The Kalman filter is the optimal minimum-variance state estimator for linear dynamic systems with Gaussian noise. In addition, the Kalman filter is the optimal linear state estimator for linear dynamic systems with non-Gaussian noise. For nonlinear systems various modifications of the Kalman filter (e.g., the extended Kalman filter, the unscented Kalman filter, and the particle filter) have been proposed as approximations to the optimal state estimator, which (in general) cannot be solved analytically. One reason that Kalman filtering is optimal for linear systems is that it uses all the available information about the system in order to obtain a state estimate.

However, in the application of state estimators, there may be known information about the system that the standard Kalman filter does not incorporate. For example, there may be state constraints (equality or inequality) that must be satisfied. The standard Kalman filter does not incorporate state constraints, and therefore it is suboptimal since it does not use all of the available information. In cases like these we can modify the Kalman filter to exploit this additional information.

There are a variety of ways to use constraint information to modify the Kalman filter, and we discuss many of them in this paper. If both the system and state constraints are linear, then
all of these different approaches result in the same state estimate, which is in fact the optimal state estimate subject to the constraints. (This is analogous to the many different derivations of the Kalman filter for standard unconstrained systems, all of which result in the same filter.) If either the system or constraints are nonlinear, then constrained Kalman filtering is suboptimal, and different approaches give different results. (This is analogous to the many different Kalman filter approximations for nonlinear systems, all of which are suboptimal, and all of which give estimation performance that is problem-dependent.)

Many examples of state-constrained systems are found in practical engineering problems. These include chemical processes [1], vision-based systems [2], [3], target tracking [4], [5], biomedical systems [6], robotics [7], navigation [8], fault diagnosis [9], and many others [10]. Constrained Kalman filtering is thus becoming a focus of increased attention in both academia and industry; see “Constrained Kalman Filtering Research.” This paper presents a survey of how state constraints can be incorporated into the Kalman filter. We discuss linear and nonlinear systems, linear and nonlinear state constraints, and equality and inequality state constraints.

**INTRODUCTION**

Consider the system model

\[
\begin{align*}
x_{k+1} &= Fx_k + w_k, \\
y_k &= Hx_k + v_k,
\end{align*}
\]

(1)

where \( k \) is the time step, \( x_k \) is the state, \( y_k \) is the measurement, \( w_k \) and \( v_k \) are the process noise and measurement noise (zero-mean with covariances \( Q \) and \( R \) respectively), and \( F \) and \( H \) are
the state transition and measurement matrices. The Kalman filter is initialized with

\[
\hat{x}_0^+ = E(x_0), \\
P_0^+ = E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T],
\]

where \( E(\cdot) \) is the expectation operator. The Kalman filter is given by the following equations [11], which are computed for each time step \( k = 1, 2, \ldots \)

\[
P_k^- = F P_{k-1}^+ F^T + Q, \\
K_k = P_k^- H^T (H P_k^- H^T + R)^{-1}, \\
\hat{x}_k^- = F \hat{x}_{k-1}^+, \\
\hat{x}_k^+ = \hat{x}_k^- + K_k (y_k - H \hat{x}_k^-), \\
P_k^+ = (I - K_k H) P_k^-,
\]

where \( I \) is the identity matrix. \( \hat{x}_k^- \) is the \textit{a priori} estimate; it is the best estimate of the state \( x_k \) given measurements up to and including time \( k - 1 \). \( \hat{x}_k^+ \) is the \textit{a posteriori} estimate; it is the best estimate of the state \( x_k \) given measurements up to and including time \( k \). \( K_k \) is called the Kalman gain. \( P_k^- \) is the covariance of the \textit{a priori} estimation error \( (x_k - \hat{x}_k^-) \), and \( P_k^+ \) is the covariance of the \textit{a posteriori} estimation error \( (x_k - \hat{x}_k^+) \).

The Kalman filter is attractive for its computational simplicity and its theoretical rigor. If the noise sequences \( \{w_k\} \) and \( \{v_k\} \) are Gaussian, uncorrelated, and white, then the Kalman filter is the filter that minimizes the two-norm of the estimation error covariance at each time step. If \( \{w_k\} \) and \( \{v_k\} \) are non-Gaussian, the Kalman filter is still the optimal \textit{linear} filter. If \( \{w_k\} \) and \( \{v_k\} \) are correlated or colored, the Kalman filter can be easily modified so that it is still optimal [11].
LINEAR CONSTRAINTS

Suppose that we want to find a state estimate \( \hat{x}_k \) that satisfies the constraints \( D\hat{x}_k = d \), or \( D\hat{x}_k \leq d \), where \( D \) is a known matrix and \( d \) is a known vector. That is, we have equality and/or inequality constraints on the state. If there are no constraints then the Kalman filter is the optimal linear estimator. But since we have state constraints, the Kalman filter can be modified to obtain better results.

Model reduction

State equality constraints can be addressed by reducing the system model parameterization [12]. As an example, consider the system

\[
\begin{align*}
    x_{k+1} &= \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 4 & -2 & 2 \end{bmatrix} x_k + \begin{bmatrix} w_{1k} \\ w_{2k} \\ w_{3k} \end{bmatrix}, \\
    y_k &= \begin{bmatrix} 2 & 4 & 5 \end{bmatrix} x_k + v_k.
\end{align*}
\]

(4)

Suppose that we also have the constraint

\[
\begin{bmatrix} 1 & 0 & 1 \end{bmatrix} x_k = 0.
\]

(5)

If we make the substitution \( x_k(3) = -x_k(1) \) in the state and measurement equations, we obtain

\[
\begin{align*}
    x_{k+1}(1) &= -2x_k(1) + 2x_k(2) + w_{1k}, \\
    x_{k+1}(2) &= 2x_k(1) + 2x_k(2) + w_{2k}, \\
    y_k &= -3x_k(1) + 4x_k(2) + v_k.
\end{align*}
\]

(6)
These equations can be written in matrix form as

\[
\begin{align*}
    x_{k+1} &= \begin{bmatrix} -2 & 2 \\ 2 & 2 \end{bmatrix} x_k + \begin{bmatrix} w_{k1} \\ w_{k2} \end{bmatrix}, \\
    y_k &= \begin{bmatrix} -3 & 4 \end{bmatrix} x_k + v_k.
\end{align*}
\] (7)

We have reduced the filtering problem with equality constraints to an equivalent but unconstrained filtering problem. The Kalman filter for this unconstrained problem is the optimal linear estimator, and so it is also the optimal linear estimator for the original constrained problem. The dimension of the problem has been reduced, and so the computational effort of the problem is reduced. One disadvantage of this approach is that the physical meaning of the state variables has been lost. Also this approach cannot be directly used for inequality constraints.

**Perfect Measurements**

State equality constraints can be treated as perfect measurements (measurements with zero measurement noise) [3]–[5]. If the constraints are given as \( D x_k = d \), where \( D \) is a known \( s \times n \) matrix \((s < n)\), and \( d \) is a known vector, then we can solve the constrained Kalman filtering problem by augmenting the measurement equation with \( s \) perfect measurements of the state.

\[
\begin{align*}
    x_{k+1} &= F x_k + w_k, \\
    \begin{bmatrix} y_k \\ d \end{bmatrix} &= \begin{bmatrix} H & D \end{bmatrix} \begin{bmatrix} x_k \\ 0 \end{bmatrix} + \begin{bmatrix} v_k \\ 0 \end{bmatrix}.
\end{align*}
\] (8)

The state equation has not changed, but the measurement equation has been augmented. The fact that the last \( s \) elements of the measurement equation are noise-free means that the *a posteriori*
Kalman filter estimate of the state will be consistent with these \( s \) measurements [13]; that is, the Kalman filter estimate will satisfy \( D \hat{x}_k^+ = d \). This approach is mathematically identical to the model reduction approach.

Note that the new measurement noise covariance will be singular. A singular noise covariance does not present theoretical problems [14]. In fact, Kalman’s original paper [15] presents an example that uses perfect measurements. However, in practice a singular noise covariance increases the possibility of numerical problems [16, p. 249], [17, p. 365]. Also the use of perfect measurements is not directly applicable to inequality constraints.

**Estimate Projection**

Another approach to constrained filtering is to begin with the standard unconstrained estimate \( \hat{x}_k^+ \) and project it onto the constraint surface [8], [18]. This can be written as

\[
\tilde{x}_k^+ = \arg\min_x (x - \hat{x}_k^+)^T W (x - \hat{x}_k^+) \text{ such that } Dx = d,
\]

(9)

where \( W \) is any positive-definite weighting matrix. The solution to this problem is

\[
\tilde{x}_k^+ = \hat{x}_k^+ - W^{-1} D^T (D W^{-1} D^T)^{-1} (D \hat{x}_k^+ - d).
\]

(10)

If the process and measurement noise is Gaussian and we set \( W = (P_k^+)^{-1} \) we obtain the maximum probability estimate of the state subject to state constraints. If we set \( W = I \) we obtain the least squares estimate of the state subject to state constraints. (This is similar to the approach used in [19] for input signal estimation.) It is shown in [8], [18] that the constrained state estimate of (10) has several interesting properties.

1) The constrained estimate is unbiased. That is, \( E(\tilde{x}_k^+) = E(x_k) \).
2) Setting $W = (P_k^+)^{-1}$ results in the minimum variance filter. That is, if $W = (P_k^+)^{-1}$ then
\[ \text{Cov}(x_k - \hat{x}_k^+) \leq \text{Cov}(x_k - \hat{x}_k^-) \] for all $\hat{x}_k^+$.

3) Setting $W = I$ results in a constrained estimate that is closer to the true state than the unconstrained estimate at every time step. That is, if $W = I$ then $||x_k - \tilde{x}_k^+||_2 \leq ||x_k - \hat{x}_k^+||_2$ for all $k$.


In [10] this result (with $W = (P_k^+)^{-1}$) is obtained in a different form along with some additional properties and generalizations. References [8], [18] assume that the *a priori* estimate is based on the unconstrained estimate so that the constrained filter is

\[
\begin{align*}
\hat{x}_k^- &= F\hat{x}_{k-1}, \\
\hat{x}_k^+ &= \hat{x}_k^- + K_k(y_k - H\hat{x}_k^-), \\
\tilde{x}_k^+ &= \hat{x}_k^+ - P_k^+ D^T (DP_k^+ D^T)^{-1}(D\hat{x}_k^+ - d). 
\end{align*}
\tag{11}
\]

Reference [10] bases the *a priori* estimate on the constrained estimate so that the constrained filter is

\[
\begin{align*}
\hat{x}_k^- &= F\tilde{x}_{k-1}, \\
\hat{x}_k^+ &= \tilde{x}_k^- + K_k(y_k - H\tilde{x}_k^-), \\
\tilde{x}_k^+ &= \hat{x}_k^+ - P_k^+ D^T (DP_k^+ D^T)^{-1}(D\hat{x}_k^+ - d). 
\end{align*}
\tag{12}
\]

It can be inductively shown that these two formulations are the same as long as the initial state estimate $\hat{x}_0^+ = \tilde{x}_0^+$. It can also be shown that these constrained estimates are the same as those obtained with the perfect measurement approach [10], [20], [21].
Extension to inequality constraints

The projection approach to constrained filtering has the advantage that it can be easily extended to inequality constraints. If we have the constraints \( Dx \leq d \), then the constrained estimate can be obtained by modifying (9) and solving the problem

\[
\hat{x}_k^+ = \arg\min_x (x - \hat{x}_k^+)^T W (x - \hat{x}_k^+) \text{ such that } Dx \leq d.
\]  

(13)

This is known as a quadratic programming problem [22], [23]. Several algorithms can solve quadratic programming problems, most of which fall in the category known as active set methods. An active set method uses the fact that it is only those constraints that are active at the solution of the problem that are significant in the optimality conditions. Assume that we have \( s \) inequality constraints, and \( q \) of the \( s \) inequality constraints are active at the solution of (13). Denote by \( \hat{D} \) and \( \hat{d} \) the \( q \) rows of \( D \) and \( q \) elements of \( d \) corresponding to the active constraints. If the correct set of active constraints were known a priori then the solution of (13) would also be a solution of the equality constrained problem

\[
\bar{x}_k^+ = \arg\min_x (x - \bar{x}_k^+)^T W (x - \bar{x}_k^+) \text{ such that } \hat{D}x = \hat{d}.
\]  

(14)

The inequality constrained problem of (13) is equivalent to the equality constrained problem of (14). Therefore all of the properties of the equality constrained state estimate enumerated above also apply to the inequality constrained state estimate.

Gain Projection

The standard Kalman filter can be derived by solving the problem [11]

\[
K_k = \arg\min_K \text{Trace} \left[ (I - KH)P_k^- (I - KH)^T + KRK \right].
\]  

(15)
The solution to this problem gives the optimal Kalman gain

\[ S_k = HP_k^{-1}H^T + R, \]
\[ K_k = P_k^{-1}H^T S_k^{-1}, \] (16)

and the state estimate measurement update is

\[ r_k = y_k - H\hat{x}_k, \]
\[ \hat{x}_k^+ = \hat{x}_k^- + K_k r_k. \] (17)

If the constraint \( D\hat{x}_k^+ = d \) is added to the problem, then the minimization problem of (15) can be written as

\[ \tilde{K}_k = \arg\min_{K_k} \text{Trace} \left[ (I - KH)P_k^{-1}(I - KH)^T + KRR \right] \text{ such that } D\hat{x}_k^+ = d. \] (18)

The solution to this constrained problem is [20], [24]

\[ \tilde{K}_k = K_k - D^T (DD^T)^{-1} (D\hat{x}_k^+ - d)(r_k^T S_k^{-1} r_k)^{-1} r_k^T S_k^{-1}. \] (19)

When this value for \( \tilde{K}_k \) is used in place of \( K_k \) in (17), the result is the constrained state estimate

\[ \bar{x}_k^+ = \hat{x}_k^+ - D^T (DD^T)^{-1} (D\hat{x}_k^+ - d). \] (20)

This is the same as the estimate projection given in (10) with \( W = I \).

**Probability Density Function Truncation**

In the PDF truncation approach, we take the PDF of the state estimate that is computed by the Kalman filter (assuming that it is Gaussian) and truncate it at the constraint edges. The constrained state estimate then becomes equal to the mean of the truncated PDF [11], [25], [26]. This approach is designed for inequality constraints on the state although it can also be applied...
to equality constraints with a simple modification. See [11, p. 222] for a graphical illustration of how this works.

This method becomes complicated when the state dimension is more than one. In that case the state estimate is normalized so that its elements are statistically independent of each other. Then the normalized constraints are applied one at a time. After all the constraints have been applied, the normalized state estimate is un-normalized to obtain the constrained state estimate. Details of the algorithm are given in [11], [26].

Soft Constraints

Soft constraints (as opposed to hard constraints) are constraints that are required to be only approximately satisfied rather than exactly satisfied. We might want to implement soft constraints in cases where the constraints are heuristic rather than rigorous, or in cases where the constraint function has some uncertainty or fuzziness. For example, suppose we have a vehicle navigation system with two states $x(1)$ (north position) and $x(2)$ (east position). We know that the vehicle is on a straight road such that $x(1) = mx(2) + b$ for known constants $m$ and $b$. But the road also has a nonzero width, so the state constraint can be written as $x(1) \approx mx(2) + b$. Furthermore, we do not know exactly how wide the road is. In this case we have an approximate equality constraint, which is referred to in the literature as a soft constraint. It can easily be argued that estimators for most practical engineering systems should be implemented with soft constraints rather than hard constraints.

Soft constraints can be implemented in Kalman filters in various ways. First, the perfect measurement approach can be extended to inequality constraints by adding small nonzero
measurement noise to the “perfect” measurements [4], [5], [27], [28]. Second, soft constraints can be implemented by adding a regularization term to the standard Kalman filter [9]. Third, soft constraints can be enforced by projecting the unconstrained estimates in the direction of the constraints rather than exactly onto the constraint surface [29].

**System Projection**

State constraints imply that there are constraints not only on the states, but also on the process noise. This leads to a modification of the initial estimation error covariance and the process noise covariance, after which the standard Kalman filter equations are implemented [30]. Given the constrained system

\[ x_{k+1} = Fx_k + w_k, \]
\[ Dx_k = d, \]

it is reasonable to suppose that the noise-free system also satisfies the constraints. That is, \( DFx_k = 0 \). But this also means that \( Dw_k = 0 \). (If these equations are not satisfied, then the noise \( w_k \) must be correlated with the state \( x_k \), which violates typical assumptions on the system characteristics.) If \( Dw_k = 0 \) then

\[ Dw_k w_k^T D^T = 0, \]
\[ E(Dw_k w_k^T D^T) = 0, \]
\[ DQD^T = 0. \] (22)
This means that $Q$ must be singular (assuming that $D$ is full rank). As a simple example consider the three-state system given at the beginning of this paper. From (4) we have

$$x_{1,k+1} + x_{3,k+1} = 5x_{1k} + 5x_{3k} + w_{1k} + w_{3k}.$$  \hfill (23)

But this means, from (5), that

$$w_{1k} + w_{3k} = 0.$$  \hfill (24)

So the covariance matrix $Q$ must be singular for this constrained system to be consistent. We must have $Dw_k = 0$, which in turn implies (22).

If the given process noise covariance $Q$ does not satisfy (22) then it should be projected onto a modified covariance $\tilde{Q}$ that does satisfy the constraint to make the system consistent. $\tilde{Q}$ then replaces $Q$ in the Kalman filter. This can be accomplished as follows [30].

1) Find the singular value decomposition of $D^T$.

$$D^T = USV^T,$$

$$= \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S_r & 0 \end{bmatrix} V^T,$$ \hfill (25)

where $S_r$ is an $r \times r$ matrix, and $r$ is the rank of $D$.

2) Compute $N = U_2U_2^T$, which is the orthogonal projector onto the null space of $D$.

3) Compute

$$\tilde{Q} = NQN.$$ \hfill (26)

This ensures that

$$D\tilde{Q}D^T = (VS^T U^T)(U_2U_2^T Q U_2U_2^T)(USV^T),$$

$$= 0,$$ \hfill (27)
and thus (22) is satisfied.

Similarly the initial estimation error covariance should be modified as

\[ \tilde{P}_0^+ = NP_0^+ N. \] (28)

It is shown in [30] that the estimation error covariance obtained by this method is less than or equal to that obtained by the estimate projection method. This is because \( \tilde{Q} \) is assumed to be the true process noise covariance, so the system projection method gives the optimal state estimate (just as the standard Kalman filter gives the optimal state estimate for an unconstrained system).

But the standard Kalman filter, and projection methods based on it, use an incorrect covariance \( Q \). If the given \( Q \) satisfies \( DQD^T = 0 \) then the standard Kalman filter estimate satisfies the state constraint \( D\hat{x}_k^+ = 0 \), and the system projection filter, the estimate projection filter, and the standard Kalman filter are all identical.

**Example 1**

Consider a navigation problem. The first two state elements are the north and east positions of a land vehicle, and the last two elements are the north and east velocities. The velocity of the vehicle is in the direction of \( \theta \), an angle measured clockwise from due east. A position-measuring device provides a noisy measurement of the vehicle’s north and east positions.

Equations for this system can be written as

\[
x_{k+1} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ T \sin \theta \\ T \cos \theta \end{bmatrix} u_k + w_k,
\]
\[ y_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} x_k + v_k, \]  

where \( T \) is the discretization step size and \( u_k \) is an acceleration input. The covariances of the process and measurement noise are \( Q = \text{diag} \left( 4, 4, 1, 1 \right) \) and \( R = \text{diag} \left( 900, 900 \right) \). The initial estimation error covariance is \( P_0^+ = \text{diag} \left( 900, 900, 4, 4 \right) \). If we know that the vehicle is on a road with a heading of \( \theta \) then we have

\[
\tan \theta = \frac{x(1)}{x(2)},
\]

\[
= \frac{x(3)}{x(4)}. \quad (30)
\]

We can write these constraints in the form \( D_i x_k = 0 \) using one of two \( D_i \) matrices.

\[
D_1 = \begin{bmatrix} 1 - \tan \theta & 0 & 0 \\ 0 & 0 & 1 - \tan \theta \end{bmatrix},
\]

\[
D_2 = \begin{bmatrix} 0 & 0 & 1 - \tan \theta \end{bmatrix}.
\quad (31)
\]

\( D_1 \) directly constrains both velocity and position. \( D_2 \) relies on the fact that velocity determines position, so when velocity is constrained then position is implicitly constrained. Note that we cannot use \( D = \begin{bmatrix} 1 - \tan \theta & 0 & 0 \end{bmatrix} \). If we did then position would be constrained but velocity would not be constrained. But it is velocity that determines position through the system equations. So this value of \( D \) would not be consistent with the state equations. In particular it violates the \( DF = D \) condition of [10].

At this point we can take several approaches to state estimation.

1) Using the given \( Q \) and \( P_0^+ \),
a) Run the standard unconstrained Kalman filter and ignore the constraints.

b) Run the perfect measurement filter.

c) Project the unconstrained estimate onto the constraint surface.

d) Use moving horizon estimation (MHE) with the constraints. (MHE will be covered later in this paper since it is a general nonlinear estimator.)

e) Use the PDF truncation method.

2) Using the projected $\tilde{Q}$ and $\tilde{P}_0^+$, run the standard Kalman filter. Since $\tilde{Q}$ and $\tilde{P}_0^+$ are consistent with the constraints, the state estimate satisfies the constraints (as long as the initial estimate $\tilde{x}_{0^+}$ satisfies the constraints). This is the system projection approach. Note that neither the perfect measurement filter, the estimate projection filter, the MHE, nor the PDF truncation filter, will change the estimate in this case, since the unconstrained estimate has been implicitly constrained by means of system projection.

In addition, with any of the constrained filters we can use either the $D_1$ or $D_2$ matrix of (31) to constrain the system.

We ran these filters on a 150 s simulation with a 3 s simulation step size. Table I shows the RMS state estimation errors (averaged for the two position states), and the RMS constraint error. Each RMS value shown is averaged over 100 Monte Carlo simulations.

Table I shows that all of the constrained filters have constraint errors that are exactly zero. All of the constrained filters perform identically when $D_1$ is used as the constraint matrix. However, when $D_2$ is used as the constraint matrix, then the perfect measurement and system projection methods perform the best.
Filter results for the linear vehicle navigation problem. The two numbers in each cell indicate the errors that were obtained using the $D_1$ and $D_2$ constraints respectively. The numbers shown are RMS errors averaged over 100 Monte Carlo simulations.

### Nonlinear Constraints

Sometimes state constraints are nonlinear. Instead of $Dx_k = d$ we have

\[
g(x_k) = h. \tag{32}\]

We can perform a Taylor series expansion of the constraint equation around $\hat{x}_k$ to obtain

\[
g(x_k) \approx g(\hat{x}_k) + g'(\hat{x}_k)(x_k - \hat{x}_k) + \frac{1}{2} \sum_{i=1}^{s} e_i(x_k - \hat{x}_k)^T g''_i(\hat{x}_k)(x_k - \hat{x}_k), \tag{33}\]
where $s$ is the dimension of $g(x)$, $e_i$ is the $i$th natural basis vector in $\mathbb{R}^s$, and the element in the $p$th row and $q$th column of the $n \times n$ matrix $g''_i(x)$ is given as

$$ [g''_i(x)]_{pq} = \frac{\partial^2 g_i(x)}{\partial x_p \partial x_q}. \quad (34) $$

Neglecting the second order term gives [3], [4], [8], [13]

$$ g'(\hat{x}_k)x_k = h - g(\hat{x}_k) + g'(\hat{x}_k)\hat{x}_k. \quad (35) $$

This is equal to the linear constraint $Dx_k = d$ if

$$ D = g'(\hat{x}_k), $$

$$ d = h - g(\hat{x}_k) + g'(\hat{x}_k)\hat{x}_k. \quad (36) $$

So all of the methods presented in the earlier sections of this paper can be used with nonlinear constraints after the constraints are linearized. Sometimes, though, we can do better than simple linearization, as discussed in the following sections.

**Second Order Expansion**

If we keep the second order term in (33) then the constrained estimation problem can be approximately written as

$$ \bar{x}_k^+ = \arg\min_x (x - \hat{x}_k^+)^TW(x - \hat{x}_k^+) \text{ such that } $$

$$ x^TM_i x + 2m_i^T x + \mu_i = 0 \quad (i = 1, \cdots, s), \quad (37) $$
where $W$ is a weighting matrix, and $M_i$, $m_i$, and $\mu_i$ are obtained from (33) as

$$M_i = \frac{g_i''(\hat{x}_k^-)}{2},$$

$$m_i = \frac{(g'_i(\hat{x}_k^-) - (\hat{x}_k^-)^T g''_i(\hat{x}_k^-))^T}{2},$$

$$\mu_i = g_i(\hat{x}_k^-) - g'_i(\hat{x}_k^-)\hat{x}_k^- + (\hat{x}_k^-)^T M_i \hat{x}_k^- - h_i.$$

(38)

This idea is similar to the way that the extended Kalman filter (EKF), which relies on linearization of the system and measurement equations, can be improved by retaining second order terms to obtain the second order EKF [11]. The optimization problem given in (37) can be solved with a numerical method. A Lagrange multiplier method for solving this problem is given in [31] for $s = 1$ and $M$ positive definite.

The smoothly constrained Kalman filter

Another approach to handling nonlinear equality constraints is the smoothly constrained Kalman filter (SCKF) [14]. This approach starts with the idea that nonlinear constraints can be handled by linearizing them and then implementing them as perfect measurements. However, the resulting estimate only approximately satisfies the nonlinear constraint. If the constraint linearization is instead applied multiple times at each measurement time then the resulting estimate should get closer to constraint satisfaction with each iteration. This is similar to the iterated Kalman filter for unconstrained estimation [11]. In the iterated Kalman filter the nonlinear system is repeatedly linearized at each measurement time; in the SCKF the nonlinear constraints are repeatedly linearized at each measurement time and then applied as measurements with increasing degrees of certainty. This idea is motivated by realizing that incorporating a measurement with a variance of 1 is equivalent to incorporating that same measurement 10
times, each with a variance of 10. Application of a scalar nonlinear constraint \( g(x) = h \) by means of the SCKF can be performed by the following algorithm, which is executed after each measurement update.

1) Initialize \( i \), the number of constraint applications, to 1. Initialize \( \hat{x} \) to \( \hat{x}_k^- \), and \( P \) to \( P_k^- \).

2) Set \( R'_0 = \alpha GP G^T \), where the \( 1 \times n \) Jacobian \( G = g'(\hat{x}) \). This is the variance with which the constraint is incorporated into the state estimate as a measurement. Note that \( GP G^T \) is the approximate (linearized) variance of \( g(x) \), so \( R'_0 \) is the fraction of this variance that is used to incorporate the constraint as a measurement. \( \alpha \) is a tuning parameter, typically between 0.01 and 0.1.

3) Set \( R'_i = R'_0 \exp(-i) \). This equation is used to gradually (with each iteration) decrease the measurement variance that is used to apply the constraint.

4) Set \( S_i = \max_j(G_j P_{jj} G_j)/(GP G^T) \). This is a normalized version of the information that is associated with the constraint. When this exceeds a given threshold \( S_{\text{max}} \) then the iteration is terminated. A typical value of \( S_{\text{max}} \) is 100. The iteration can also be terminated after a predetermined number of constraint applications \( i_{\text{max}} \) (since there is not yet a convergence proof for the SCKF). When the iteration terminates, set \( \hat{x}_k^+ = \hat{x} \) and \( P_k^+ = P \).

5) Incorporate the constraint as a measurement using

\[
K = PG^T(GPG^T + R'_i)^{-1},
\]
\[
\hat{x} = \hat{x} + K(h - g(\hat{x})),
\]
\[
P = P(I - KG).
\]

These are the standard Kalman filter equations for a measurement update, but the measurement that we are incorporating is the “not-quite-perfect” measurement of the
constraint.

6) Compute the updated Jacobian $G = g'(\hat{x})$. Increment $i$ by one and go to step 3 to continue the iteration.

The above loop needs to be executed once for each inequality constraint.

**Moving Horizon Estimation**

MHE, first suggested in [32], is based on the fact that the Kalman filter solves the following optimization problem [33]–[35].

$$\{\hat{x}_k^+\} = \arg\min_{\{x_k\}} ||x_0 - \hat{x}_0||^2_{I_0^+} + \sum_{k=1}^N ||y_k - H x_k||^2_{R^{-1}} + \sum_{k=0}^{N-1} ||x_{k+1} - F x_k||^2_{Q^{-1}},$$  

(40)

where $\{\hat{x}_k^+\}$ is the sequence of estimates $\hat{x}_0^+, \cdots, \hat{x}_N^+$, and $I_0^+ = \left(P_0^+\right)^{-1}$. This is a quadratic programming problem. The $\{\hat{x}_k^+\}$ sequence that solves this problem gives the optimal smoothed estimate of the state given the measurements $\{y_1, \cdots, y_N\}$.

This motivates a similar method for general nonlinear constrained estimation. Given

$$x_{k+1} = f(x_k) + w_k,$$

$$y_k = h(x_k) + v_k,$$

$$g(x) = 0,$$

(41)

solve the following optimization problem [33], [34], [36].

$$\min_{\{x_k\}} ||x_0 - \hat{x}_0||^2_{I_0^+} + \sum_{k=1}^N ||y_k - h(x_k)||^2_{R^{-1}} + \sum_{k=0}^{N-1} ||x_{k+1} - f(x_k)||^2_{Q^{-1}} \text{ such that } g(\{x_k\}) = 0,$$

(42)

where by an abuse of notation we use $g(\{x_k\})$ to mean $g(x_k)$ for $k = 1, \cdots, N$. This constrained nonlinear optimization problem can be solved by a variety of methods [22], [37], [38], therefore
all of the theory that applies to the particular optimization algorithm that is used also applies to the constrained state estimation problem. The difficulty is the fact that the dimension of the problem increases with time. With each measurement that is obtained, the number of independent variables increases by $n$ (where $n$ is the number of state variables). MHE therefore limits the time span of the problem to decrease the computational effort. The MHE problem can be written as

$$\min_{\{x_k\}} \|x_M - \hat{x}_M^+\|_I^2 + \sum_{k=M+1}^{N} \|y_k - h(x_k)\|_{R^{-1}}^2 + \sum_{k=M}^{N-1} \|x_{k+1} - f(x_k)\|_{Q^{-1}}^2$$

such that $g(\{x_k\}) = 0$, \hspace{1cm} (43)

where $N - M + 1$ is the horizon size. The dimension of this problem is $n(N - M + 1)$. The horizon size is chosen to give a tradeoff between estimation accuracy and computational effort.

$I_M^+ = \left(P_M^+\right)^{-1}$, and $P_M^+$ (which is an approximation of the covariance of the estimation error of $\hat{x}_M^+$) is obtained from the standard Kalman filter recursion [11].

\begin{align*}
F_{k-1} & = \left. \frac{\partial f}{\partial x}\right|_{\hat{x}_{k-1}} , \\
H_k & = \left. \frac{\partial h}{\partial x}\right|_{\hat{x}_k}, \\
P_k^- & = F_{k-1}P_{k-1}^+F_{k-1}^T + Q, \\
K_k & = P_k^-H_k^T(H_kP_k^-H_k^T + R)^{-1}, \\
P_k^+ & = (I - K_kH_k)P_k^- .
\end{align*}

(44)

Some stability results related to MHE are given in [39]. MHE is attractive in the generality of its formulation, but this generality results in large computational effort (even for small horizons) compared to the other constrained filters discussed in this paper.

Another difficulty with MHE is its assumption of an invertible $P_0^+$ in (40) and (42), and
an invertible $P_M^+$ in (43). We saw in (28) that the estimation error covariance for a constrained system is singular. The MHE therefore cannot use the true estimation error covariance; instead it uses the covariance of the unconstrained filter as shown in (44), which makes it suboptimal even for linear systems. Nevertheless the fact that MHE does not make any linear approximations (except implicitly in the nonlinear optimizer) makes it a powerful estimator.

Recursive nonlinear dynamic data reconciliation and combined predictor-corrector optimization [1] are other approaches to constrained state estimation that are similar to MHE. These methods are essentially MHE with a horizon size of one. However the ultimate goal of these methods is data reconciliation (that is, output estimation) rather than state estimation, and they also include parameter estimation.

**Unscented Kalman filtering**

The unscented Kalman filter (UKF) is a filter for nonlinear systems that is based on two fundamental principles [11], [40]. First, although it is difficult to perform a nonlinear transformation of a PDF, it is easy to perform a nonlinear transformation of a vector. Second, it is not difficult to find a set of vectors in state space whose sample PDF approximates any given PDF. The UKF operates by producing a set of vectors called sigma points. The UKF uses between $n + 1$ and $2n + 1$ sigma points, where $n$ is the dimension of the state. The sigma points are transformed and combined in a special way in order to obtain an estimate of the state and an estimate of the covariance of the state estimation error.

Constraints can be incorporated into the unscented Kalman filter (UKF) by treating the constraints as perfect measurements. This can be done in various ways.
1) One possibility is to base the *a priori* state estimate on the unconstrained UKF *a posteriori* state estimate from the previous time step [10], [41]. In this case the standard (unconstrained) UKF runs independently of the constrained UKF. At each measurement time the state estimate of the unconstrained UKF is combined with the constraints (treated as perfect measurements) to obtain a constrained *a posteriori* UKF estimate. This is referred to as the projected UKF (PUKF) and is analogous to (11) for linear systems and constraints. Note that nonlinear constraints can be incorporated as perfect measurements in a variety of ways (e.g., linearization, second order expansion [31], unscented transformation [2], or the SCKF, which is an open research problem).

2) Another approach is to base the *a priori* state estimate on the constrained UKF *a posteriori* state estimate from the previous time step [10]. At each measurement time the state estimate of the unconstrained UKF is combined with the constraints (treated as perfect measurements) to obtain a constrained *a posteriori* UKF estimate. This constrained *a posteriori* estimate is then used as the initial condition for the next time update. This is referred to as the equality-constrained UKF (ECUKF) and is also identical to the measurement-augmentation UKF in [10]. This is analogous to (12) for linear systems and constraints. A similar filter was explored in [2] where it was argued that the covariance of the constrained estimate should be *larger* than that of the unconstrained estimate (since the unconstrained estimate approximates the minimum variance estimate).

3) The two-step UKF (2UKF) [2] projects each *a posteriori* sigma point onto the constraint surface to obtain constrained sigma points. The state estimate is obtained by taking the weighted mean of the sigma points in the usual way, and the resulting estimate is then projected onto the constraint surface. (Note that the mean of constrained sigma points does
not itself necessarily satisfy a nonlinear constraint.) 2UKF is unique in that the estimation error covariance increases after the constraints are applied. The argument for this is that the unconstrained estimate is the minimum variance estimate, so changing the estimate via constraints should increase the covariance. Furthermore, if the covariance decreases with the application of constraints (e.g., using the algorithms in [8], [41]) then the covariance can easily become singular, which could lead to numerical problems with the matrix square root algorithm of the unscented transformation.

4) Unscented recursive nonlinear dynamic data reconciliation (URNDDR) [42] is similar to 2UKF. URNDDR proceeds as follows.

a) The a posteriori sigma points are projected onto the constraint surface, and their weights are modified based on their distances from the a posteriori state estimate.

b) The modified a posteriori sigma points are passed through the dynamic system in the usual way to obtain the a priori sigma points at the next time step.

c) The next set of a posteriori sigma points are obtained using a nonlinear constrained MHE with a horizon size of 1. This requires the solution of a nonlinear constrained optimization problem for each sigma point.

d) The a posteriori state estimate and covariance are obtained by combining the sigma points in the normal way.

The constraints are thus used in two different ways for the a posteriori estimates and covariances. The URNDDR is called the sigma point interval UKF in [41].

5) The constrained UKF (CUKF) is identical to the standard UKF, except a nonlinear constrained MHE with a horizon size of 1 is used to obtain the a posteriori estimate [41]. Sigma points are not projected onto the constraint surface, and constraint information is
not used to modify covariances.

6) The constrained interval UKF (CIUKF) combines the sigma point constraints of URNDDR with the measurement update of the CUKF [41]. That is, the CIUKF is the same as URNDDR except instead of using MHE to constrain the a posteriori sigma points, the unconstrained sigma points are combined to form an unconstrained estimate, and then MHE is used to constrain the estimate.

7) The interval UKF (IUKF) combines the post-measurement projection step of URNDDR with the measurement update of the standard unconstrained UKF [41]. That is, the IUKF is the same as URNDDR except that it skips the MHE-based constraint of the a posteriori sigma points. Equivalently, IUKF is also the same as CIUKF except that it skips the MHE-based constraint of the a posteriori state estimate.

8) The truncated UKF (TUKF) combines the PDF truncation approach described earlier in this paper with the UKF [41]. After each measurement update of the UKF, the PDF truncation approach is used to generated a constrained state estimate and covariance. The constrained estimate is used as the initial condition for the following time update.

9) The truncated interval UKF (TIUKF) adds the PDF truncation step to the a posteriori update of the IUKF [41]. As with the TUKF, the constrained estimate is used as the initial condition for the following time update.

**Interior point approaches**

A new approach to inequality constrained state estimation called interior point likelihood maximization (IPLM) has been recently proposed [43]. This approach is based on interior point methods, which are fundamentally different from active set methods for constraint enforcement.
Active set methods for inequality constraints, as discussed earlier in this paper, proceed by solving equality-constrained subproblems and then checking if the constraints of the original problem are satisfied. One difficulty with active set methods is that computational effort grows exponentially with the number of constraints. Interior point approaches solve inequality-constrained problems by iterating using a Newton’s method that is applied to a certain subproblem. The approach in [43] relies on linearization. It also has the disadvantage that the problem grows linearly with the number of time steps. However, this difficulty could be potentially be addressed by limiting the horizon size, similar to MHE.

Example 2

This example is taken from [10]. A discretized model of a pendulum can be written as

\[\begin{align*}
\theta_{k+1} &= x_k + T \omega_k, \\
\omega_{k+1} &= \omega_k - (T g/L) \sin \theta_k, \\
y_k &= \begin{bmatrix} \theta_k \\ \omega_k \end{bmatrix} + v_k, \\
\end{align*}\]

(45)

where \(\theta\) is angular position, \(\omega\) is angular velocity, \(T\) is the discretization step size, \(g\) is the acceleration due to gravity, and \(L\) is the pendulum length. By conservation of energy we have

\[-m g L \cos \theta_k + m L^2 \omega_k^2 / 2 = \text{constant}.\]

(46)

This is a nonlinear constraint on the states \(\theta_k\) and \(\omega_k\). We use \(L = 1\), \(T = 0.05\) (trapezoidal integration), \(g = 9.81\), \(m = 1\), and \(x_0 = \begin{bmatrix} \pi/4 \\ \pi/50 \end{bmatrix}^T\). The covariance of the measurement noise is \(R = \text{diag} \left( 0.01, 0.01 \right)\). The initial estimation error covariance is \(P_0^+ = \text{diag} \left( 1, 1 \right)\).

We do not use process noise in the system simulation, but in the Kalman filters we use
\[ Q = \text{diag} \left( 0.007^2, 0.007^2 \right) \] to help with convergence. We compare several approaches to state estimation, including the following.

1) Run the standard unconstrained EKF and ignore the constraints.

2) Linearize the constraint and:
   a) Run the perfect measurement EKF.
   b) Project the unconstrained EKF estimate onto the constraint surface.
   c) Use the PDF truncation method on the EKF estimate.

3) Use a second order expansion of the constraint to project the EKF estimate onto the constraint surface.

4) Use nonlinear MHE with the nonlinear constraint.

5) Use the SCKF.

6) Use the UKF and:
   a) Ignore the constraint.
   b) Project the \textit{a posteriori} estimate onto the constraint surface using a linearized expansion of the constraint.
   c) Use the constrained \textit{a posteriori} estimate to obtain the \textit{a priori} estimate at the next time step.
   d) Use the two-step UKF projection.

Note that the corrected \( \tilde{Q} \) and \( \tilde{P}_0^+ \) (obtained via first order linearization and system projection) could be used with any of the filtering approaches listed above.

Table II shows the RMS state estimation errors (averaged for the two states) and the RMS constraint error. Each RMS value shown is averaged over 100 Monte Carlo simulations.
Table II shows that MHE performs the best relative to estimation error. However this is at a high computational expense. Matlab’s Optimization Toolbox has constrained nonlinear optimization solvers that can be used for MHE, but for this example we used SolvOpt [44], [45]. If computational expense is a consideration then the equality constrained UKF performs the next best. However UKF implementations can also be expensive because of the sigma point calculations that are required. We see that several of the estimators result in constraint errors that are essentially zero. The constraint errors and estimation errors are positively correlated, but small constraint errors do not guarantee that the estimation errors are small.

CONCLUSION

The number of algorithms for constrained state estimation can be overwhelming. The reason that there are so many different algorithms is because the problem can be viewed from so many different perspectives. A linear relationship between states implies a reduction of the state dimension, hence the model reduction approach. State constraints can be viewed as perfect measurements, hence the perfect measurement approach. Constrained Kalman filtering can be viewed as a constrained likelihood maximization problem or a constrained least squares problem, hence the projection approaches. If we start with the unconstrained estimate and then incorporate the constraints to adjust the estimate we get the general projection approach and PDF truncation. If we realize that state constraints affect the relationships between the process noise terms we get the system projection approach.

Nonlinear systems and constraints have all the possibilities of nonlinear estimation, combined with all the possibilities for solving general nonlinear equations. This gives rise to the
<table>
<thead>
<tr>
<th>Filter Type</th>
<th>RMS Estimation Error ($Q$, $\tilde{Q}$)</th>
<th>RMS Constraint Error ($Q$, $\tilde{Q}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>unconstrained</td>
<td>0.0411, 0.0253</td>
<td>0.1167, 0.0417</td>
</tr>
<tr>
<td>perfect measurement</td>
<td>0.0316, 0.0905</td>
<td>0.0660, 0.0658</td>
</tr>
<tr>
<td>estimate projection</td>
<td>0.0288, 0.0207</td>
<td>0.0035, 0.0003</td>
</tr>
<tr>
<td>MHE, horizon size 2</td>
<td>0.0105, 0.0067</td>
<td>0.0033, 0.0008</td>
</tr>
<tr>
<td>MHE, horizon size 4</td>
<td>0.0089, 0.0067</td>
<td>0.0044, 0.0007</td>
</tr>
<tr>
<td>system projection</td>
<td>N/A, 0.0250</td>
<td>N/A, 0.0241</td>
</tr>
<tr>
<td>PDF truncation</td>
<td>0.0288, 0.0207</td>
<td>0.0035, 0.0003</td>
</tr>
<tr>
<td>2nd order constraint</td>
<td>0.0288, 0.0204</td>
<td>0.0001, 0.0000</td>
</tr>
<tr>
<td>SCKF</td>
<td>0.0270, 0.0235</td>
<td>0.0000, 0.0000</td>
</tr>
<tr>
<td>unconstrained UKF</td>
<td>0.0400, 0.0237</td>
<td>0.1147, 0.0377</td>
</tr>
<tr>
<td>projected UKF</td>
<td>0.0280, 0.0192</td>
<td>0.0046, 0.0007</td>
</tr>
<tr>
<td>equality constrained UKF</td>
<td>0.0261, 0.0173</td>
<td>0.0033, 0.0004</td>
</tr>
<tr>
<td>two-step UKF</td>
<td>0.0286, 0.0199</td>
<td>0.0005, 0.0000</td>
</tr>
</tbody>
</table>

**TABLE II**

Filter results for the nonlinear pendulum example. The two numbers in each cell indicate the errors that were obtained when $Q$ and $\tilde{Q}$ respectively were used in the filter. The numbers shown are RMS errors averaged over 100 Monte Carlo simulations.
EKF, the UKF, MHE, and particle filtering for estimation. Then any of these estimators can be combined with various approaches for handling constraints, including first order linearization (which includes the SCKF). If first order linearization is used then any of the approaches discussed above for handling linear constraints can be used. In addition, since there are multiple steps in state estimation (the \textit{a priori} step and the \textit{a posteriori} step), we can use one approach at one step and another approach at another step. The total number of possible constrained estimators seems to grow exponentially with the number of nonlinear estimation approaches and with the number of constraint handling options.

Theoretical and simulation results indicate that all of the constrained filters for linear systems and linear constraints perform identically, if the constraints are complete. So in spite of the many approaches to the problem, we have a pleasingly parsimonious unification. However, if the constraints are not complete, then the perfect measurement and system projection methods performed best in our particular simulation example.

For nonlinear systems and nonlinear constraints, our simulation results indicate that of all the algorithms we investigated, MHE results in the smallest estimation error. However, this is at the expense of programming effort and computational effort that is orders of magnitude higher than other methods. Given this caveat, it is not obvious what the “best” constrained estimation algorithm is, and it generally depends on the application. The possible approaches to constrained state estimation can be delineated by following a flowchart that asks questions about the system type and the constraint type; see “Constrained Kalman Filtering Possibilities.”

Constrained state estimation is well-established but there are still many interesting possibilities for future work, among which are the following.
1) From Example 2 we saw that the linearly constrained filters performed identically for the $D_1$ constraint, but differently for the $D_2$ constraint. Some of these equivalences have already been proven, but conditions under which the various approaches are identical have not yet been completely established.

2) The second order constraint approximation was developed in [31] (and implemented in this paper) in combination with the estimate projection filter. The second order constraint approximation could also be combined with other filters, such as MHE, the UKF, and the SCKF.

3) Algorithms for solving the second order constraint approach should be developed and investigated for the case of multiple constraints.

4) More theoretical results related to convergence and stability are needed for nonlinear constrained filters such as SCKF, MHE, and UKF.

5) MHE should be modified so that it can use the optimal (singular) estimation error covariance (obtained using system projection) in its cost function.

6) Second order Kalman filtering could be combined with MHE to get a more accurate approximation of the estimation error covariance.

7) Various combinations of the approaches discussed in this paper could be explored. For example, PDF truncation could be combined with MHE, the SCKF could be combined with the UKF, and so forth.

8) For the nonlinear system of Example 2 we used a first-order approximation for system projection to obtain $\tilde{Q}$ and $\tilde{P}_0^+$, but a second-order approximation should give better results.

9) Particle filtering is a state estimation approach that is outside the scope of this paper, but it has obvious applications to constrained estimation. There is a lot of room for advancement.
in the theory and implementation of constrained particle filters.

10) Interior point methods for constrained state estimation have just begun to be explored. Further work in this area could include higher order expansion of nonlinear system and constraint equations in interior point methods, moving horizon interior point methods, the use of additional interior point theory and algorithms beyond those used in [43], and generalization of the convergence results given in [43].

The results presented in this paper can be reproduced by downloading Matlab source code from http://academic.csuohio.edu/simond/ConstrKF.
REFERENCES


methods,” submitted for publication.


filters for interval-constrained nonlinear systems,” submitted for publication.


SIDEBAR S1: CONSTRAINED KALMAN FILTERING RESEARCH

A literature search using INSPEC, a science and engineering research database, reveals the recent increase in research related to constrained Kalman filtering. Figure S1 shows the number of published papers with a title containing “Kalman” or some form of the word “filter”, and some form of the word “constrain”. Such a simple search excludes many papers that deal with constrained Kalman filtering, and includes some that do not deal with constrained Kalman filtering, but the general trend shown in Figure S1 is still illuminating.

Figure S1. The number of published papers with a title containing “Kalman” or some form of the word “filter”, and some form of the word “constrain”. The solid line is a polynomial fit to the data showing the projected future growth of such research.

SIDEBAR S2: CONSTRAINED KALMAN FILTERING POSSIBILITIES

Although it is not possible to determine \textit{a priori} the “best” constrained filter for a given problem, Figure S2 summarizes the possible approaches that can be taken for various
combinations of system type and constraint type. The acronyms used in the flowchart are given below, and the reference numbers show where the relevant equations can be found.

2E = second order expansion of nonlinear constraints [31]

2UKF = two-step UKF [2]

CIUKF = constrained IUKF [41]

ECUKF = equality constrained EKF [10]


GP = gain projection [20], [24]

IPLM = interior point likelihood maximization [43]

IUKF = interval UKF [41]

MHE = moving horizon estimation [33], [39]


PDFT = probability density function truncation [11]

PM = perfect measurement [11]

PUKF = projected UKF [41]

SCKF = smoothly constrained Kalman filter [14]

SP = system projection [30]

TIUKF = truncated IUKF [41]

TUKF = truncated UKF [41]

UKF = unscented Kalman filter [11], [40]

URNDDR = unscented recursive nonlinear dynamic data reconciliation [42]

Note that some of the acronyms refer only to filter methods, some refer only to
constraint incorporation methods, and some refer to a combination filter/constraint incorporation algorithm. In addition, sometimes the same acronym (algorithm) can refer to both a filter without constraints, and also a filter/constraint handling combination. For example, MHE can be used as an unconstrained state estimator, and then the MHE estimate can be modified by incorporating constraints using EP; or MHE can be used as a constrained state estimator by incorporating the constraints into the MHE cost function.

**AUTHOR INFORMATION**

Dan Simon is an Associate Professor at Cleveland State University in Cleveland, Ohio. His research areas include robust control, computer intelligence, motor control, and robotics. He worked for 14 years in industry before beginning his career in academia. His industrial experience includes work in the aerospace, automotive, biomedical, GPS, process control, and software engineering fields. He is the author of the book *Optimal State Estimation* (John Wiley & Sons, 2006).
Figure S2. Possible filter and constraint-handling choices for various combinations of system types and constraint types. Note that some of the acronyms refer only to filter options, some refer only to constraint incorporation options, and some refer to a combination filter/constraint incorporation algorithm.