therefore, if the initial pdf of the state is a MGP, then the *a posteriori* pdf is a linear combination of MGPs.

Theorem 1: If at time $t-1$ the *a posteriori* pdf of the state is a weighted MGP sum with N_{t-1} terms as follows:

$$p(x_{t-1}|y_{0:t-1}) = \sum_{i=1}^{N_{t-1}} \rho_{t-1}^i \tilde{\Gamma}(x_{t-1}, \hat{x}_{t-1|t-1}^i, P_{t-1|t-1}^i) \quad (5)$$

then, at time t , the *a posteriori* pdf is again a weighted MGP sum with $2 \times N_{t-1}$ terms defined by

$$\begin{aligned} p(x_t|y_{0:t}) &= \sum_{i=1}^{N_{t-1}} \rho_{t-1}^{i,1} \tilde{\Gamma}(x_t, \hat{x}_{t|t}^{i,1}, P_{t|t}^i) + \sum_{i=1}^{N_{t-1}} \rho_{t-1}^{i,2} \tilde{\Gamma}(x_t, \hat{x}_{t|t}^{i,2}, P_{t|t}^i) \end{aligned}$$

where

$$\hat{x}_{t|t-1}^i = F\hat{x}_{t-1|t-1}^i \quad (6a)$$

$$P_{t|t-1}^i = FP_{t-1|t-1}^i F^T + Q \quad (6b)$$

$$\hat{x}_{t|t}^{i,1} = \hat{x}_{t|t-1}^i + K_t^i (y_t - H\hat{x}_{t|t-1}^i) \quad (7a)$$

$$\hat{x}_{t|t}^{i,2} = -\hat{x}_{t|t-1}^i + K_t^i (y_t + H\hat{x}_{t|t-1}^i) \quad (7b)$$

$$\Sigma_{t|t-1}^i = HP_{t|t-1}^i H^T + R \quad (8a)$$

$$K_t^i = P_{t|t-1}^i H^T (\Sigma_{t|t-1}^i)^{-1} \quad (8b)$$

$$P_{t|t}^i = P_{t|t-1}^i - P_{t|t-1}^i H^T (\Sigma_{t|t-1}^i)^{-1} HP_{t|t-1}^i \quad (8c)$$

$$\tilde{\rho}_t^{i,1} = \rho_{t-1}^i \mathbb{I}_{[y_t]_1 > 0} \Gamma(y_t - H\hat{x}_{t|t-1}^i, \Sigma_{t|t-1}^i) \quad (9a)$$

$$\tilde{\rho}_t^{i,2} = \rho_{t-1}^i \mathbb{I}_{[y_t]_1 > 0} \Gamma(y_t + H\hat{x}_{t|t-1}^i, \Sigma_{t|t-1}^i) \quad (9b)$$

$$\rho_t^{i,1} = \frac{\tilde{\rho}_t^{i,1}}{\sum_{i=1}^{N_{t-1}} (\tilde{\rho}_t^{i,1} + \tilde{\rho}_t^{i,2})}$$

$$\rho_t^{i,2} = \frac{\tilde{\rho}_t^{i,2}}{\sum_{i=1}^{N_{t-1}} (\tilde{\rho}_t^{i,1} + \tilde{\rho}_t^{i,2})}$$

Proof: The so-called Chapman-Kolmogorov theorem applied to Equation (5) leads to

$$\begin{aligned} p(x_t|y_{0:t-1}) &= \int_{\mathbb{R}^n} p(x_t|x_{t-1}) p(x_{t-1}|y_{0:t-1}) dx_{t-1} \\ &= \sum_{i=1}^{N_{t-1}} \rho_{t-1}^i \int_{\mathbb{R}^n} \tilde{\Gamma}(x_t, Fx_{t-1}, Q) \\ &\quad \times \tilde{\Gamma}(x_{t-1}, \hat{x}_{t-1|t-1}^i, P_{t-1|t-1}^i) dx_{t-1} \end{aligned}$$

Corollary 5 allows one to rewrite this expression as follows:

$$p(x_t|y_{0:t-1}) = \sum_{i=1}^{N_{t-1}} \rho_{t-1}^i \tilde{\Gamma}(x_t, \hat{x}_{t|t-1}^i, P_{t|t-1}^i)$$

where

$$\begin{aligned} \hat{x}_{t|t-1}^i &= F\hat{x}_{t-1|t-1}^i \\ P_{t|t-1}^i &= FP_{t-1|t-1}^i F^T + Q \end{aligned}$$

The correction is then obtained by using the Bayes rule, that is, omitting the scaling term:

$$\begin{aligned} p(x_t|y_{0:t}) &\propto p(y_t|x_t) p(x_t|y_{0:t-1}) \\ &= \sum_{i=1}^{N_{t-1}} \rho_{t-1}^i \tilde{\Gamma}(y_t, Hx_t, R) \tilde{\Gamma}(x_t, \hat{x}_{t|t-1}^i, P_{t|t-1}^i) \end{aligned}$$

Lemma 4 then leads to the result. $\blacksquare$

Clearly, the number of MGP needed for the representation of the filtering solution grows exponentially with time, as it is the case in Gaussian sum filtering ([12] [1]). Note that the MGP number increase is due to the correction step while the Gaussian number pdf increase is due to the prediction step when using the GSF technique.

As in the GSF case, one has to use approximations to reduce the increase in MGP number. The simplest technique consists of maintaining only the $N_{\max}$ MGPs with greatest weights ρ_t^i and pruning the other MGPs, $N_{\max}$ being the number of MGP well-matched with the computational power allocated to the application.

B. Computing the estimator

If at step t , the *a posteriori* pdf is

$$p(x_t|y_{0:t}) = \sum_{i=1}^{N_t} \rho_t^i \tilde{\Gamma}(x_t, \hat{x}_{t|t}^i, P_{t|t}^i)$$

Then the minimum variance estimator of the constrained component of the state $[x]_1$ has the following expression

(see Appendix I, Proposition 8)

$$\begin{aligned} [\hat{x}_{t|t}]_1 &= \sum_{i=1}^{N_t} \rho_t^i \left(2 [P_{t|t}^i]_1^1 \Gamma\left(\left[\hat{x}_{t|t}^i\right]_1, [P_{t|t}^i]_1^1\right) \right. \\ &\quad \left. + \operatorname{erf}\left(\left[\hat{x}_{t|t}^i\right]_1 / [P_{t|t}^i]_1^1\right) \left[\hat{x}_{t|t}^i\right]_1 \right) \end{aligned}$$

where $[\bar{x}]_1$ stands for the first component of $\bar{x}$ and $[P]_1^1$ stands for the variance of the first component of the covariance matrix P . It is based on the so-called erf function defined by

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-\xi^2) d\xi$$

Note that if $[\hat{x}_{t|t}^i]_1$ is far from the bound (zero), then the erf function is quite near unity and $\Gamma\left(\left[\hat{x}_{t|t}^i\right]_1, [P_{t|t}^i]_1^1\right)$ is near zero. Therefore, $[\hat{x}_{t|t}]_1$ is almost equal to the mean of the unconstrained filter, say $\sum_{i=1}^{N_t} \rho_t^i [\hat{x}_{t|t}^i]_1$.

On the other hand, it is easy to check that the other component means $[\hat{x}_{t|t}]_{k,k \neq 1}$ are equal to $\sum_{i=1}^{N_t} \rho_t^i [\hat{x}_{t|t}^i]_k$, the same values that in the unconstrained case.

C. The nonlinear case

Suppose now that the dynamic system is nonlinear:

$$\begin{aligned} x_t &= f(x_{t-1}) + w_t \\ y_t &= h(x_t) + v_t \end{aligned}$$

In this case, a very popular approximation is based on the linear approximation of functions f and h and leads to the so-called EKF. In our case, such an approximation is still relevant provided by the function h to be an odd function.

Theorem 2: If the function h is an odd function and if at time $t-1$ the *a posteriori* pdf of the state is a weighted MGP sum with N_{t-1} terms, Expression (5), then at time t the *a posteriori* pdf can be approximated by a weighted MGP sum with $2 \times N_{t-1}$ terms defined by

$$p(x_t|y_{0:t}) \simeq \sum_{i=1}^{N_{t-1}} \rho_t^{i,1} \tilde{\Gamma}(x_t, \hat{x}_{t|t}^{i,1}, P_{t|t}^i) + \sum_{i=1}^{N_{t-1}} \rho_t^{i,2} \tilde{\Gamma}(x_t, \hat{x}_{t|t}^{i,2}, P_{t|t}^i)$$

where all the terms of this expression are obtained by replacing in Theorem 1 Kalman filters by EKFs.

Proof: The proof is almost the same as those of Theorem 1 using Lemma 6 and Corollary 7. Just replace $F\hat{x}_{t-1|t-1}^i$ with $f(\hat{x}_{t-1|t-1}^i)$ in Expression (6a), F with

the Jacobian of f evaluated at $\hat{x}_{t-1|t-1}^i$ in Expression (6b), $H\hat{x}_{t|t-1}^i$ with $h(\hat{x}_{t|t-1}^i)$ in Expressions (7) and (9) and finally H with the Jacobian of h evaluated at $\hat{x}_{t|t-1}^i$ in Expression (8). ■

III. SIMULATION RESULTS

To illustrate the relevance of such a filtering approach, let us consider the stereovision issue in a very simple case, that is, a mono-dimensional stereovision system. Recall that the classical range estimation of a salient point is achieved via disparity measurement, that is, the difference between the pixel location corresponding to the salient point on the two camera image planes. This disparity d_t is then related to the range r_t as follows:

$$d_t = \frac{bf}{r_t}$$

where b stands for the distance between the two cameras and f is the focal distance of the lens. If f is expressed in pixel units, considering the measurement error equal to one pixel, the measurement may be modeled as follows:

$$y_t = \frac{bf}{r_t} + v_t$$

where $f = N/2$, N is the total number of pixels, and where v_t is a Gaussian white noise with variance equal to unity. Recall that if the measurement $y_t < 0$, the value is changed in $y_t = 0$.

Let us consider that the salient point speed (s_t) moves according to a first order dynamic system. The range (r_t) evolution may be represented by the following equations:

$$\begin{aligned} r_t &= r_{t-1} + s_{t-1}\Delta t \\ s_t &= ks_{t-1} + w_t \end{aligned}$$

where $k = 0.95$ refers to a friction factor and w_t is Gaussian white noise with variance Q . Δt stands for the sampling period.

We then apply Theorem 2 to achieve the filtering of x_t (it is clear that the output function is an odd function). The growth of the MGP number is treated by pruning MGPs with lowest weights and the total number of MGPs is fixed to $N_{\max} = 2^3 = 8$. The distance between cameras is fixed at $b = 0.2\text{m}$ and the pixel number at $N = 640$. The true initial range is equal to $r_0 = 2\text{m}$ and the initial pdf is defined by $p(r_0) = \tilde{\Gamma}(r_0, \bar{r}_0, P_0)$ with $\bar{r}_0 = c/y_0$ and $\sqrt{P_0} = 60\text{m}$. The measurement frequency is 1Hz ($\Delta t = 1\text{s}$). The standard deviation of w_t is set to $10 \times \sqrt{\Delta t}$.

Our filtering algorithm (MEKF) has been compared with the classical EKF, the algorithm including the projection utility suggested in [11] and [6] (CEKF) and

the pseudo-measurement approach (PEKF) suggested in [15]. Note that, for this example, the approaches in [14] and [7] lead to the same results than the CEKF approach. A typical example of range and speed estimations is shown Figures 2 and 3 respectively.

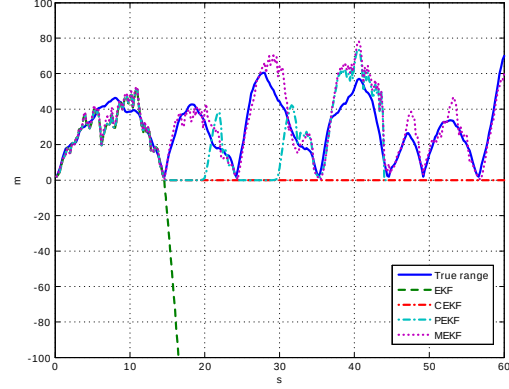


Fig. 2. Range estimation

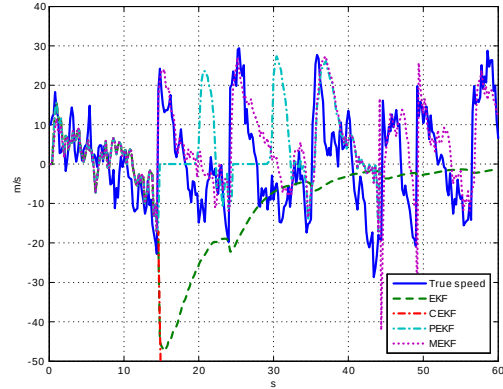


Fig. 3. Speed estimation

Clearly, when the true range comes near zero, the EKF range estimator diverges when it locks on a negative value. It is then unable to go back to a positive estimate as it uses the Jacobian of $1/r$, which has the wrong sign. In the same situation (true range near zero), the CEKF is unable to get away from zero as the speed estimation has the wrong sign. Finally, the PEKF range estimator remains zero until the real range reaches a value far from zero. On the other hand, our algorithm seems to be always efficient, even near zero. Note that when the true range is far from zero and before any divergence, all filters are equivalent as they come down to the unconstrained EKF.

IV. CONCLUSION

A new pdf devoted to perform positive variables has been proposed here. This pdf appears to be well-adapted to filtering when one component of the variable to be estimated is assumed to be positive. A new filtering algorithm based on these pdfs has been proposed compared with EKF in an academic example. One should imagine an extension to constraints defined by several bounds (an interval, for instance) provided to grow the Gaussian pdf number devoted to represent the MGP.

APPENDIX I

TRUNCATED GAUSSIAN FORMULA

Lemma 3: The MGP defined by Expression (3) is a normalized pdf.

Proof: One has

$$\begin{aligned} \int_{\mathbb{R}^n} \tilde{\Gamma}(x, \bar{x}, P) dx &= \int_{\mathbb{R}^+} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \Gamma(x - \bar{x}, P) dx \\ &\quad + \int_{\mathbb{R}^+} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \Gamma(x + \bar{x}, P) dx \\ &= \int_{\mathbb{R}^+} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \Gamma(x - \bar{x}, P) dx \\ &\quad + \int_{\mathbb{R}^-} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \Gamma(x - \bar{x}, P) dx \\ &= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \Gamma(x - \bar{x}, P) dx = 1 \end{aligned}$$

using the fact that $\Gamma(x, P) = \Gamma(-x, P)$. ■

Lemma 4: The product of two truncated Gaussian pdf is a linear combination of truncated Gaussian pdfs:

$$\begin{aligned} &\tilde{\Gamma}(y, Hx, R) \tilde{\Gamma}(x, \mu, P) \\ &= \mathbb{I}_{[y^1 \geq 0]} \Gamma(y - H\mu, \Sigma) \tilde{\Gamma}(x, \mu + K(y - H\mu), B) \\ &\quad + \mathbb{I}_{[y^1 \geq 0]} \Gamma(y + H\mu, \Sigma) \tilde{\Gamma}(x, -\mu + K(y + H\mu), B) \end{aligned}$$

where

$$\begin{aligned} \Sigma &= HPH^T + R \\ K &= PH^T \Sigma^{-1} \\ B &= P - PH^T \Sigma^{-1} HP \end{aligned}$$

Proof: Let us develop the truncated Gaussian pdf product as follows

$$\begin{aligned} &\tilde{\Gamma}(y, Hx, R) \tilde{\Gamma}(x, \mu, P) \\ &= \mathbb{I}_{[y^1 \geq 0]} \mathbb{I}_{[x^1 \geq 0]} (\Gamma(y - Hx, R) \Gamma(x - \mu, P) \\ &\quad + \Gamma(y + Hx, R) \Gamma(x + \mu, P) \\ &\quad + \Gamma(y - Hx, R) \Gamma(x + \mu, P) \\ &\quad + \Gamma(y + Hx, R) \Gamma(x - \mu, P)) \end{aligned} \quad (11)$$

Recall the classical Gaussian product formula usually used to compute the Kalman filter [5]

$$\begin{aligned} &\Gamma(y - Hx, R) \Gamma(x - \mu, P) \\ &= \Gamma(y - H\mu, \Sigma) \Gamma(x - (\mu + K(y - H\mu)), B) \end{aligned}$$

where

$$\begin{aligned} \Sigma &= HPH^T + R \\ K &= PH^T \Sigma^{-1} \\ B &= P - PH^T \Sigma^{-1} HP \end{aligned}$$

Let us take into account the first two Gaussian pdf products of Expression (11). Note that the second term is derived from the first one by replacing $y \rightarrow -y$ and $\mu \rightarrow -\mu$. Thus, one has

$$\begin{aligned} &\Gamma(y - Hx, R) \Gamma(x - \mu, P) + \Gamma(y + Hx, R) \Gamma(x + \mu, P) \\ &= \Gamma(y - H\mu, \Sigma) \Gamma(x - (\mu + K(y - H\mu)), B) \\ &\quad + \Gamma(x + (\mu + K(y - H\mu)), B) \end{aligned}$$

In the same manner, taking into account the last two components of Expression (11) leads to

$$\begin{aligned} &\Gamma(y - Hx, R) \Gamma(x + \mu, P) + \Gamma(y + Hx, R) \Gamma(x - \mu, P) \\ &= \Gamma(y + H\mu, \Sigma) \Gamma(x - (\mu + K(y + H\mu)), B) \\ &\quad + \Gamma(x + (\mu + K(y + H\mu)), B) \end{aligned}$$

Reporting these two expressions in Equation (11) leads to the result. ■

Corollary 5: The convolution of two MGP is a MGP

$$\int_{\mathbb{R}^n} \tilde{\Gamma}(y, Hx, R) \tilde{\Gamma}(x, \mu, P) dx = \tilde{\Gamma}(y, H\mu, \Sigma)$$

with

$$\Sigma = HPH^T + R$$

Proof: The integration of Expression (10) leads directly to the result, each MGP $\tilde{\Gamma}(x, \mu + K(y - H\mu), B)$ and $\tilde{\Gamma}(x, -\mu + K(y + H\mu), B)$ being normalized. ■

Lemma 6: If h is an odd function, the product of MGP $\tilde{\Gamma}(y, h(x), R) \tilde{\Gamma}(x, \mu, P)$ can be approximated as follows

$$\begin{aligned} &\tilde{\Gamma}(y, h(x), R) \tilde{\Gamma}(x, \mu, P) \\ &\simeq \mathbb{I}_{[y^1 > 0]} \Gamma(y - h(\mu), \Sigma(\mu)) \\ &\quad \times \tilde{\Gamma}(x, \mu + K(\mu)(y - h(\mu)), B(\mu)) \\ &\quad + \mathbb{I}_{[y^1 \geq 0]} \Gamma(y + h(\mu), \Sigma(\mu)) \\ &\quad \times \tilde{\Gamma}(x, -\mu + K(\mu)(y + h(\mu)), B(\mu)) \end{aligned} \quad (12)$$

where

$$\Sigma(\mu) = H(\mu)PH(\mu)^T + R \quad (13a)$$

$$K(\mu) = PH(\mu)^T \Sigma^{-1} \quad (13b)$$

$$B(\mu) = P - PH(\mu)^T \Sigma^{-1} H(\mu)P \quad (13c)$$

$H(\mu)$ being the Jacobian of the function h .

Proof: The development of the product is similar to that of Expression (11), just replacing Hx with $h(x)$. The linearization of $h(x)$ at the neighborhood of μ leads to

$$\begin{aligned} & \Gamma(y - h(x), R) \Gamma(x - \mu, P) \\ & \simeq \Gamma(y - h(\mu), \Sigma(\mu)) \\ & \quad \times \Gamma(x - (\mu + K(\mu)(y - h(\mu))), B(\mu)) \end{aligned} \quad (14)$$

where Σ , K and B are defined by Expression (13). In the same manner, the second product of Expression (11) becomes

$$\begin{aligned} & \Gamma(y + h(x), R) \Gamma(x + \mu, P) \\ & \simeq \Gamma(y + h(-\mu), \Sigma(-\mu)) \\ & \quad \times \Gamma(x - (-\mu + K(-\mu)(y - h(-\mu))), B(-\mu)) \end{aligned} \quad (15)$$

after linearization of $h(x)$ at the neighborhood of $-\mu$. The function h being odd, the Jacobian $H(\mu)$ is even and $\Sigma(-\mu) = \Sigma(\mu)$, $K(-\mu) = K(\mu)$ and $B(\mu) = B(-\mu)$. Equations (14) and (15) can then be combined in

$$\Gamma(y - h(\mu), \Sigma(\mu)) \tilde{\Gamma}(x, \mu + K(\mu)(y - h(\mu)), B(\mu))$$

The same computation is applied to the second product. ■

Corollary 7: If h is an odd function, one has

$$\int_{\mathbb{R}^n} \tilde{\Gamma}(y, h(x), R) \tilde{\Gamma}(x, \mu, P) dx \simeq \tilde{\Gamma}(y, h(\mu), \Sigma(\mu))$$

with

$$\Sigma(\mu) = H(\mu) P H(\mu)^T + R$$

where H is the Jacobian of h .

Proof: The integration of Expression (12) leads directly to the result. ■

Proposition 8: The mean $[\hat{x}]_1$ of the constrained state $[x]_1$ with pdf $\tilde{\Gamma}(x, \bar{x}, P)$ has the following expression

$$[\hat{x}]_1 = 2[P]_1^1 \Gamma([\bar{x}]_1, [P]_1^1) + \text{erf}\left([\bar{x}]_1 / \sqrt{2[P]_1^1}\right) [\bar{x}]_1$$

Proof: The expectation of $[x]_1$ is defined by

$$\begin{aligned} [\hat{x}]_1 & \triangleq \int_{\mathbb{R}^+} [x]_1 \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \Gamma(x - \bar{x}, P) dx \\ & \quad + \int_{\mathbb{R}^+} [x]_1 \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \Gamma(x + \bar{x}, P) dx \\ & = \int_{\mathbb{R}^+} [x]_1 \Gamma([x]_1 - [\bar{x}]_1, [P]_1^1) d[x]_1 \\ & \quad + \int_{\mathbb{R}^+} [x]_1 \Gamma([x]_1 + [\bar{x}]_1, [P]_1^1) d[x]_1 \end{aligned}$$

$\Gamma([x]_1 - [\bar{x}]_1, [P]_1^1)$ being the marginal pdf of $[x]_1$. Recall that

$$\int_0^\infty x \Gamma(x - \bar{x}, P) dx = \frac{1}{\sqrt{2\pi P}} \int_0^\infty x \exp\left(-\frac{1}{2}(x - \bar{x})^2 / P\right) dx$$

Noting that

$$\begin{aligned} x \exp\left(-\frac{1}{2}(x - \bar{x})^2 / P\right) & = \bar{x} \exp\left(-\frac{1}{2}(x - \bar{x})^2 / P\right) \\ & \quad - P \frac{d}{dx} \left(\exp\left(-\frac{1}{2}(x - \bar{x})^2 / P\right) \right) \end{aligned}$$

an integration by part leads to

$$\begin{aligned} & \int_0^\infty x \Gamma(x - \bar{x}, P) dx \\ & = \frac{1}{\sqrt{2\pi P}} \bar{x} \int_0^\infty \exp\left(-\frac{1}{2}(x - \bar{x})^2 / P\right) dx \\ & \quad - \frac{P}{\sqrt{2\pi P}} \left[\exp\left(-\frac{1}{2}(x - \bar{x})^2 / P\right) \right]_0^\infty \\ & = P \Gamma(\bar{x}, P) + \frac{1}{\sqrt{2\pi P}} \bar{x} \int_0^\infty \exp\left(-\frac{1}{2}(x - \bar{x})^2 / P\right) dx \\ & = P \Gamma(\bar{x}, P) + \frac{1}{2} \left(1 + \text{erf}\left(\bar{x} / \sqrt{2P}\right) \right) \bar{x} \end{aligned}$$

Therefore

$$\begin{aligned} [\hat{x}]_1 & = 2[P]_1^1 \Gamma([\bar{x}]_1, [P]_1^1) \\ & \quad + \frac{1}{2} \left(1 + \text{erf}\left([\bar{x}]_1 / \sqrt{2[P]_1^1}\right) \right) [\bar{x}]_1 \\ & \quad - \frac{1}{2} \left(1 - \text{erf}\left([\bar{x}]_1 / \sqrt{2[P]_1^1}\right) \right) [\bar{x}]_1 \end{aligned}$$

which leads to the result. ■

REFERENCES

- [1] D. L. Alspach and H. W. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximation," *IEEE Transactions on Automatic Control*, vol. 17, no. 4, pp. 439–448, 1972.
- [2] J.-M. Brankart, "Optimal nonlinear filtering and smoothing assuming truncated Gaussian probability distributions in a reduced dimension space of adaptive size," MEOM-LEGI, Grenoble, France, Tech. Rep., 2006.
- [3] J. De Geeter, H. Van Brussel, and J. De Schutter, "A smoothly constrained Kalman filter," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 10, pp. 1171–1177, 1997.
- [4] E. Haseltine and J. Rawlings, "A critical evaluation of extended Kalman filtering and moving horizon estimation," Texas-Wisconsin Modeling and Control Consortium (TWMCC), Tech. Rep., 2003.
- [5] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York and London: Academic Press, 1970, vol. 64.
- [6] R. Kandepe, I. L., and B. Foss, "Constrained state estimation using the unscented Kalman filter," in *16th Mediterranean Conference on Control and Automation*, Ajaccio, France, 2008, pp. 1453–1458.

- [7] D. Massicote, R. Z. Morawski, and A. Barwicz, "Incorporation of a positivity constraint into a Kalman-filter-based algorithm for correction of spectrometric data," *IEEE Transactions on Instrumentation and Measurement*, vol. 44, no. 1, pp. 2–7, 1995.
- [8] A. Monin and G. Salut, "Minimum variance estimation of parameters constrained by bounds," *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 246–248, 2001.
- [9] C. Rao, J. Rawlings, and D. Mayne, "Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 246–258, 2003.
- [10] P. Richards, "Constrained Kalman filtering using pseudo-measurements," in *IEE Colloquium on Algorithms for Target Tracking*, 1995, pp. 75–79.
- [11] D. Simon and D. L. Simon, "Kalman filtering with inequality constraints for turbofan engine health estimation," *IEE Proceedings - Control Theory and Applications*, vol. 153, no. 2, pp. 371–378, 2006.
- [12] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, pp. 465–479, 1971.
- [13] M. Tahk and J. Speyer, "Target tracking subject to kinematic constraints," *IEEE Transactions on Automatic Control*, vol. 35, no. 3, 1990.
- [14] J. K. Tugnait, "Constrained signal restoration via iterated extended Kalman filtering," *IEEE transactions on Acoustics, Speech and Signal Processing*, vol. 33, no. 2, pp. 472–475, 1985.
- [15] S. Ungarala, E. Dolence, and K. Li, "Constrained extended Kalman filter for nonlinear state estimation," in *8th International IFAC Symposium on Dynamics and Control of Process Systems*, Cancun, Mexico, 2007, pp. 63–68.