Margin requirements for first page
Paper size this page A4

94 pt
1.306 in
33.2 mm

Constrained Posterior Cramér-Rao Bound for Discrete-Time Systems [1]

Leif Erik Andersson *,
Lars Imsland *, **Edmund Førland Brekke** *, **Francesco Scibilia** **

* *Department of Engineering Cybernetics, Norwegian University of Science and Technology, 7491 Trondheim, Norway. (e-mail: {lars.imsland, leif.e.andersson}@itk.ntnu.no)*
** *Statoil ASA, Statoil Research Center, Arkitekt Ebbells veg 10, 7053 Ranheim, Norway. (e-mail: fsci@statoil.com)*

Abstract: This article presents a Cramér-Rao lower bound for the discrete-time filtering problem under linear state constraints. A simple recursive algorithm is presented that extends the computation of the Cramér-Rao lower bound found in previous literature by one additional step in which the full-rank Fisher Information matrix is projected onto the tangent hyperplane of the constraint set. This makes it possible to compute the constrained Cramér-Rao lower bound for the discrete-time filtering problem without reparametrization of the original problem to remove redundancies in the state vector, which improves insights into the problem. It is shown that in case of a positive-definite Fisher Information Matrix the presented constrained Cramér-Rao bound is lower than the unconstrained Cramér-Rao bound. The bound is evaluated on an example.

1. INTRODUCTION

40 pt
0.556 in
14.1 mm

Discrete-time state estimation arise in adaptive control, identification as well as in model-based control, where it is usually prior to the control prediction step. In general, it is challenging to build an optimal estimator for such systems. Hence, it is necessary to turn to one of the many suboptimal filter techniques (Galdos, 1980).

A common strategy to design such an estimator is to use the Bayesian approach. Closed form solutions exist only for a few cases, and often it is necessary to approximate the Bayesian solution numerically. However, filters based on such approximations lead to estimates that deviate from the ideal solution (Šimandl et al., 2001). Lower bounds on the mean-square error of an estimate can give an indication of performance limitations. Consequently, it can be used to determine whether imposed performance requirements are realistic or not (Galdos, 1980; Šimandl et al., 2001; Tichavsky et al., 1998). The Cramér-Rao bound (CRB) given as the inverse of the Fisher information matrix (FIM) presents such a lower bound for dynamic models. However, in time-varying systems the estimated parameter vector (the estimated state vector) has to be considered random since it corresponds to an underlying nonlinear, randomly driven model (Tichavsky et al., 1998). In Van Trees (1968) the CRB was extended for random parameter estimation. The CRB was successfully applied in state estimation for discrete-time non-linear stochastic dynamic systems in Galdos (1980) and Bobrovsky and Zakai (1975). The basic principle for both bounds is to construct a suitable Gaussian system for which the mean-square estimation error is a lower bound to that of the original system. Another approach to compute the CRB for filtering problems in discrete-time non-linear systems was proposed by Tichavsky et al. (1998). They assume the state history as random parameter and obtain the CRB for the state as lower right block of the CRB for the complete state history (Šimandl et al., 2001). They referred to the obtained bound as posterior CRB (PCRB).

40 pt
0.556 in
14.1 mm

In some applications, prior knowledge in form of linear equality constraints is available. This information should result in improved estimates and a lower CRB (Marzetta, 1993). One way to find the CRB under constraints is to reparametrize the original problem and remove redundancies in the parameter vector. However, this approach may be difficult to implement, and may hinder insights into the original unconstrained problem (Stoica and Ng, 1998). Gorman and Hero (1990) proposed a convenient way to compute the constrained CRB for static problems. The bound equals the bound of the original unconstrained problem minus a correction matrix. The same result with a different proof was presented by Marzetta (1993). Another constrained CRB, which also holds for singular Fisher Information matrices, was presented by Stoica and Ng (1998). This theory was extended for complex parameters by Jagannatham and Rao (2004) and biased estimators by Ben-Haim and Eldar (2009).

This article connects the theory about constrained CRB with the PCRB for discrete-time systems presented by Tichavsky et al. (1998).

The article proceeds as follows: An overview over the PCRB is given in Section 2. In Section 3 the constrained CRB is introduced. How to compute the constrained CRB recursively is established in Section 4. In Section 5 it is proven that the constrained CRB is smaller or equal

[1] This work was supported by Statoil ASA, and in part by Centre for Autonomous Marine Operations and Systems (CoE AMOS, RCN project no. 223254)

68 pt
0.944 in
24 mm

Margin requirements for the other pages
Paper size this page A4

68 pt
0.944 in
24 mm

than the unconstrained CRB, while in Section 6 it is demonstrated how to compute the constrained CRB for a linear Gaussian case. A numerical example is discussed in Section 7 followed by a conclusion.

2. THE CRB FOR THE NON-LINEAR FILTERING PROBLEM

This section follows mainly the posterior CRB presented in Tichavsky et al. (1998), which was also summarized in Šimandl et al. (2001). Consider the discrete-time non-linear filtering problem

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{w}_k) \tag{1a}$$
$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k), \tag{1b}$$

where k is the time index, and $\mathbf{x}_k \in \mathbb{R}^n$ and $\mathbf{z}_k \in \mathbb{R}^y$ represent the state and measurement vectors, respectively. The vectors $\mathbf{w}_k \in \mathbb{R}^n$ and $\mathbf{v}_k \in \mathbb{R}^y$ are mutually independent white processes and $\mathbf{f}_k$ and $\mathbf{h}_k$ are non-linear functions. They may depend on the time k. The white noise processes are described by known probability density functions (pdf) $p(\mathbf{w}_k)$ and $p(\mathbf{v}_k)$. Furthermore, it is assumed that the initial state $\mathbf{x}_0$ has a known pdf $p(\mathbf{x}_0)$.
Let the complete state and measurement histories up to the time instant k be denoted as $\mathbf{X}_k = [\mathbf{x}_0^T, \mathbf{x}_1^T, \ldots, \mathbf{x}_k^T]^T$ and $\mathbf{Z}_k = [\mathbf{z}_0^T, \mathbf{z}_1^T, \ldots, \mathbf{z}_k^T]^T$, respectively. The joint pdf of state and measurement histories $p(\mathbf{X}_k, \mathbf{Z}_k)$ may be written as $p(\mathbf{X}_k, \mathbf{Z}_k) = p(\mathbf{Z}_k|\mathbf{X}_k)p(\mathbf{X}_k)$. Respecting that, the stochastic system (1) is a Markov process, the logarithm of this pdf can be expressed as

$$\ln p(\mathbf{X}_k, \mathbf{Z}_k) = \ln p(\mathbf{x}_0) + \sum_{i=0}^{k} \ln p(\mathbf{z}_i|\mathbf{x}_i) + \sum_{i=1}^{k} \ln p(\mathbf{x}_i|\mathbf{x}_{i-1}). \tag{2}$$

If the expectation and derivatives exist, the FIM for this system can be computed as

$$\mathbf{J}_{k|k}(\mathbf{X}_k) = -E\left(\nabla_{\mathbf{X}_k}[\nabla_{\mathbf{X}_k} \ln p(\mathbf{X}_k, \mathbf{Z}_k)]^T\right), \tag{3}$$

where we know that the mean-square error matrix (MSEM) is bounded by the inverse of the FIM

$$\mathbf{\Sigma}_{k|k} = E\left\{(\mathbf{X}_k - \hat{\mathbf{X}}_k)(\mathbf{X}_k - \hat{\mathbf{X}}_k)^T\right\} \geq \mathbf{J}_{k|k}^{-1}. \tag{4}$$

Following the notation used in Šimandl et al. (2001) and in order to simplify the derivation of the filtering estimate let us introduce the following $n \times n$ matrices

$$\mathbf{K}_{k+1}^{k} = E\{-\nabla_{\mathbf{x}_k}^{\mathbf{x}_k} \ln p(\mathbf{x}_{k+1}|\mathbf{x}_k)^T]\}, \tag{5a}$$
$$\mathbf{K}_{k+1}^{k,k+1} = E\{-\nabla_{\mathbf{x}_k}^{\mathbf{x}_{k+1}} \ln p(\mathbf{x}_{k+1}|\mathbf{x}_k)^T]\} = [\mathbf{K}_{k+1}^{k+1,k}]^T, \tag{5b}$$
$$\mathbf{K}_{k+1}^{k+1} = E\{-\nabla_{\mathbf{x}_{k+1}}^{\mathbf{x}_{k+1}} \ln p(\mathbf{x}_{k+1}|\mathbf{x}_k)^T]\}, \tag{5c}$$
$$\mathbf{L}_{k}^{k} = E\{-\nabla_{\mathbf{x}_k}^{\mathbf{x}_k} \ln p(\mathbf{y}_k|\mathbf{x}_k)]^T\}, \tag{5d}$$

where we will define

$$\mathbf{D}_k = \mathbf{L}_k^k + \mathbf{K}_k^k + \mathbf{K}_{i+1}^k. \tag{6}$$

The lower index in (5) is the time instant of the state described by the transition pdf, while the upper index represents the states for which the derivatives of the transition pdf are performed.
Using (2) and the introduced notations (5) the FIM (4) decomposes into four blocks

$$\mathbf{J}_{k|k}(\mathbf{X}_k) = \left[\begin{array}{cccc|c} \mathbf{D}_0 & \mathbf{K}_1^{0,1} & & & \\ \mathbf{K}_1^{1,0} & \ddots & \ddots & & \\ & \ddots & \mathbf{D}_{k-1} & & \mathbf{K}_k^{k-1,k} \\ \hline & & & \mathbf{K}_k^{k,k-1} & \mathbf{L}_k^k + \mathbf{K}_k^k \end{array}\right] = \left[\begin{array}{c|c} \mathbf{J}_{k|k}^{1,1} & \mathbf{J}_{k|k}^{1,2} \\ \hline \mathbf{J}_{k|k}^{2,1} & \mathbf{J}_{k|k}^{2,2} \end{array}\right], \tag{7}$$

where the zero elements have been left empty and for $k = 0$ it holds $\mathbf{J}_{0|0}(\mathbf{x}_0) = \mathbf{L}_0^0 + \mathbf{K}_0^0$. The blocks of the FIM represent the state history decomposed as $\mathbf{X}_k = [\mathbf{X}_{k-1}^T, \mathbf{x}_k^T]^T$. Following this notation, we can see that the time update can be expressed as (Šimandl et al., 2001)

$$\mathbf{J}_{k+1|k}(\mathbf{X}_{k+1}) = \left[\begin{array}{cc|c} \mathbf{J}_{k|k}^{1,1} & \mathbf{J}_{k|k}^{1,2} & 0 \\ \mathbf{J}_{k|k}^{2,1} & \mathbf{J}_{k|k}^{2,2} + \mathbf{K}_{k+1}^{k} & \mathbf{K}_{k+1}^{k,k+1} \\ \hline 0 & \mathbf{K}_{k+1}^{k+1,k} & \mathbf{K}_{k+1}^{k+1} \end{array}\right] \tag{8}$$

and the measurement update as

$$\mathbf{J}_{k|k}(\mathbf{X}_k) = \left[\begin{array}{c|c} \mathbf{J}_{k|k-1}^{1,1} & \mathbf{J}_{k|k-1}^{1,2} \\ \hline \mathbf{J}_{k|k-1}^{2,1} & \mathbf{J}_{k|k-1}^{2,2} + \mathbf{L}_k^k \end{array}\right]. \tag{9}$$

The dimension of the FIM (8) and (9) increase at each iteration. Furthermore, it can be easily seen that $\mathbf{J}_{k|k-1}(X_k)$ and $\mathbf{J}_{k|k}(X_k)$ are equal except for the lower-right corner block, which is $\mathbf{K}_k^k$ compared to $\mathbf{K}_k^k + \mathbf{L}_k^k$.
Applying (4) to (9) a formulation for the inequality for the MSEM of a filtering estimate at time k can be obtained

$$E\left\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T\right\} \geq \mathbf{C}_{k|k} = [\mathbf{J}_{k|k}^{-1}(\mathbf{X}_k)]_{22}, \tag{10}$$

where $\mathbf{C}_{k|k}$ is the PCRB of an estimate $\hat{\mathbf{x}}_{k|k}$. Using the matrix inversion lemma (A.1)-(A.2) and $\mathbf{J}_{k|k-1}^{2,2} = \mathbf{K}_k^k$ we obtain

$$\mathbf{C}_{k|k}^{-1} = \mathbf{L}_k^k + \mathbf{K}_k^k - \mathbf{J}_{k|k-1}^{2,1}[\mathbf{J}_{k|k-1}^{1,1}]^{-1}\mathbf{J}_{k|k-1}^{1,2} \tag{11}$$

for the measurement update. For the time-update with the same matrix inversion lemma the following can be obtained

$$\begin{aligned} \mathbf{C}_{k+1|k}^{-1} &= \mathbf{K}_{k+1}^{k+1} - \begin{pmatrix} 0 & \mathbf{K}_{k+1}^{k+1,k} \end{pmatrix} \begin{pmatrix} \mathbf{J}_{k|k}^{1,1} & \mathbf{J}_{k|k}^{1,2} \\ \mathbf{J}_{k|k}^{2,1} & \mathbf{J}_{k|k}^{2,2} + \mathbf{K}_{k+1}^{k} \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \mathbf{K}_{k+1}^{k,k+1} \end{pmatrix} \\ &= \mathbf{K}_{k+1}^{k+1} - \\ &\quad \mathbf{K}_{k+1}^{k+1,k} \left(\mathbf{J}_{k|k}^{2,2} + \mathbf{K}_{k+1}^{k} - \mathbf{J}_{k|k}^{1,2}[\mathbf{J}_{k|k}^{1,1}]^{-1}\mathbf{J}_{k|k}^{2,1}\right)^{-1} \mathbf{K}_{k+1}^{k,k+1}. \end{aligned} \tag{12}$$

With (11) and $\mathbf{J}_{k|k-1}^{1,1} = \mathbf{J}_{k|k}^{1,1}$, $\mathbf{J}_{k|k-1}^{1,2} = \mathbf{J}_{k|k}^{1,2}$ and $\mathbf{J}_{k|k-1}^{2,2} = \mathbf{K}_k^k$ this can be reduced to

$$\mathbf{C}_{k+1|k}^{-1} = \mathbf{K}_{k+1}^{k+1} - \mathbf{K}_{k+1}^{k+1,k}\left(\mathbf{K}_{k+1}^{k} + \mathbf{C}_{k|k}^{-1}\right)^{-1}\mathbf{K}_{k+1}^{k,k+1}. \tag{13}$$

With (12) the measurement update can also be reduced to

$$\mathbf{C}_{k|k}^{-1} = \mathbf{C}_{k|k-1}^{-1} + \mathbf{L}_k^k, \tag{14}$$

which completes the recursion to compute the CRB for the time and measurement update.

3. CONSTRAINED CRB

In this section, the constrained CRB is introduced. This section will closely follow the derivation presented in Stoica and Ng (1998). However, the difference between this paper and Stoica and Ng (1998) is that Stoica and Ng (1998) assumed that a vector of non-random parameters

40 pt
0.556 in
14.1 mm

40 pt
0.556 in
14.1 mm

68 pt
0.944 in
24 mm

Margin requirements for the other pages
Paper size this page A4

68 pt
0.944 in
24 mm

is estimated. In this paper, on the other hand, a vector of random parameters $\mathbf{X}$ based on a vector of observations $\mathbf{Z}$ is estimated. It is required that the estimate $\hat{\mathbf{X}}$ satisfies l ($l < n$) continuously differentiable constraints,

$$\mathbf{f}(\hat{\mathbf{X}}) = 0. \tag{15}$$

It is further assumed that the set $\{\mathbf{X} | f(\mathbf{X}) = 0\}$ is non-empty. The gradient matrix of the constraints can be defined as

$$\mathbf{F} = \frac{\partial \mathbf{f}(\mathbf{X})}{\partial \mathbf{X}^T}, \tag{16}$$

where it is assumed that (16) has full rank for any $\mathbf{X}$ satisfying (15). Thus, there exist an $n \times (n-l)$ matrix $\mathbf{U}$ such that

$$\mathbf{F}\mathbf{U} = 0, \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}, \tag{17}$$

where it is assumed that $\mathbf{U}$ is independent of $\mathbf{X}$. While this is restrictive, it always holds for linear constraints. The likelihood function is given by (2), and let us denote

$$\mathbf{\Delta} = \nabla_{\mathbf{X}} \ln p(\mathbf{X}, \mathbf{Z}). \tag{18}$$

In that case, the FIM (3) is alternatively given as

$$\mathbf{J} = E(\mathbf{\Delta}\mathbf{\Delta}^T). \tag{19}$$

If the following conditions exist

(1) $\partial \frac{p(\mathbf{X},\mathbf{Z})}{\partial x_i^j}$ is absolutely integrable with respect to $\mathbf{X}$ and $\mathbf{Z}$ for $i = 0, \ldots, k$ and $j = 0, \ldots, n$.
(2) $\partial^2 \frac{p(\mathbf{X},\mathbf{Z})}{\partial x_i^j}$ is absolutely integrable with respect to $\mathbf{X}$ and $\mathbf{Z}$ for $i = 0, \ldots, k$ and $j = 0, \ldots, n$.
(3) The conditional expectation of the error, given $\mathbf{X}$, is

$$\mathbf{B}(\mathbf{X}) = \int_{-\infty}^{\infty} [\hat{\mathbf{X}} - \mathbf{X}] p(\mathbf{Z}|\mathbf{X}) d\mathbf{Z}, \tag{20}$$

where $\mathbf{B}(\mathbf{X})$ denotes the bias of the estimate. It is assumed that

$$\begin{aligned} &\lim_{x_i^j \to \infty} \mathbf{B}(\mathbf{X}) p(\mathbf{X}) = 0, && \text{for } i = 0, \ldots, k \\ &\lim_{x_i^j \to -\infty} \mathbf{B}(\mathbf{X}) p(\mathbf{X}) = 0, && \text{and } j = 0, \ldots, n \end{aligned} \tag{21}$$

With the conditions it can be shown that the following holds (Van Trees and Bell, 2013)

$$E((\hat{\mathbf{X}} - \mathbf{X})\mathbf{\Delta}^T) = \mathbf{I}, \tag{22}$$

which is the fact that is required such that

$$E((\hat{\mathbf{X}} - \mathbf{X})\mathbf{\Delta}^T)\mathbf{U}\mathbf{U}^T = \mathbf{U}\mathbf{U}^T. \tag{23}$$

Lemma 1. If the condition holds, with $\mathbf{U}$ defined in (17) and in case $\mathbf{U}^T\mathbf{J}\mathbf{U}$ is non-singular the constrained CRB is given as

$$E\left\{(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^T\right\} \geq \mathbf{U}(\mathbf{U}^T\mathbf{J}\mathbf{U})^{-1}\mathbf{U}^T. \tag{24}$$

Proof. Let $\mathbf{W}$ be an arbitrary $n \times n$ matrix. Then

$$\begin{aligned} E\Big\{&(\mathbf{X} - \hat{\mathbf{X}} - \mathbf{W}\mathbf{U}\mathbf{U}^T\mathbf{\Delta})(\mathbf{X} - \hat{\mathbf{X}} - \mathbf{W}\mathbf{U}\mathbf{U}^T\mathbf{\Delta})^T\Big\} \\ &= E\Big\{(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^T\Big\} - \mathbf{W}\mathbf{U}\mathbf{U}^T - \\ &\quad \mathbf{U}\mathbf{U}^T\mathbf{W}^T + \mathbf{W}\mathbf{U}\mathbf{U}^T\mathbf{J}\mathbf{U}\mathbf{U}^T\mathbf{W}^T \geq 0, \end{aligned} \tag{25}$$

where the equality follows from (23) and the fact that $\mathbf{U}$ is independent of $\mathbf{X}$. The inequality is a consequence of the positive semi-definiteness of the covariance matrix of $\mathbf{X} - \hat{\mathbf{X}} - \mathbf{W}\mathbf{U}\mathbf{U}^T\mathbf{\Delta}$. For the rest of the proof the derivation in Stoica and Ng (1998) can be followed.

4. THE CONSTRAINED CRB FOR THE NON-LINEAR FILTERING PROBLEM

Theorem 2. The constrained CRB for the non-linear filtering problem can be computed by a time update

$$\mathbf{C}_{k|k-1}^{-1} = \mathbf{K}_k^k - \mathbf{K}_k^{k,k-1}\left(\mathbf{K}_k^{k-1} + \tilde{\mathbf{C}}_{k-1|k-1}^{-1}\right)^{-1}\mathbf{K}_k^{k-1,k}, \tag{26}$$

a measurement update

$$\mathbf{C}_{k|k}^{-1} = \mathbf{C}_{k|k-1}^{-1} + \mathbf{L}_k^k, \tag{27}$$

and a constraint update

$$\tilde{\mathbf{C}}_{k|k} = \mathbf{U}_k(\mathbf{U}_k^T\mathbf{C}_{k|k}^{-1}\mathbf{U}_k)^{-1}\mathbf{U}_k^T. \tag{28}$$

Proof. It will be shown that the constrained CRB of the whole system (4) results in the same CRB as given by the recursion (26)-(28). This will be shown for the first time step and afterwards in will be shown that this holds also for every following time step.

The FIM for whole state and measurement history after the first time step is given by the following $2n \times 2n$ matrix

$$\mathbf{J}_{1|1}(\mathbf{X}_1) = \begin{pmatrix} \mathbf{K}_0^0 + \mathbf{L}_0^0 + \mathbf{K}_1^0 & \mathbf{K}_1^{0,1} \\ \mathbf{K}_1^{1,0} & \mathbf{L}_1^1 + \mathbf{K}_1^1. \end{pmatrix} \tag{29}$$

It is assumed that the constraints (15) depend only on the states at each time step. This gives a $2n \times 2(n-l)$ matrix

$$\mathbf{U}(\mathbf{X}_1) = \begin{pmatrix} \mathbf{U}(\mathbf{x}_0) & 0 \\ 0 & \mathbf{U}(\mathbf{x}_1) \end{pmatrix} = \begin{pmatrix} \mathbf{U}_0 & 0 \\ 0 & \mathbf{U}_1, \end{pmatrix} \tag{30}$$

where the matrix $\mathbf{U}_i$ of each state is on the diagonal while all other entries are zero. Using (24), (30) and the matrix inversion lemma (A.1) on (29) to compute the constrained CRB for state $\mathbf{x}_1$ we obtain

$$\begin{aligned} \tilde{\mathbf{C}}_{1|1} = \mathbf{U}_1\Big[&\mathbf{U}_1^T\left(\mathbf{L}_1^1 + \mathbf{K}_1^1\right)\mathbf{U}_1 - \\ &\mathbf{U}_1^T\mathbf{K}_1^{1,0}\mathbf{U}_0\left[\mathbf{U}_0^T\left(\mathbf{K}_0^0 + \mathbf{L}_0^0 + \mathbf{K}_1^0\right)\mathbf{U}_0\right]^{-1}\mathbf{U}_0^T\mathbf{K}_1^{0,1}\mathbf{U}_1\Big]^{-1}\mathbf{U}_1^T. \end{aligned} \tag{31}$$

With the recursion (26)-(28) the constrained CRB for state $\mathbf{x}_1$ is

$$\begin{aligned} \tilde{\mathbf{C}}_{1|1} =& \mathbf{U}_1 \cdot \\ &\left[\mathbf{U}_1^T\left(\mathbf{L}_1^1 + \mathbf{K}_1^1 - \mathbf{K}_1^{1,0}\left[\mathbf{J}_{0|0} + \mathbf{K}_1^0\right]^{-1}\mathbf{K}_1^{0,1}\right)\mathbf{U}_1\right]^{-1}\mathbf{U}_1^T. \end{aligned} \tag{32}$$

Comparing (31) and (32) it can be seen that the only thing left to prove is

$$\mathbf{U}_0\left[\mathbf{U}_0^T\left(\mathbf{K}_0^0 + \mathbf{L}_0^0 + \mathbf{K}_1^0\right)\mathbf{U}_0\right]^{-1}\mathbf{U}_0^T \equiv \left[\mathbf{J}_{0|0} + \mathbf{K}_1^0\right]^{-1} \tag{33}$$

The left-hand side of (33) is rewritten slightly and the binomial inverse theorem (B.1) applied

$$\begin{aligned} &\mathbf{U}_0\left[\mathbf{U}_0^T\left(\mathbf{K}_0^0 + \mathbf{L}_0^0\right)\mathbf{U}_0 + \mathbf{U}_0^T\mathbf{K}_1^0\mathbf{U}_0\right]^{-1}\mathbf{U}_0^T \\ =&\mathbf{U}_0\Bigg[\left(\mathbf{U}_0^T\left(\mathbf{K}_0^0 + \mathbf{L}_0^0\right)\mathbf{U}_0\right)^{-1} - \left(\mathbf{U}_0^T\left(\mathbf{K}_0^0 + \mathbf{L}_0^0\right)\mathbf{U}_0\right)^{-1}\mathbf{U}_0^T \cdot \\ &\left(\mathbf{I} + \mathbf{K}_1^0\mathbf{U}_0\left(\mathbf{U}_0^T\left(\mathbf{K}_0^0 + \mathbf{L}_0^0\right)\mathbf{U}_0\right)^{-1}\mathbf{U}_0^T\right)^{-1}\mathbf{K}_1^0\mathbf{U}_0 \cdot \\ &\left(\mathbf{U}_0^T\left(\mathbf{K}_0^0 + \mathbf{L}_0^0\right)\mathbf{U}_0\right)^{-1}\Bigg]\mathbf{U}_0^T, \end{aligned} \tag{34}$$

where $\mathbf{B} = \mathbf{U}_0^T$, $\mathbf{D} = \mathbf{K}_1^0$ and $\mathbf{C} = \mathbf{U}_0$. Given that

$$[\mathbf{J}_{0|0}]^{-1} = \mathbf{U}_0\left(\mathbf{U}_0^T\left(\mathbf{K}_0^0 + \mathbf{L}_0^0\right)\mathbf{U}_0\right)^{-1}\mathbf{U}_0^T, \tag{35}$$

40 pt
0.556 in
14.1 mm

40 pt
0.556 in
14.1 mm

68 pt
0.944 in
24 mm

Margin requirements for the other pages
Paper size this page A4
68 pt 0.944 in 24 mm
40 pt 0.556 in 14.1 mm
40 pt 0.556 in 14.1 mm
68 pt 0.944 in 24 mm

the binomial inverse theorem (B.1) is also applied to the right-hand side of (33)

$$\begin{aligned}&\left[\mathbf{J}_{0|0}+\mathbf{K}_1^0\right]^{-1}\\ =&\mathbf{U}_0\left(\mathbf{U}_0^T\left(\mathbf{K}_0^0+\mathbf{L}_0^0\right)\mathbf{U}_0\right)^{-1}\mathbf{U}_0^T-\mathbf{U}_0\left(\mathbf{U}_0^T\left(\mathbf{K}_0^0+\mathbf{L}_0^0\right)\mathbf{U}_0\right)^{-1}\mathbf{U}_0^T\cdot\\ &\mathbf{B}\left(\mathbf{I}+\mathbf{DCU}_0\left(\mathbf{U}_0^T\left(\mathbf{K}_0^0+\mathbf{L}_0^0\right)\mathbf{U}_0\right)\mathbf{B}\right)^{-1}\mathbf{DCU}_0\\ &\mathbf{U}_0\left(\mathbf{U}_0^T\left(\mathbf{K}_0^0+\mathbf{L}_0^0\right)\mathbf{U}_0\right)^{-1}\mathbf{U}_0^T,\end{aligned} \tag{36}$$

where $\mathbf{B}=\mathbf{I}$, $\mathbf{D}=\mathbf{K}_1^0$ and $\mathbf{C}=\mathbf{I}$. Comparing (34) and (36) we see that (33) holds.
For k time steps it can be obtained in a similar fashion as before (29) - (32) that

$$\mathbf{U}_{k-1}\left[\mathbf{U}_{k-1|0}^T\mathbf{J}_{k|k-1}^{1,1}\mathbf{U}_{k-1|0}\right]\mathbf{U}_{k-1}^T\equiv[\mathbf{J}_{k-1|k-1}+\mathbf{K}_{|}^0]^{-1}. \tag{37}$$

Applying the matrix inversion lemma to the left-hand side plus the fact that the matrix $\mathbf{U}$ is block-diagonal and setting in the recursive equations to the right-hand side, (37) can be reduced to (33).

5. REDUCTION OF THE CONSTRAINED CRB

In this section, it is established that the unconstrained CRB (4) is never smaller than the constrained CRB (24). Consequently, adding information about constraints reduces the CRB. This can be shown in case $\mathbf{J}_{k|k}$ is positive definite. In that case as shown in Stoica and Ng (1998); Khatri (1966) the constrained CRB (24) can be written as

$$\mathbf{U}(\mathbf{U}^T\mathbf{J}\mathbf{U})^{-1}\mathbf{U}^T=\left(\mathbf{I}-\mathbf{J}^{-1}\mathbf{F}^T(\mathbf{F}\mathbf{J}^{-1}\mathbf{F}^T)^{-1}\mathbf{F}\right)\mathbf{J}^{-1}=\mathbf{Q}\mathbf{J}^{-1}, \tag{38}$$

where $\mathbf{F}$ is given in (16) and indices and arguments were omitted for national convenience. It can be shown that $\mathbf{Q}$ is idempotent which means

$$\mathbf{Q}=\mathbf{Q}\mathbf{Q}. \tag{39}$$

Furthermore, $\mathbf{I}-\mathbf{Q}$ is idempotent and $\mathbf{J}^{-1}$ as well as $\mathbf{Q}\mathbf{J}^{-1}$ are symmetric. With the conditions, it can be proven that the constrained CRB reduces in the sense that the subtraction of constraint from the unconstrained CRB is non-negative definite (Gorman and Hero, 1990).

$$\begin{aligned}\mathbf{Q}\mathbf{J}^{-1}&=\mathbf{J}^{-1}-(\mathbf{I}-\mathbf{Q})\mathbf{J}^{-1}\\ &=\mathbf{J}^{-1}-(\mathbf{I}-\mathbf{Q})(\mathbf{I}-\mathbf{Q})\mathbf{J}^{-1}\\ &=\mathbf{J}^{-1}-(\mathbf{I}-\mathbf{Q})\mathbf{J}^{-1}(\mathbf{I}-\mathbf{Q})\\ &\leq\mathbf{J}^{-1}\end{aligned} \tag{40}$$

6. COMPUTATION OF THE CONSTRAINED CRB

The constrained CRB can be computed with the matrices (5) and the recursive equations (26) - (28). In case of a positive definite FIM $\mathbf{J_{k|k}}$ (28) can be exchanged with (38) which may reduce the computational burden. Important to consider is that the constrained CRB is singular. Consequently, the constrained CRB is not invertible. This is not an issue in the recursion, but (26) can only be computed with the help of a matrix inversion lemma like (A.3).

6.1 Linear Gaussian case

Let us consider the linear system with additive Gaussian state and measurement noise with linear equality constraints

$$\mathbf{x}_{k+1}=\mathbf{F}_k\mathbf{x}_k+\mathbf{w}_k \tag{41a}$$
$$\mathbf{z}_k=\mathbf{H}_k\mathbf{x}_k+\mathbf{v}_k, \tag{41b}$$
$$0=\mathbf{A}_k\mathbf{x}_k, \tag{41c}$$

where $\mathbf{A}_k$ is the gradient matrix of the constraints. The process covariance and measurement covariance is given by the positive definite matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$, respectively. For this special case, the matrices (5) have an analytical solution

$$\mathbf{K}_{k+1}^k=\mathbf{F}_k^T\mathbf{Q}_k^{-1}\mathbf{F}_k, \tag{42a}$$
$$\mathbf{K}_{k+1}^{k,k+1}=-\mathbf{F}_k^T\mathbf{Q}_k^{-1}=[\mathbf{K}_{k+1}^{k+1,k}]^T, \tag{42b}$$
$$\mathbf{K}_{k+1}^{k+1}=\mathbf{Q}_k^{-1}, \tag{42c}$$
$$\mathbf{L}_k^k=\mathbf{H}_k^T\mathbf{R}^{-1}\mathbf{H}. \tag{42d}$$

Using the matrix inversion lemma (A.3) the recursive equations (26) and (27) can be computed as

$$\mathbf{C}_{k|k-1}=\mathbf{Q}_k+\mathbf{F}_k\tilde{\mathbf{C}}_{k-1|k-1}\mathbf{F}_k^T, \tag{43a}$$
$$\mathbf{C}_{k|k}=\mathbf{C}_{k|k-1}-\mathbf{C}_{k|k-1}\mathbf{H}^T\left(\mathbf{R}+\mathbf{H}\mathbf{C}_{k|k-1}\mathbf{H}^T\right)^{-1}\mathbf{H}\mathbf{C}_{k|k-1}, \tag{43b}$$

which are the Kalman filter time and measurement update equations (Simon, 2006). The constrained update step (28) should be computed using the right-hand side of (38)

$$\tilde{\mathbf{C}}_{k|k}=\mathbf{C}_{k|k}-\mathbf{C}_{k|k}\mathbf{A}_k^T\left(\mathbf{A}_k\mathbf{C}_{k|k}\mathbf{A}_k^T\right)^{-1}\mathbf{A}_k\mathbf{C}_{k|k}, \tag{44}$$

since it avoids the otherwise needed inversion of (43b).

7. NUMERICAL EXAMPLE

In Simon (2010) a four state navigation problem with equality constraints is presented. The first two states present the positions and the last two states the velocities in north and east direction, respectively. The velocity of the vehicle is in the direction of θ, an angle measured anti-clockwise from the east axis. A sensor provides noisy measurement of the vehicle's north and east positions. The equations for this system can be written as

$$\mathbf{x}_{k+1}=\begin{pmatrix}1&0&T&0\\0&1&0&T\\0&0&1&0\\0&0&0&1\end{pmatrix}\mathbf{x}_k+\begin{pmatrix}0\\0\\T\sin\theta\\T\cos\theta\end{pmatrix}u_k+\mathbf{w}_k, \tag{45a}$$
$$\mathbf{y}_k=\begin{pmatrix}1&0&0&0\\0&1&0&0\end{pmatrix}\mathbf{x}_k+\mathbf{v}_k, \tag{45b}$$

where T is the discretization step size and u_k is the acceleration input. The covariance of process and measurement noise are

$$\mathbf{Q}=\mathrm{diag}(4,4,1,1),\quad \mathbf{R}=\mathrm{diag}(900,900) \tag{46}$$

and the initial estimation error covariance is

$$\mathbf{P}_0^+=\mathrm{diag}(900,900,4,4), \tag{47}$$

where it holds

$$[\mathbf{P}_0^+]^{-1}=\mathbf{L}_0^0+\mathbf{K}_0^0. \tag{48}$$

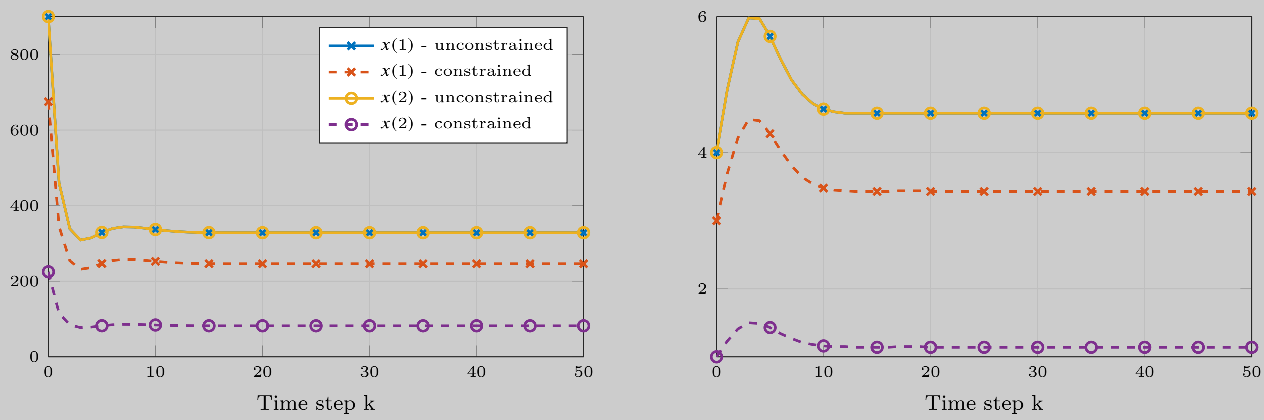
Since we know that the vehicle is on a road with a heading angle θ the following holds

$$\tan\theta=\frac{x(1)}{x(2)}=\frac{x(3)}{x(4)}. \tag{49}$$

The constraints of the system can be expressed in the form $\mathbf{D}_i\mathbf{x}_k$ using either

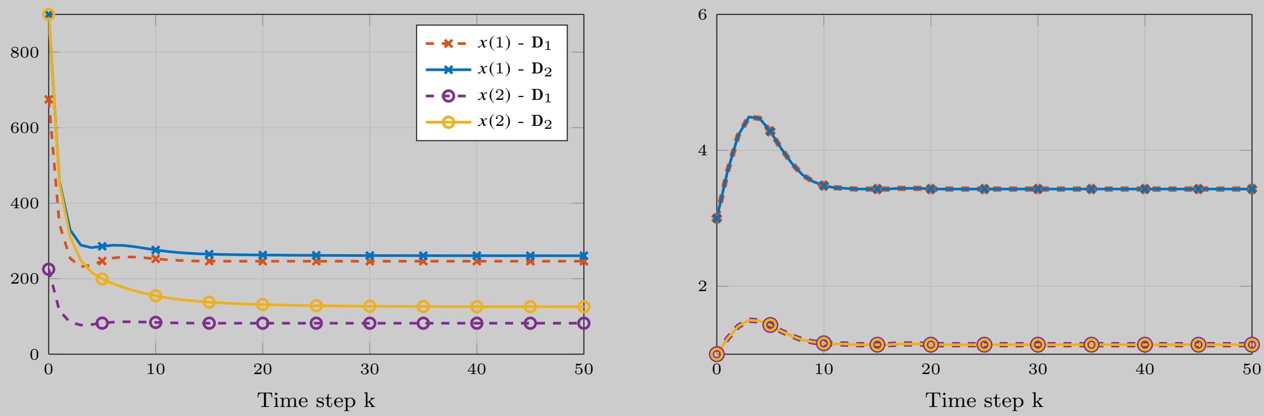
$$\mathbf{D}_1=\begin{pmatrix}1&-\tan\theta&0&0\\0&0&1&-\tan\theta\end{pmatrix} \tag{50}$$

Margin requirements for the other pages
Paper size this page A4
68 pt
0.944 in
24 mm



(a) Positions (b) Velocities

Fig. 1. Constrained and unconstrained CRB for the state components of the navigation example.

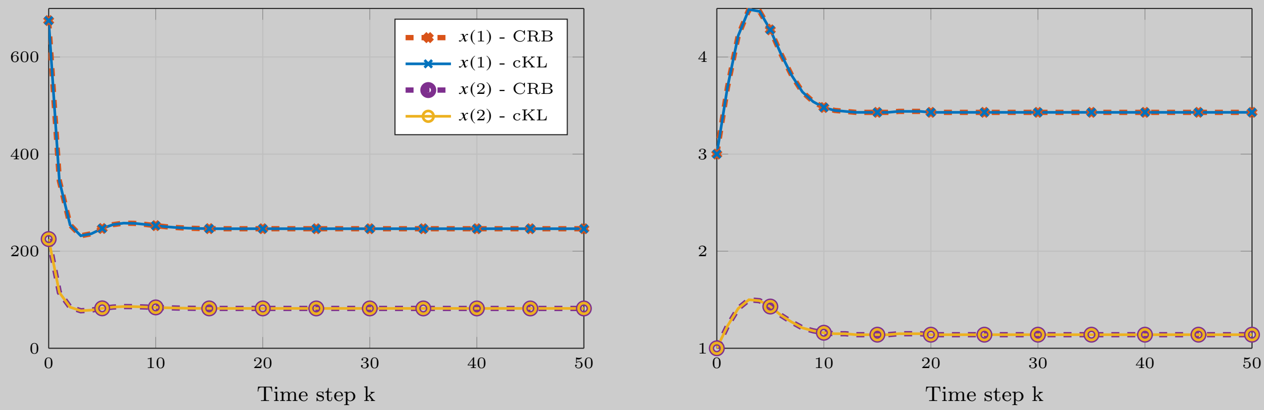


(a) Positions (b) Velocities

40 pt
0.556 in
14.1 mm

40 pt
0.556 in
14.1 mm

Fig. 2. Constrained CRB for the state components of the navigation example using either $\mathbf{D}_1$ or $\mathbf{D}_2$.



(a) Positions (b) Velocities

Fig. 3. Constrained CRB and error covariance of constrained Kalman filter (cKL) using system projection method. The system is constrained using $\mathbf{D}_1$.

or

$$\mathbf{D}_2 = \begin{pmatrix} 0 & 0 & 1 & -\tan\theta \end{pmatrix}. \tag{51}$$

Consequently, the system has either one or two equality constraints.

The recursive equations (43)-(44) are used to compute the constrained CRB for the system with $\mathbf{D}_1$ as constrained gradient matrix of the system. The CRB for all states is shown in Fig. 1. The CRB for the constrained system reduces since the constrained information is used to compute the bound. In a similar way the constrained CRB for the system using $\mathbf{D}_2$ as gradient matrix of the constraints can be computed. Since $\mathbf{D}_1$ provides more information about the constraints on the position while the same information as $\mathbf{D}_2$ about the velocities, the constrained CRB using

68 pt
0.944 in
24 mm

Margin requirements for the other pages
Paper size this page A4
68 pt
0.944 in
24 mm

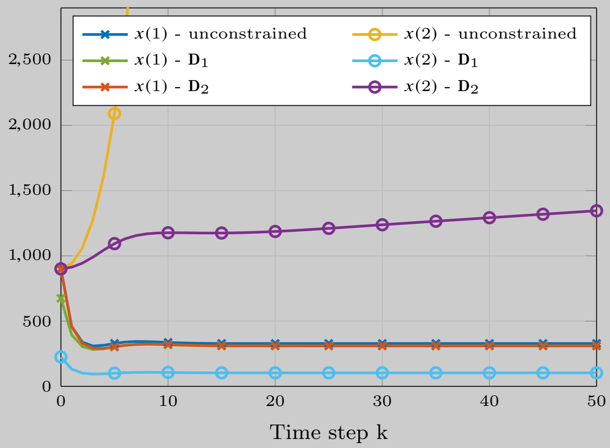


Fig. 4. CRB for the state components of the position when only $x(1)$ is measured.

$\mathbf{D}_1$ is lower as the one using $\mathbf{D}_2$ for the position estimates (Fig. 2a) while the same for the velocity estimates (Fig. 2b).

The computed constrained CRB can be reached with equality if the constrained Kalman filter using the system projection method presented by Simon (2010) is used (Fig. 3). In this case, the initial estimation error covariance and the process noise covariance are projected onto the constrained surface such that they are consistent with the state constraints. Of all constrained linear estimators considered in Simon (2010) the Kalman filter using the system projection method reached the lowest error covariance for the presented problem.

Consider the problem where the sensor has a fault and only is able to provide a measurement of the vehicle's north position. It can be easily seen that the unconstrained problem becomes unobservable. This is not the case for the constrained problem using $\mathbf{D}_1$. However, the problem is also unobservable if $\mathbf{D}_2$ is used since the constrained CRB of $x(2)$ grows without bounds (Fig. 4). The CRB of $x(2)$ of the unconstrained case grows quickly, while the CRB of $x(2)$ using $\mathbf{D}_2$ considerably slower to infinity.

40 pt
0.556 in
14.1 mm

8. CONCLUSION

A simple recursive algorithm to compute the CRB for the non-linear filtering problem under linear state constraints was presented. The CRB version of previous literature was extended by one additional step that projects the unconstrained CRB onto the tangent hyperplane of the constrained set. Two different equations were presented to perform the constraint update. It was shown that this update step reduces the CRB compared to the unconstrained CRB. The constrained CRB was illustrated on a navigation problem, where it was shown that the constrained CRB reduces compared to the unconstrained one. Moreover, it was shown that a Kalman filter using a system projection method reaches the constrained CRB for the considered example. In addition, it was demonstrated that the constrained CRB can easily be used to investigate observability of a constrained problem.

Appendix A. MATRIX INVERSION LEMMA

The inverse of a 2×2 block matrix is given by

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{F}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{E}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{F} & \mathbf{E} \end{pmatrix}, \qquad \text{(A.1)}$$

where

$$\mathbf{E} = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B},$$
$$\mathbf{F} = \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}. \qquad \text{(A.2)}$$

Furthermore, the following holds

$$(\mathbf{A} + \mathbf{B}\mathbf{D}\mathbf{C})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}^{-1} + \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1}, \qquad \text{(A.3)}$$

provided $\mathbf{A}^{-1}$ exists.

Appendix B. BINOMIAL INVERSE THEOREM

The binomial inverse theorem is a more general formula of (A.3), which also exist in cases of a singular matrix $\mathbf{D}$ (Henderson and Searle, 1981)

$$(\mathbf{A}^{-1} + \mathbf{B}\mathbf{D}\mathbf{C})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{I} + \mathbf{D}\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{D}\mathbf{C}\mathbf{A}^{-1}. \qquad \text{(B.1)}$$

40 pt
0.556 in
14.1 mm

REFERENCES

Ben-Haim, Z. and Eldar, Y.C. (2009). On the constrained cramér–rao bound with a singular fisher information matrix. *IEEE Signal Processing Letters*, 16(6), 453–456.

Bobrovsky, B. and Zakai, M. (1975). A lower bound on the estimation error for markov processes. *IEEE Transactions on Automatic Control*, 20(6), 785–788.

Galdos, J. (1980). A cramér-rao bound for multidimensional discrete-time dynamical systems. *IEEE Transactions on Automatic Control*, 25(1), 117–119.

Gorman, J.D. and Hero, A.O. (1990). Lower bounds for parametric estimation with constraints. *IEEE Transactions on Information Theory*, 36(6), 1285–1301.

Henderson, H.V. and Searle, S.R. (1981). On deriving the inverse of a sum of matrices. *Siam Review*, 23(1), 53–60.

Jagannatham, A.K. and Rao, B.D. (2004). Cramer-rao lower bound for constrained complex parameters. *IEEE Signal Processing Letters*, 11(11), 875–878.

Khatri, C.G. (1966). A note on a manova model applied to problems in growth curve. *Annals of the Institute of Statistical Mathematics*, 18(1), 75–86.

Marzetta, T.L. (1993). A simple derivation of the constrained multiple parameter cramer-rao bound. *IEEE Transactions on Signal Processing*, 41(6), 2247–2249.

Šimandl, M., Královec, J., and Tichavskỳ, P. (2001). Filtering, predictive, and smoothing cramér–rao bounds for discrete-time nonlinear dynamic systems. *Automatica*, 37(11), 1703–1716.

Simon, D. (2010). Kalman filtering with state constraints: a survey of linear and nonlinear algorithms. *Control Theory & Applications, IET*, 4(8), 1303–1318.

Simon, D. (2006). *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley.

Stoica, P. and Ng, B.C. (1998). On the cramér-rao bound under parametric constraints. *IEEE Signal Processing Letters*, 5(7), 177–179.

Tichavsky, P., Muravchik, C.H., and Nehorai, A. (1998). Posterior cramér-rao bounds for discrete-time nonlinear filtering. *IEEE Transactions on signal processing*, 46(5), 1386–1396.

Van Trees, H.L. (1968). *Detection, estimation, and modulation theory*. John Wiley & Sons.

Van Trees, H.L. and Bell, K.L. (2013). *Detection estimation and modulation theory, pt. I*. Wiley.

68 pt
0.944 in
24 mm