



LUND UNIVERSITY

Log-concave Observers

Henningsson, Toivo; Åström, Karl Johan

Published in:

17th International Symposium on Mathematical Theory of Networks and Systems, 2006

2006

[Link to publication](#)

Citation for published version (APA):

Henningsson, T., & Åström, K. J. (2006). Log-concave Observers. In *17th International Symposium on Mathematical Theory of Networks and Systems, 2006*

Total number of authors:

2

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Log-concave Observers

Toivo Henningsson, Karl Johan Åström

Abstract—The Kalman filter is the optimal state observer in the case of linear dynamics and Gaussian noise. In this paper, the observer problem is studied when process noise and measurements are generalized from Gaussian to log-concave. This generalization is of interest for example in the case where observations only give information that the signal is in a given range. It turns out that the optimal observer preserves log-concavity. The concept of strong log-concavity is introduced and two new theorems are derived to compute upper bounds on optimal observer covariance in the log-concave case. The theory is applied to a system with threshold based measurements, which are log-concave but far from Gaussian.

Keywords—Observers, Stochastic systems, Filtering and estimation

I. INTRODUCTION

The Kalman filter (see [1], [2]) is one of the most widely used schemes for state estimation from noisy measurements. It is optimal for linear measurements and Gaussian noise, but it is often applied in a more general setting. Although the Extended Kalman filter (see [3]) often works well in practice, sometimes it does not, and it is in general not easy to see how altered conditions change the observer problem.

In this paper, a particular generalization is investigated where measurements and noise are allowed to be log-concave (see [4], [5], [6], [7]). The model of log-concave measurements is applicable in many instances where the assumption of independent additive measurement noise is too limited, for instance with heavy quantization, or with the problem of event based sampling discussed in [8].

Within this framework, the problem of moving horizon ML/MAP estimation becomes a convex optimization problem, see [9]. This paper will however focus on the covariance of the Bayesian Observer, which is investigated and compared with the Kalman filter.

Strongly log-concave functions are introduced as a means to quantify observer properties. Two new theorems are applied to derive upper bounds on optimal observer covariance.

Toivo Henningsson and Karl Johan Åström are with Department of Automatic Control, Faculty of Engineering, Lund University, Box 118, SE 221 00 Lund, Sweden. Email: toivo.henningsson@control.lth.se, karl_johan.astrom@control.lth.se

It turns out that the observer problem is still quite well behaved so that, especially with some insight gained in the analysis, a Kalman filter might often be usable for this more general measurement setting. For a more thorough treatment, see [10].

The paper is organized as follows. A motivating example is presented in section II. The notion of log-concavity is introduced in section III, where we state the main results as theorem 1 and 2. In section IV we treat the observer problem. The results in section III are used to investigate the observer properties. Finally in section V the results are applied to the example.

II. EXAMPLE: A MEMS ACCELEROMETER

Consider an accelerometer based on the following design. A test mass is suspended to move freely in one dimension and is affected by an external acceleration. Sensors detect deviations from the origin exceeding a detection threshold and report the sign of the deviation. An input signal is available to accelerate the test mass so as to keep it close to the origin. The aim of the design is to estimate the external acceleration as accurately as possible.

The discrete time dynamics are given by

$$x(k) = \begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} x(k-1) + \begin{pmatrix} \frac{1}{2}h^2 \\ h \end{pmatrix} u(k-1) + v(k-1),$$

where x is the state, u the input signal, v the external acceleration and h the sampling period. The state consists of position x_1 and velocity x_2 . With the external acceleration as a white noise disturbance, sampling yields v to be Gaussian white noise with covariance

$$P_N = \sigma^2 \begin{pmatrix} \frac{1}{3}h^3 & \frac{1}{2}h^2 \\ \frac{1}{2}h^2 & h \end{pmatrix},$$

where σ^2 is the process noise intensity.

The measurements are given by

$$y(k) = \begin{cases} \text{sign}(x_1(k)), & |x_1(k)| \geq 1 \\ 0, & \text{otherwise,} \end{cases}$$

which is the only non-classical assumption used in the model. The output $y(k)$ is not readily described as a linear combination of state and uncorrelated measurement noise, making a straightforward application of Kalman filter theory difficult.

In fact, it is not at all obvious what properties to expect for this observer problem; will the observer error remain bounded, how large will it be, how does it depend on the measurement sequence, how complex observer is necessary, and so on. To answer questions about the observer problem, the Bayesian observer for the system will be analyzed.

Other examples where similar measurement conditions apply are when measurements are coarsely quantized or come in the form of level triggered events.

III. LOG-CONCAVITY

Many results are available on general log-concavity, see for instance [4], [5], [6], and [7]. The book [11] contains much material on convex functions that can easily be transferred to the log-concave case. Here, only the properties that are most relevant in the context of this paper will be stated.

A log-concave function is a function with concave logarithm. Log-concave functions are well suited for applying convexity theory to probability densities; many common densities are log-concave and several useful operations preserve log-concavity. In contrast, no probability density on \mathbb{R}^n is either convex or concave since probability densities have a finite integral while convex and concave functions on \mathbb{R}^n do not.

Definition 1 (Log-concave Function): A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is logarithmic concave or *log-concave*, iff $f(x) \geq 0$, f has convex support and $\ln(f(x))$ is concave on this support.

For some simple examples of log-concave functions see figure 1, and for some counterexamples figure 2. Among common log-concave densities are Gaussian and exponential densities.

Log-concave functions are unimodal, meaning that the superlevel sets $\{x; f(x) \geq a\}, a \in \mathbb{R}$ are convex. Many attractive properties of log-concave functions are analogous to those for concave functions. A useful fact is that multiplication takes the place of addition so that the product of two log-concave functions is log-concave. Another very useful result derived by Prékopa is

Proposition 1 (Prékopa): Let $f(x, y)$ be jointly log-concave in $x \in \mathbb{R}^m, y \in \mathbb{R}^n$. Then the integral

$$g(x) = \int f(x, y) dy$$

is a log-concave function of x .

Proof. See [4] and [5].

This theorem implies for instance that the marginal densities of log-concave densities are log-concave, and that the convolutions of log-concave functions are log-concave. It will be central in the proof of theorems 1 and 2.

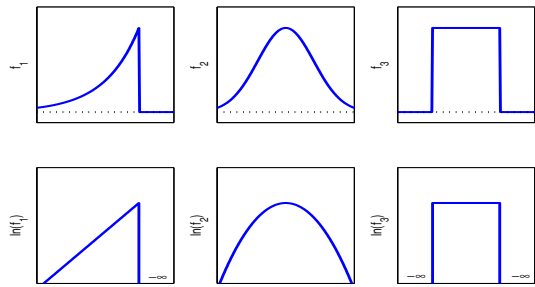


Fig. 1. Some examples of log-concave functions in one variable; the function is plotted above and its logarithm below. The dotted line is $f = 0$, and $\ln(0)$ is taken to be $-\infty$. f_1 : Truncated exponential function, f_2 : Gaussian function, f_3 : rectangular window.

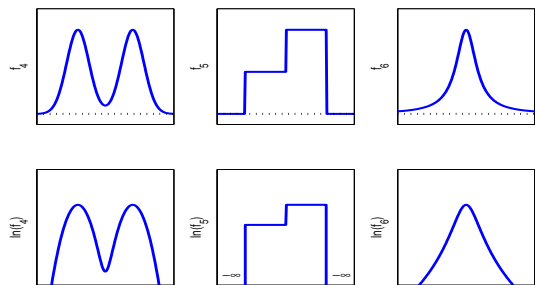


Fig. 2. Some examples of functions in one variable that are *not* log-concave; the function is plotted above and its logarithm below. f_4 : Not unimodal, f_5 : Discontinuous on interior of support, $f_6(x) = \frac{1}{1+x^2}$: sub-exponential decay.

A. Strong log-concavity

Log-concavity is in its nature only a qualitative property. To allow for quantitative statements, the following class of functions is introduced.

Definition 2 (Strongly Log-concave Function): Let $P \in \mathbb{R}^{n \times n}$ be positive definite and define the set

$$\mathcal{LC}(P^{-1}) = \left\{ f; f_0(x) = \frac{f(x)}{e^{-\frac{1}{2}x^T P^{-1}x}} \text{ log-concave} \right\}.$$

The function f is *strongly log-concave* of strength P^{-1} iff $f \in \mathcal{LC}(P^{-1})$.

All strongly log-concave functions are log-concave, bounded and go to zero as $|x| \rightarrow \infty$ at least as fast as a Gaussian function.

Membership in $\mathcal{LC}(P^{-1})$ can be seen as an inequality, in the sense that

$$\begin{aligned} f &\in \mathcal{LC}(P^{-1}), \quad P \leq R \\ \implies f &\in \mathcal{LC}(R^{-1}). \end{aligned}$$

The inclusion $f \in \mathcal{LC}(P^{-1})$ is *tight* iff $\mathcal{LC}(P^{-1})$ is a subset of all $\mathcal{LC}(R^{-1})$ that contain f .

A Gaussian density with covariance P is tightly in $\mathcal{LC}(P^{-1})$, and can be seen as the *corresponding Gaussian* to this class. The definition states that any strongly log-concave function can be written

as the product of a log-concave function and a corresponding Gaussian. Also, the following properties hold:

Theorem 1 (Encapsulation Property): If $f \in \mathcal{LC}(F^{-1})$ and $g \in \mathcal{LC}(G^{-1})$ then

$$\begin{aligned} f(Ax + b) &\in \mathcal{LC}(A^T F^{-1} A) \\ (f * g)(x) &\in \mathcal{LC}((F + G)^{-1}) \\ f(x) \cdot g(x) &\in \mathcal{LC}(F^{-1} + G^{-1}), \end{aligned}$$

where $x, b \in \mathbb{R}^n, A \in \mathbb{R}^{n \times n}$ and $f * g$ is the convolution of f and g .

Proof. See appendix A.

The inclusions are as narrow as the premises allow, being tight when f and g are the corresponding Gaussians.

Theorem 2 (Covariance Bound): If $f \in \mathcal{LC}(P^{-1})$ is a probability density then

$$V = \int (x - m_x)(x - m_x)^T f(x) dx \leq P,$$

where $m_x = \int x f(x) dx$. The bound is tight for the corresponding Gaussian.

Proof. See appendix B.

The matrix expressions for strength of log-concavity correspond exactly to the way that the operations propagate inverse covariances for Gaussian functions. By the latter theorem, the inverse strength of log-concavity is an upper bound on the covariance.

The theorems form a chain of inequalities that can be used to propagate upper bounds on covariance under the operations of affine transformation, convolution and multiplication. For more properties of strongly log-concave functions, see [10].

IV. LOG-CONCAVE OBSERVERS

The observer problem that will be considered is for processes with linear dynamics and log-concave noise and measurements, as defined below.

The dynamics are given by

$$x(k) = Ax(k-1) + Bu(k-1) + v(k-1),$$

where x is the state, u the input and v the process noise. The noise has log-concave probability density f_N . The matrices A and B , as well as f_N may be time-varying.

The measurements are described by the stochastic variables $Y(k)$,

$$f_{Y(k)|X(k)}(y|x(k)) = f_M(y, x(k)),$$

where the measurement function f_M is log-concave in x for each y and may be time-varying.

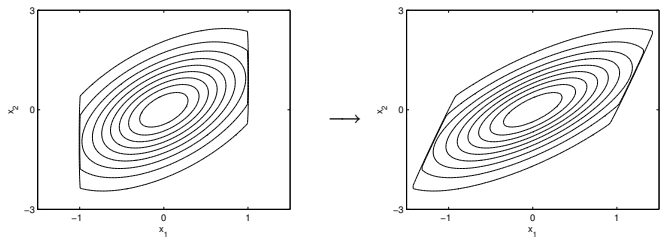


Fig. 3. Illustration of the dynamics update for the MEMS accelerometer observer. The transformation amounts to a shear in this case.

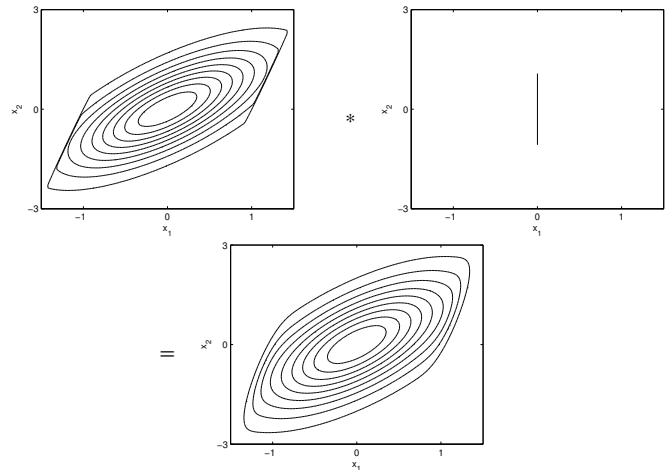


Fig. 4. Illustration of the process noise update for the MEMS accelerometer observer. The Gaussian noise enters almost exclusively in the x_2 direction.

A. The Bayesian observer

As a basis for the analysis, the online Bayesian observer for estimation of $x(k)$ from the history of y and u will be considered. The state of the observer at any time is fully described by the function

$$f_k(x) = f_{X_k|y_{1:k}, f_{X_0}}(x),$$

where $y_{1:k}$ is the measurement history and f_{X_0} is the assumed initial density.

The observer update from f_{k-1} to f_k is best described in three steps taking into account dynamics, process noise, and measurements:

$$f_k^d(x) \propto f_{k-1}(A^{-1}x - A^{-1}Bu(k-1)), \quad (1)$$

$$f_k^{\text{dn}}(x) = (f_N * f_k^d)(x), \quad (2)$$

$$f_k(x) \propto f_M(y(k), x) \cdot f_k^{\text{dn}}(x), \quad (3)$$

where \propto denotes proportionality and A is assumed to be invertible. For the derivation, see [10]. The dynamics update corresponds to an affine transformation, the noise update to a convolution with f_N , and the measurement update to a multiplication with $f_M(y(k), \cdot)$. For an illustration, see figures 3, 4 and 5. If f_{X_0}, f_N and $f_M(y, \cdot)$ are Gaussian, the observer updates (1)-(3) reduce to a Kalman filter.

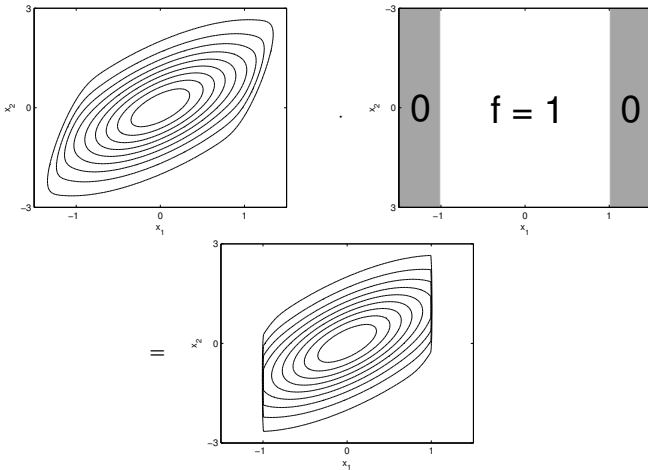


Fig. 5. Illustration of the measurement update for the MEMS accelerometer observer. The measurement is $y = 0$.

B. Properties

Since log-concavity is preserved under affine parameter transformation, convolution, and multiplication, all f_k are log-concave if f_{X_0} is log-concave.

Theorem 1 can be used to propagate upper bounds on observer covariance. This approach can be used to assess the merits of a particular sensor setup, or together with some information about the localization of f_k to give state estimates with error bounds. The computations of covariance propagation have the structure of a Kalman filter applied to corresponding Gaussians.

V. AN APPLICATION

The MEMS accelerometer will now be used to illustrate how the theory can be applied in the analysis of a concrete observer problem.

A. Analytical covariance bounds

The accelerometer has linear dynamics and log-concave noise and measurements. The process noise density f_N is Gaussian with covariance P_N , so that $f_N \in \mathcal{LC}(P_N^{-1})$. The measurement function $f_M(y, x)$ is log-concave in x for all y , see figure 6.

Applying theorem 1 directly leads in this case to a highly conservative covariance bound, achieved when completely ignoring the measurements. The bound grows cubically with time. Grid based finite difference simulations of the Bayesian observer do however indicate that the covariance is bounded, and if the output changes frequently, small.

The reason why the bound is so conservative is that f_M is not strongly log-concave for any y ; strength of log-concavity being the only measure of information that the theorem considers. In lack of stronger proven results, a slight approximation will allow to account for the major source of state information.

The most important source of state information under normal conditions is the events when y goes

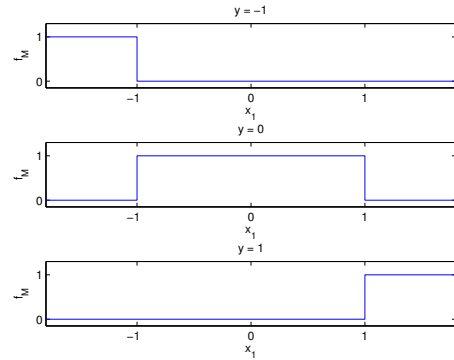


Fig. 6. The measurement function $f_M(y, x)$ for the MEMS accelerometer describing the relative probability of state x when $y = -1, 0, -1$. The function is independent of x_2 .

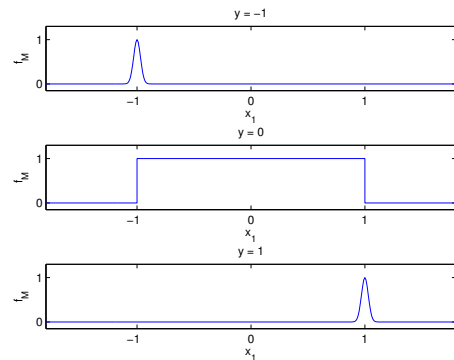


Fig. 7. The modified measurement function $\hat{f}_M(y, x)$ for the MEMS accelerometer when $y = -1, 0, -1$. For $y = \pm 1$, the function has been changed to a narrow Gaussian centered on the detection threshold.

from being 0 to ± 1 , at which time x_1 is known to be almost exactly equal to y . This can be modeled as a Gaussian measurement of x_1 with expectation y and variance σ_M^2 .

The variance σ_M^2 will depend on the process noise and uncertainty in velocity, but will be small when h is small. The modified measurement function \hat{f}_M can be seen in figure 7. For events, $\hat{f}_M(\pm 1, \cdot) \in \mathcal{LC}(Q_M)$ where

$$Q_M = \begin{pmatrix} \sigma_M^{-2} & 0 \\ 0 & 0 \end{pmatrix},$$

and otherwise $\hat{f}_M(0, \cdot) \in \mathcal{LC}(0)$.

Under this approximation, the variance of the optimal estimate \hat{x}_2 of x_2 right after an event can be shown to satisfy

$$V(\hat{x}_2) \leq \frac{1}{3}\sigma^2 t + 2\sigma_M^2 t^{-2},$$

where t is the time since the last event. For the derivation, see [10].

The bound illustrates that the accuracy of the accelerometer depends strongly on the rate of events.

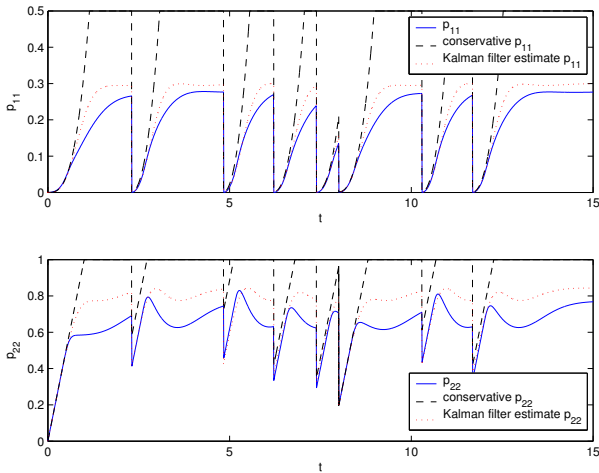


Fig. 8. Observer covariances during a simulation for the MEMS accelerometer: grid filter, approximate upper bound and tuned Kalman filter.

If the objective of control is good measurements, the controller should keep the rate above a certain level, for instance sending the test mass bouncing in a ping pong fashion between the detection boundaries.

B. Kalman filter approximation

A Kalman filter was tuned to give a reasonable approximation of the Bayesian observer. The crucial issue was to assign the covariance of the measurement $y = 0$. While a single measurement $y = 0$ predicts x_1 to have expectation zero with variance $\sigma^2 = \frac{1}{3}$, there is much less additional information in the measurement $y = 0$ at the next time step.

In this case it is reasonable to design the Kalman filter by choosing the stationary variance p_{11}^{stat} of x_1 when $y = 0$. The variance would typically be $p_{11}^{\text{stat}} = \frac{1}{\sqrt{3}}$ (rectangular distribution) or a little less. From solving the Riccati equation, it is found that the measurement variance $\sigma_o^2 h^{-1}$ for $y = 0$ must be chosen according to

$$\sigma_o = 2^{-1/3} (p_{11}^{\text{stat}})^{2/3} \sigma^{-1/3},$$

where σ^2 is the process noise intensity.

C. Simulation

Figure 8 shows a comparison of actual and predicted variances for a simulation of the Bayesian observer. The variance σ_M^2 was chosen so that the approximate upper bound would always be conservative. The upper bound is quite tight some time after each event, but then diverges. The variance of the tuned Kalman filter appears to be an only mildly conservative approximation of the actual variance. As long as the rate of events is reasonably high, the approximate upper bound is very tight.

A simple control law was devised to control the rate of events, and simulations were run for different rates to compare observer performance for

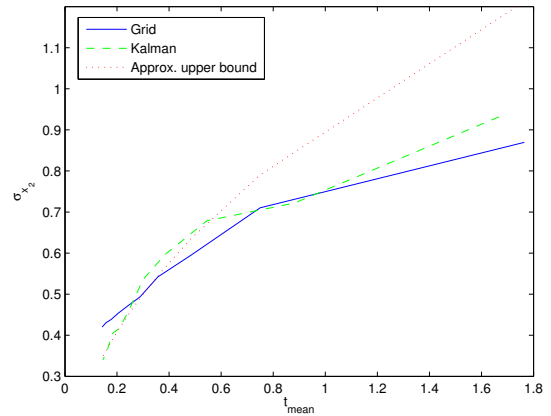


Fig. 9. RMS x_2 estimation error as a function of mean time between events: grid filter estimation error, tuned Kalman filter estimation error, and approximate upper bound. For too high event rates, the grid filter suffers from discretization problems.

the grid filter and the tuned Kalman filter. Figure 9 shows the observer error as a function of mean time between events t_{mean} . The grid filter is slightly better than the tuned Kalman filter and considerably better than the approximate covariance bound down to values of t_{mean} around 0.4.

For lower t_{mean} it seems that the grid filter scheme encounters discretization issues. At the same time, the tuned Kalman filter comes very close to the approximate upper bound which appears to be very tight in this region, indicating that the observer problem is very similar to the Kalman filter case for high rates. This similarity is not surprising since when the covariance is small, the bulk of probability mass is concentrated in a small region which is only seldom affected by the non Gaussian measurements.

Thus it is seen that the upper bound derived from the theory is quite tight when the rate of events is high and that if the inherent correlation in the non Gaussian measurements is considered, a Kalman filter can be applied as a close to optimal observer.

Example 1 (Quantized measurements): In the previous example it was necessary to rely on approximations because the measurement functions were not strongly log-concave. If the measurement function can be chosen freely, much stronger results are possible.

Consider the general problem of estimating a scalar variable from a series of independent identically distributed quantized measurements. The objective is to find a conditional measurement distribution, or measurement function, that is in some sense optimal. Using strength of log-concavity as an optimality criterion one can formulate the following problem:

Let the independent measurements y be dis-

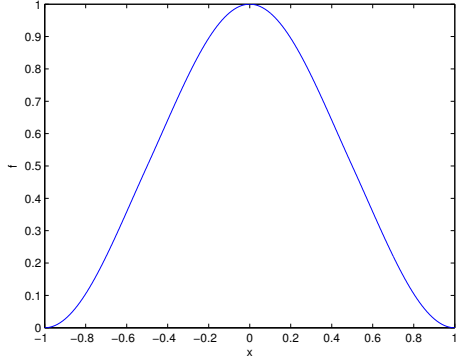


Fig. 10. The measurement function in example 1. The function is Gaussian when $|x| \leq \frac{1}{2}$ and zero when $|x| \geq 1$.

tributed according to

$$f_{Y|x}(y|x) = f(x - y), \quad y \in \mathbb{Z},$$

where x is the variable to be estimated. Find a function $f \in \mathcal{LC}(p^{-1})$, where $p > 0$ is as small as possible, such that

$$\begin{aligned} f(x) &\geq 0, \\ f(-x) &= f(x), \\ \sum_{k=-\infty}^{\infty} f(x - k) &= 1. \end{aligned}$$

The solution is given by the function

$$f(x) = \begin{cases} 2^{-4|x|^2}, & |x| \leq \frac{1}{2}, \\ 1 - 2^{-4(1-|x|)^2}, & \frac{1}{2} < |x| \leq 1, \\ 0, & \text{otherwise,} \end{cases}$$

satisfying $f(x) \in \mathcal{LC}(8 \ln(2))$. The function is plotted in figure 10, and in log-scale in figure 11. A series of n measurements with f as measurement function is guaranteed to yield a probability density in $\mathcal{LC}(n \cdot 8 \ln(2))$ and therefore a variance satisfying $\sigma^2 \leq \frac{1}{n \cdot 8 \ln(2)}$.

VI. CONCLUSION

Log-concavity is a powerful tool when dealing with probability densities. The generalization to allow log-concave densities in the observer widens the range of application considerably compared to the Kalman filter. Although no closed form solution exists in the general case, the observer problem is still very accessible to mathematical treatment.

Regarding observability and observer performance, strongly log-concave functions together with theorems 1 and 2 can be applied to derive simple bounds on achievable observer covariance.

An in-depth treatment of the log-concave case gives a greater understanding of the performance of

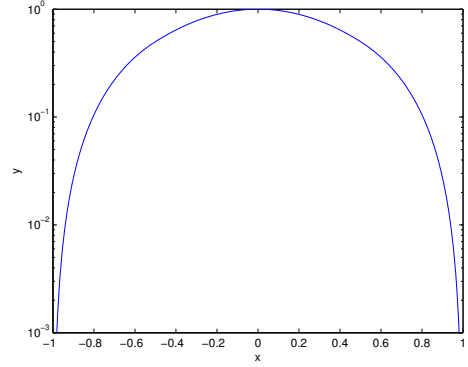


Fig. 11. The measurement function in example 1 in log-scale. The logarithm is clearly concave, being quadratic when $|x| \leq \frac{1}{2}$.

an Extended Kalman filter. In design of instrumentation, striving for log-concave measurement functions can facilitate the observer problem.

VII. ACKNOWLEDGEMENT

The authors would like to thank Anders Rantzer, Maria Karlsson, Peter Alriksson and Ather Gattami for helpful comments.

REFERENCES

- [1] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME - Journal of Basic Engineering*, Vol. 82: pp. 35-45, 1960.
- [2] R.E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," *Transactions of the ASME - Journal of Basic Engineering*, Vol. 83: pp. 95-107, 1961.
- [3] A. Gelb, editor, *Applied Optimal Estimation*. MIT Press, 1974.
- [4] A. Prékopa, "Logarithmic concave measures with application to stochastic programming," *Acta Scientiarum Mathematicarum*, 32, 301-315, 1971.
- [5] A. Prékopa, "On logarithmic concave measures and functions," *Acta Scientiarum Mathematicarum*, 33, 335-343, 1973.
- [6] M. Bagnoli and T. Bergstrom, "Log-concave probability and its applications" (January 1, 2004). Department of Economics, UCSB. Ted Bergstrom. Paper 1989D. Available at <http://repositories.cdlib.org/ucsbecon/bergstrom/1989D/>
- [7] M.Y. An, "Log-concave probability distributions: theory and statistical testing." Technical report, Economics Department, Duke University, Durham, N.C. 27708-0097, 1995.
- [8] K. J. Åström and B. Bernhardsson, "Comparison of Riemann and Lebesgue Sampling for First-order Stochastic Systems," in *IEEE Conference on Decision and Control*, 2002.
- [9] T. Schön, F. Gustafsson and A. Hansson, "A note on state estimation as a convex optimization problem," in *Proceedings of ICASSP, Hongkong, PRC*, 2003.
- [10] Toivo Henningsson, "Logarithmic Concave Observers." Master's thesis ISRN LUTFD2/TFRT-5747-SE, May 2005. Department of Automatic Control, Faculty of Engineering, Lund University, Lund, Sweden.
- [11] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge: Cambridge University Press, 2004. Available at <http://www.ee.ucla.edu/~vandenbe/cvxbook.html>.

APPENDIX

A. PROOF OF THEOREM 1

The proofs are based on the fact that a function f is in $\mathcal{LC}(F^{-1})$ iff it can be factored as

$$f(x) = e^{-\frac{1}{2}x^T F^{-1}x} f_0(x), \quad (4)$$

where $f_0(x)$ is log-concave. This follows from the definition.

The proofs for affine transformation and multiplication are straightforward, while the proof for convolution is a little more involved.

A. Affine transformation

Let $f \in \mathcal{LC}(F^{-1})$, $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ and $y = Ax + b$. Then

$$\begin{aligned} g(x) &= f(Ax + b) \\ &= e^{-\frac{1}{2}(Ax+b)^T F^{-1}(Ax+b)} \cdot f_0(y) \\ &= e^{-\frac{1}{2}(x^T A^T F^{-1} Ax + 2b^T F^{-1} Ax + b^T F^{-1} b)} \cdot f_0(y) \\ &= e^{-\frac{1}{2}x^T (A^T F^{-1} A)x} \cdot \underbrace{\left(e^{-\frac{1}{2}b^T F^{-1} b} e^{-(A^T F^{-1} b)^T x} f_0(y) \right)}_{g_0(x)}. \end{aligned}$$

We see that g_0 is the product of a constant, an exponential function and a log-concave function, since log-concavity is preserved under affine parameter transformation. Then g_0 is log-concave because each of the factors is log-concave. Thus $g \in \mathcal{LC}(A^T F^{-1} A)$.

B. Convolution

For the proof we need the following matrix identity. Let A, B and C be positive definite matrices such that $C^{-1} = A^{-1} + B^{-1}$, or $C = A(A + B)^{-1}B$. Let x, y and $z = y - (A + B)^{-1}Bx$ be vectors. Then

$$\begin{aligned} z^T (A + B)z \\ = y^T (A + B)y - 2x^T B y + x^T B (A + B)^{-1} B x \end{aligned}$$

and

$$\begin{aligned} x^T C x + z^T (A + B)z \\ = x^T A (A + B)^{-1} B x + z^T (A + B)z \\ = x^T B x + y^T (A + B)y - 2x^T B y \\ = y^T A y + (x - y)^T B (x - y), \end{aligned}$$

that is,

$$y^T A y + (x - y)^T B (x - y) = x^T C x + z^T (A + B)z, \quad (5)$$

which can be seen as completion of squares in x .

Let $f \in \mathcal{LC}(F^{-1})$ and $g \in \mathcal{LC}(G^{-1})$. Then

$$\begin{aligned} h(x) &= (f * g)(x) \\ &= \int f(y)g(x - y) dy \\ &= \int e^{-\frac{1}{2}y^T F^{-1}y} e^{-\frac{1}{2}(x-y)^T G^{-1}(x-y)} \cdot f_0(y)g_0(x - y) dy \\ &= \int e^{-\frac{1}{2}\left(y^T F^{-1}y + (x-y)^T G^{-1}(x-y)\right)} \cdot f_0(y)g_0(x - y) dy. \end{aligned}$$

Applying (5) with $A = F^{-1}$, $B = G^{-1}$ and $C = H^{-1}$ yields $H^{-1} = (F + G)^{-1}$ and

$$\begin{aligned} h(x) &= \int e^{-\frac{1}{2}\left(x^T H^{-1}x + z^T (F^{-1} + G^{-1})z\right)} \cdot f_0(y)g_0(x - y) dy \\ &= e^{-\frac{1}{2}x^T H^{-1}x} \underbrace{\int e^{-\frac{1}{2}z^T (F^{-1} + G^{-1})z} \cdot f_0(y)g_0(x - y) dy}_{h_0(x)}. \end{aligned}$$

The integrand is log-concave since it is the product of a Gaussian function and two log-concave functions, and thus h_0 is log-concave according to theorem 1. This proves that $h \in \mathcal{LC}(H^{-1}) = \mathcal{LC}((F + G)^{-1})$.

C. Multiplication

Let $f \in \mathcal{LC}(F^{-1})$ and $g \in \mathcal{LC}(G^{-1})$. Then

$$\begin{aligned} h(x) &= f(x)g(x) \\ &= e^{-\frac{1}{2}x^T F^{-1}x} e^{-\frac{1}{2}x^T G^{-1}x} \cdot f_0(x)g_0(x) \\ &= e^{-\frac{1}{2}x^T (F^{-1} + G^{-1})x} \cdot h_0(x), \end{aligned}$$

where f_0 and g_0 are log-concave and $h_0(x) = f_0(x)g_0(x)$. Thus h_0 is log-concave and so $h \in \mathcal{LC}(F^{-1} + G^{-1})$.

APPENDIX

B. PROOF OF THEOREM 2

The factorization (4) will be central also in this proof. Consider first the theorem in one dimension. Let $f \in \mathcal{LC}(p^{-1})$, $p > 0$ be a nonincreasing probability density defined for $x \geq 0$. Then f can be factored as

$$f(x) = e^{-\frac{1}{2}p^{-1}x^2} f_0(x),$$

where $f_0(x)$, $x \geq 0$ is log-concave. The right derivative $f'(0)$ exists since any convex function is almost everywhere differentiable which transfers trivially to log-concave functions. Furthermore $f'_0(0) = f'(0) \leq 0$, and since f_0 is log-concave it is nonincreasing for all $x \geq 0$.

Let $C > 0$ be defined such that

$$\int_0^\infty C e^{-\frac{1}{2}p^{-1}x^2} dx = \int_0^\infty \underbrace{e^{-\frac{1}{2}p^{-1}x^2} f_0(x)}_{f(x)} dx = 1.$$

Then, since $f_0(x)$ is nonincreasing, there must exist some $x_0 > 0$ such that

$$\begin{aligned} f_0(x) &\geq C, \quad x < x_0 \\ f_0(x) &\leq C, \quad x > x_0. \end{aligned}$$

The second moment of f is

$$\begin{aligned}
& \int_0^\infty x^2 f(x) dx \\
&= \int_0^\infty x^2 C e^{-\frac{1}{2}P^{-1}x^2} dx + \int_0^\infty x^2 (f(x) - C e^{-\frac{1}{2}P^{-1}x^2}) dx \\
&= p + \int_0^\infty x^2 e^{-\frac{1}{2}P^{-1}x^2} (f_0(x) - C) dx \\
&= p + \int_0^\infty e^{-\frac{1}{2}P^{-1}x^2} (x_0^2 + (x^2 - x_0^2)) (f_0(x) - C) dx \\
&\leq p + x_0^2 \int_0^\infty e^{-\frac{1}{2}P^{-1}x^2} (f_0(x) - C) dx \\
&= p,
\end{aligned}$$

where we have used that $(x^2 - x_0^2)(f_0(x) - C) \leq 0$. Thus the second moment of f around $x = 0$ is $\leq p$.

Now assume that $f(x) \in \mathcal{LC}(P^{-1})$ is an arbitrary strongly log-concave function in one dimension that assumes its maximum value at $x = M_x$. All strongly log-concave functions are bounded and go to zero as $|x| \rightarrow \infty$, so if f does not assume its maximum it can be made to do so by changing the value in one point, which does not affect integrals of f and preserves strong log-concavity. Then $g(x) = f(x - M_x)$ can be written as a convex combination of two probability densities in $\mathcal{LC}(P^{-1})$ such that one has support on $x < 0$ and is nondecreasing and one has support on $x \geq 0$ and is nonincreasing. The second moment of g around 0 is a convex combination of the moments of the two densities, and so

$$\int (x - M_x)^2 f(x) dx \leq p.$$

Since the covariance of the density f is the minimum of the second moment around any point,

$$\int (x - m_x)^2 f(x) dx = \min_y \int (x - y)^2 f(x) dx \leq p,$$

where m_x is expectation of the density. This proves the theorem in one dimension.

For the proof in R^n we shall need another matrix inequality. In the Cauchy-Schwartz inequality $(u^T v)^2 \leq (u^T u)(v^T v)$, let $u = P^{-\frac{1}{2}}x$ and $v = P^{\frac{1}{2}}e_z$, where $P > 0$, $\|e_z\| = 1$. This yields

$$\begin{aligned}
& (x^T e_z)^2 \leq (x^T P^{-1}x)(e_z^T P e_z) \\
& \implies x^T e_z (e_z^T P e_z)^{-1} e_z^T x \leq x^T P^{-1}x \\
& \implies Q_r = e_z (e_z^T P e_z)^{-1} e_z^T \leq P^{-1}. \tag{6}
\end{aligned}$$

Now consider a density $f \in \mathcal{LC}(P^{-1})$, $P > 0$. Without loss of generality, assume the expectation to be zero. The covariance is then

$$V = \int x x^T f(x) dx,$$

and for a given unit vector e_z ,

$$e_z^T V e_z = \int (e_z^T x)^2 f(x) dx = \int_{t \in \mathbb{R}} t^2 \underbrace{\int_{y \perp e_z} f(te_z + y) dy}_{g(t)} dt,$$

where $x = te_z + y$ and $g(t)$ is the marginal density of f in the e_z direction, having zero expectation. We see that

$$\begin{aligned}
g(t) &= \int_{y \perp e_z} e^{-\frac{1}{2}x^T P^{-1}x} f_0(x) dy \\
&= \int_{y \perp e_z} e^{-\frac{1}{2}x^T Q_r x} e^{-\frac{1}{2}x^T (P^{-1} - Q_r)x} f_0(x) dy \\
&= e^{-\frac{1}{2}(e_z^T P e_z)^{-1}t^2} \underbrace{\int_{y \perp e_z} e^{-\frac{1}{2}x^T (P^{-1} - Q_r)x} f_0(x) dy}_{g_0(t)},
\end{aligned}$$

since $y^T e_z = 0$ so that $x^T Q_r x = te_z^T Q_r e_z t = (e_z^T P e_z)^{-1}t^2$. From (6) $P^{-1} - Q_r \geq 0$ so that g_0 is log-concave. Thus $g \in \mathcal{LC}((e_z^T P e_z)^{-1})$ so that

$$e_z^T V e_z \leq e_z^T P e_z \implies V \leq P,$$

which proves the theorem.

The bound is tight for the corresponding Gaussian by definition.