A Parametric Programming Approach to Moving Horizon State Estimation

Mark L. Darby and Michael Nikolaou

Department of Chemical Engineering, University of Houston, Houston, TX 77204-4004

Submitted to Automatica, July 10, 2005

Keywords: Moving horizon, receding horizon, state estimation, parametric programming, Kalman filter, constraints

Abstract

We propose a solution to moving-horizon state estimation that incorporates inequality constraints in both a systematic and computationally efficient way akin to Kalman filtering. The proposed method allows the on-line constrained optimization problem involved in moving-horizon state estimation to be solved off-line, with only simple function evaluations required for real-time implementation. The method is illustrated via simulations on a system that has appeared in moving-horizon state estimation literature.

1 Author to whom all correspondence should be addressed. Nikolaou@uh.edu +1 713 743 4309
1 Introduction

Rapidly determining the values of a system’s states in real time is a necessary element in model-based applications such as control, monitoring, fault detection and diagnosis. When states cannot be measured directly in real time, the state values must be estimated from noisy real-time measurements. For linear systems with Gaussian noise, the celebrated Kalman filter (Kalman, 1960) provides a recursive solution to the real-time minimum-variance estimation problem, given prior knowledge of the distributions of the initial states, disturbances, and measurement noise. The Kalman filter has also been applied to nonlinear systems in the form of the extended Kalman filter (EKF), which is based on linearization of the non-linear model around the current mean and covariance estimate. However, the EKF may exhibit poor convergence properties (Maybeck, 1982; Haseltine & Rawlings, 2005). Following on the success of real-time optimization over a moving horizon used in model-predictive control, moving horizon estimation (MHE) via real-time optimization was suggested as a practical method for addressing model nonlinearities and inequality constraints in state estimation while keeping the size of the real-time optimization problem finite (Muske, Rawlings & Lee, 1993; Robertson, Lee & Rawlings, 1996; Rao, Rawlings & Lee, 2001; Rao & Rawlings, 2002; Simon & Simon, 2003). In addition, it was shown that constrained state estimators have additional useful properties, such as producing unbiased state estimates and smaller state error covariance (Simon & Simon, 2003). Including inequality constraints in MHE provides a mechanism to improve the estimation based on process knowledge (e.g., flow rates and compositions must be greater than or equal to zero), and can also help compensate for poor choices in the prior distributions. While the power of MHE has been clearly demonstrated, the computational requirements of real-time constrained optimization that MHE entails may render it impractical in cases where computing power is limited and/or data sampling rates are too fast. For example, in real-time monitoring and diagnostics for aircraft engines – a case where Kalman filtering is widespread and constraints on estimates are frequently known a priori – data are typically collected at rates over tens of Hz, leaving little time for state estimation via real-time optimization.

To address the computational efficiency issues posed by MHE, we propose in this paper an MHE approach that completely bypasses real-time optimization. At each time point, the proposed algorithm consults a look-up table to select a closed-form expression from a finite collection, and then it uses that expression to calculate the state estimate. Both the look-up table and the collection of closed-form expressions for state estimation are constructed once via off-line optimization. Our approach parallels the multi-parametric programming approach proposed for constrained model predictive control (MPC) of linear systems (Pistikopoulos, Dua, Bozinis, Bemporad & Morari, 2002). In summary, we show that the real-time optimization performed by MHE for linear systems with linear inequality constraints and quadratic objective can be expressed as a multi-parametric quadratic programming (mp-QP) problem, where the parameters appear only in the constraints. Further, we show that there is a finite number of linear polytopes, such that depending on what polytope the parameter values belong, a corresponding closed-form expression can be employed for state estimation. Even though for relatively high-dimensional systems the finite number of linear polytopes is unwieldy, due to
combinatorial complexity, that number is manageable for relatively low-dimensional systems. Thus, the linear polytopes can be characterized once, rather than in real time, via off-line optimization. Subsequently, at each measurement point in time the appropriate polytope can be selected and the corresponding closed-form expression for state estimation can be employed, completely bypassing real-time optimization.

In the rest of the paper, we briefly present elements of Kalman filtering, MHE, and mp-QP that are relevant to this work. Next, we show how mp-QP can be applied to MHE. Finally, we demonstrate the applicability of the proposed approach by presenting simulations on systems that have appeared in literature, and discuss future developments.

2 Background

2.1 System description

Let a dynamic system be described by the following discrete-time model

\[ x_{k+1} = f_k(x_k, u_k) + g_k(x_k, u_k)w_k \] (1)
\[ y_k = h_k(x_k) + v_k \] (2)

where the time point \( k \) takes integer values; the vector \( x_k \in \mathbb{R}^n \) is the state; the vector \( u_k \in \mathbb{R}^n \) is the known input; the vectors \( w_k \in \mathbb{R}^m \) and \( v_k \in \mathbb{R}^m \) are random variables (often assumed to be independent and Gaussian with zero mean) representing the state and measured output disturbances, respectively; \( y_k \in \mathbb{R}^v \) is the measured output vector; is the measured output disturbance; \( f_k(\cdot) \) and \( g_k(\cdot) \) represent the discretized form of an ODE or PDE and may not be available in closed form. It is also assumed that the states and disturbances satisfy the following constraints:

\[ D_x x_k \leq d_x, \quad D_w w_k \leq d_w, \quad D_v v_k \leq d_v \] (3)

where the matrices \( D_x, D_w, D_v \) are known. The origin is assumed to satisfy the constraints on \( w_k \) and \( v_k \). Typically the constraints take the form of upper and lower bounds of the vector components (e.g., \( x_k^{\text{min}} \leq x_k^{\max} \)), where the subscript \( i \) refers to the \( i \)-th component of the vector \( x \). It has been shown (Robertson & Lee, 2002) that constraints also allow one to incorporate non-Gaussian distributions (e.g., asymmetric or truncated distributions) in the state estimation problem.

2.2 Moving horizon state estimation and Kalman filtering

Moving horizon estimation is performed by solving the following optimization problem in real time at each discrete time point \( t \):
\[ \{ \hat{x}_{t-N+1|t}, \ldots, \hat{x}_{t+1|t}, \hat{w}_{t-N+1|t}, \ldots, \hat{w}_{t+1|t}, \hat{v}_{t-N+1|t}, \ldots, \hat{v}_{t+1|t} \} = \]

\[
\arg\min_{\hat{y}_{t-N+1|t}, \ldots, \hat{w}_{t+1|t}} \left\{ \sum_{k=t-N+1}^{t} \hat{v}_{k|t}^T R^{-1} \hat{v}_{k|t} + \sum_{k=t-N+1}^{t} \hat{w}_{k|t}^T Q^{-1} \hat{w}_{k|t} + \right. \\
\left. (\hat{x}_{t-N+1|t} - \bar{x}_{t-N+1|t-N})^T P_{t-N+1}^{-1} (\hat{x}_{t-N+1|t} - \bar{x}_{t-N+1|t-N}) \right\}
\]

subject to

\[
\hat{x}_{k|t} = f_k(\hat{x}_{k|t}, u_k, \hat{w}_{k|t}), \quad k = t-N+1, \ldots, t-1
\]

\[
y_k = h_k(\hat{x}_{k|t}) + \hat{v}_{k|t}, \quad k = t-N+1, \ldots, t
\]

\[
D_x \hat{x}_{k|t} \leq d_x, \quad k = t-N+1, \ldots, t
\]

\[
D_w \hat{w}_{k|t} \leq d_w, \quad k = t-N+1, \ldots, t-1
\]

\[
D_v \hat{v}_{k|t} \leq d_v, \quad k = t-N+1, \ldots, t
\]

where \( \hat{s}_{k|t} \) denotes the estimated value of \( s \) at time point \( k \), given measurements up to and including time \( t \); \( \bar{x}_{t-N+1|t-N} \) is the "best" estimate of the state vector at discrete time \( t - N + 1 \) predicted at time \( t - N \), e.g. as

\[
\bar{x}_{t-N+1|t-N} = E\left[ f_k(\hat{x}_{t-N+1|t-N}, u_{t-N+1}, \hat{w}_{t-N+1}) \right];
\]

and the positive definite matrices \( Q, R, \) and \( P_{t-N+1} \) are the covariances of the state disturbance, output disturbance, and state, respectively. The resulting \( \hat{x}_{t-N+1|t-N} \) is the optimal filtered state estimate, while \( \{ \hat{x}_{t-N+1|t-N} \} \) are the optimal smoothed estimates of past states.

The objective function minimized in eqn. (4) allows error to be distributed between the measurements (two summation terms) and model (third term in eqn. (4)) according to the respective model confidences quantified by the inverses of the covariance matrices \( Q, R, \) and \( P_{t-N+1} \). The third term of the objective function in eqn. (4) represents a penalty on the deviation from the state value previously model-predicted at time \( t - N + 1 \) using moving horizon estimation data up to and including time \( t - N \). The pair \( (\bar{x}_{t-N+1|t-N}, P_{t-N+1}) \) summarizes the model-based influence of old data not explicitly considered in the current finite horizon, i.e. data corresponding to time points before \( t - N + 1 \). This is the estimation problem counterpart of the terminal cost in receding horizon model predictive control (Mayne, Rawlings, Rao & Scokaert, 2000; Rawlings, 2000; Nikolaou, 2001).

A key issue for the moving horizon estimator becomes how to update or propagate \( (\bar{x}_{t-N+1|t-N}, P_{t-N+1}) \) based on previous data. In probabilistic terms (Robertson, Lee & Rawlings, 1996; Rao, Rawlings & Lee, 2001) this is a problem of updating the conditional density function of the state.
For a linear time-invariant system without inequality constraints and with horizon extending from the initial time 1 to current time \( t \), the optimization to solve for state estimation at \( t \) is

\[
\{ \hat{x}_{1|1}, \hat{x}_{2|2}, \ldots, \hat{x}_{t|t}, \hat{w}_{1|1}, \hat{w}_{2|2}, \ldots, \hat{w}_{t|t}, \hat{\nu}_{1|1}, \hat{\nu}_{2|2}, \ldots, \hat{\nu}_{t|t} \} =
\arg\min_{\hat{x}_{t|t}, \ldots, \hat{x}_{1|1}} \left\{ \sum_{k=1}^{t} \hat{w}_{k|k} R^{-1} \hat{w}_{k|k}^T + \sum_{k=1}^{t-1} \hat{\nu}_{k|k} Q^{-1} \hat{\nu}_{k|k}^T + \left( \hat{x}_{t|t} - \bar{x}_{t|t} \right)^T P_t^T \left( \hat{x}_{t|t} - \bar{x}_{t|t} \right) \right\}
\]

subject to

\[
\hat{x}_{k|k} = A \hat{x}_{k|k} + Bu_k + G \hat{w}_{k|k}, \quad k = 1, \ldots, t-1
\]

\[
y_k = C \hat{x}_{k|k} + \hat{\nu}_{k|k}, \quad k = 1, \ldots, t
\]

where \( P_t \) is the covariance matrix of the initial state estimate \( \hat{x}_{1|1} \). In that case, it can be shown (Sorenson, 1970) that the solution of the optimization problem of eqns. (9) through (11) can be expressed recursively via the Kalman filter equations, as

\[
\hat{x}_{k+1|k+1} = A \hat{x}_{k|k} + Bu_k + GE \left[ \hat{w}_{k|k} \right] + K_k \left( y_k - C \hat{x}_{k|k} \right), \quad \hat{x}_{1|1} = \bar{x}_{1|1}
\]

\[
K_k = A P_k C^T \left( CP_k C^T + R \right)^{-1}
\]

\[
P_{k+1} = GQG^T + \left[ AP_k A^T \right] \left( CP_k C^T + R \right)^{-1} CP_k A^T
\]

where \( P_t \) is the initial covariance of the state, as in eqn. (9). If no prior knowledge about the state is available, then \( P_t = \infty \) and the last term in eqn. (9) is omitted. Note that if the state disturbance \( w \) has non-zero mean, an additional forcing term is present in eqn. (12).

For a nonlinear system or linear system with constraints it is generally impossible to derive analytical expressions for the conditional density. As a result, the Kalman filter is typically used to approximate \( P_{t-N+1} \) and \( \bar{x}_{t-N+1|t-N} \) in eqn. (4) via eqns. (14) and

\[
\bar{x}_{t-N+1|t-N} = A \hat{x}_{t-N|t-N} + Bu_{t-N} + G_t E \left[ w_{t-N|t-N} \right]
\]

respectively (p. 318, Stengel, 1994).

Note also that the steady-state solution of eqn. (14) may also be used as an approximation for \( P_{t-N+1} \). We will make use of this in the next section, where we will assume a constant covariance matrix, \( P \). Ideally this corresponds to the steady-state Kalman filter covariance matrix. This choice is justified based on the results of (Rao, Rawlings & Lee, 1999), who show that selecting the steady-state Kalman filter covariance ensures that the linearly constrained moving horizon estimator is an asymptotically stable observer of the actual system, even in the presence of constraints. Their results assume a setting with a deterministic linear model (no state or output noise) subject to linear inequality constraints on \( x_k \). When modeling error is present, a value
larger than the steady-state Kalman filter covariance $P$ may be required to ensure robust stability.

The summations in the objective function in eqn. (4) assume a complete history of measurements starting at $t - N + 1$, tacitly assuming that $t \geq N$. For the start-up phase of implementation of the algorithm, i.e. $t < N$, one can execute smaller sized batch problems similar to eqn. (4) and allow the number of measurements to increase until $t = N$. At $t = N + 1$, $\hat{x}_{10}$ is replaced by $\hat{x}_{20}$, $y_{N+1}$ is brought in and $y_1$ discarded (Robertson, Lee & Rawlings, 1996).

A potential problem with using the Kalman filter approximation in moving horizon estimation is that a poor initial state covariance matrix, $P_1$ may lead to instability as a result of overweighting past data relative to newer data. Convergence and stability issues have been analyzed (Rao, Rawlings & Lee, 2001) via forward dynamic programming and the concept of arrival cost (the analog of cost to go in forward dynamic programming). It turns out that stability considerations place an upper bound on the arrival cost, which in turn places a lower bound on the initial covariance matrix, $P_1$.

### 2.3 Multiparametric Programming

Consider the following quadratic program (QP)

$$\min_M \left\{ \frac{1}{2} M^T H M + X^T F M \right\}$$

$$DM \leq T + EX$$

(16)

where $M$ is an unknown vector of optimization (decision) variables; $X$ is a known parameter vector belonging to a set $X$; and the matrices $H = H^T > 0$, $F$, $D$, $T$, and $E$ are constant. If we complete the squares and define the transformation

$$Z = M + H^{-1} F^T X,$$

(17)

then the QP of eqn. (16) is equivalent to the following multi-parametric quadratic program (mp-QP):

$$V_c(X) = \min_Z \frac{1}{2} Z^T H Z$$

$$DZ \leq T + SX$$

(18)

where $S = E + DH^{-1} F^T$. Notice that the parameter vector $X$, which appears in both the objective function and constraints in eqn. (16), only appears in the right-hand side of the constraints in eqn. (18).

The advantage of writing the QP in the form of eqn. (18) is that it allows the optimal $Z$ (and therefore the optimal $M$, via eqn. (17)) to be explicitly expressed in terms of an affine function of $X$ as

$$Z_{opt}(X) = H^{-1} D_A^T (D_A H^{-1} D_A^T)^{-1} (T_A + S_A X)$$

(19)
where the matrices $D_A$, $T_A$, and $S_A$ take values from a finite collection of values (generated by the matrices $H$, $F$, $D$, $T$, and $E$ in eqn. (16)) according to what set of constraints in eqn. (18) are active (hence the subscript $A$).

What set of constraints are active is determined by the value of the parameter vector $X$ in a direct way, as follows: Each set of active constraints generates a corresponding critical region $CR \subset X$ for the parameter vector $X$, defined as the linear polytope

$$CR = \left\{ X \mid G_i Z_{opt}(X) \leq T_i + S_i X, \lambda_{i}(X) \geq 0 \right\}$$

(20)

where the matrices $G_i$, $T_i$, and $S_i$ are determined from $H$, $F$, $D$, $T$, and $E$ in eqn. (16) according to the corresponding set of inactive constraints; $Z_{opt}(X)$ is given by eqn. (19); and

$$\lambda_{i}(X) \equiv -(D_A H^{-1} D_A^T)^{-1}(T_A + S_A X)$$

(21)

(see Appendix A for completeness). Therefore, depending on the linear polytope, $CR \subset X$, to which the parameter vector $X$ belongs, a set of corresponding constraints will be active, and corresponding values for the matrices $D_A$, $T_A$, $S_A$ will have to be used to express $Z_{opt}(X)$ in terms of eqn. (19). Consequently, if the entire collection of polytopes $CR$ have already been characterized (in terms of a collection of linear inequality constraint sets), then for repeated solution of the optimization problem in eqn. (16) as $X$ varies in $X$, one can

(a) check the two inequalities in eqn. (20) to identify the polytope $CR$ to which the parameter vector $X$ belongs; and

(b) for the polytope identified in the above step (a) use the corresponding matrices $D_A$, $T_A$, $S_A$ to compute $Z_{opt}(X)$ via eqn. (19).

To characterize the collection of linear polytopes $CR$, one can proceed as follows (Bemporad, Morari, Dua & Pistikopoulos, 2002; Pistikopoulos, Dua, Nikolaos, Bemporad & Morari, 2002):

(a) For a given value of the parameter vector $X$ in $X$, determine the matrices $G_i$, $T_i$, $S_i$, $D_A$, $T_A$, and $S_A$, and characterize the (convex) linear polytope, $CR \subset X$, in terms of the linear inequality constraints in eqn. (20).

(b) Reverse one of the constraints in the set of linear inequality constraints of eqn. (20), to construct a new set of linear inequalities.

(c) Solve a feasibility linear programming (LP) problem to find a new feasible vector $X$ in $X$ that satisfies the inequalities constructed in the above step (b). If no feasible vector $X$ in $X$ exists, redo step (b).

(d) Go to step (a), unless the process has reached its guaranteed termination in a finite number of steps and all linear polytopes $CR \subset X$ have been characterized.

The result of the above procedure is a partitioning of the set $X$ into a collection of linear polytopes $CR \subset X$, in each of which the optimal solution $M_{opt}(X)$ of eqn. (16) is
provided via the same formula (eqn. (19) and substitution into eqn. (17)) for all \( X \) in that polytope. Thus, significant computational savings may be realized in cases where a series of QP problems, eqn. (16), have to be solved for different values of \( X \), provided that the collection of all linear polytopes \( CR \) for the partitioning of \( X \) is not prohibitively large, so that it can be computed once and stored for later use.

The number of polytopes required to partition the set \( X \) can be determined as follows. If the maximum number of constraints that can be active for one execution of the QP is denoted by \( q \), then the possible number of active constraints, \( m \), can vary from 0 to \( q \). The number of possible combinations of \( m \) constraints out of \( q \) is:

\[
\binom{q}{m} = \frac{q!}{(q-m)!m!}
\]  

(22)

Therefore the maximum number of polytopes is

\[
\sum_{m=0}^{q} \binom{q}{m} = 2^q
\]  

(23)

3 Constrained Moving Horizon Estimation as Multiparametric Optimization

In this section we are using the background presented in section 2 to develop our main result, namely formulate the MHE problem for a linear system with quadratic objective and linear inequality constraints as an mp-QP problem.

The linear system is modeled by eqns. (10) (with \( k = t - N + 1, \ldots, t - 1 \)) and (11) (with \( k = t - N + 1, \ldots, t \)). The quadratic objective is eqn. (4) with \( \bar{x}_{t-N+1|t-N} \) given by eqn. (15). For the covariance matrix \( P_{t-N+1} \) we use the steady-state approximation given by eqn. (14).

To covert the above MHE problem to the mp-QP form, we substitute the model equations, eqns. (10) and (11) into the MHE objective function, eqn. (4), so that only the inequality constraints, eqn. (7), related to the random variables \( x_k, w_k, \) and \( v_k \) remain. At time \( t \), the solution to the above MHE problem depends on the past measured output sequence \( y_{t-N+1}, \ldots, y_t \), the past input sequence \( u_{t-N+1}, \ldots, u_{t-1} \), and the previously calculated \( \bar{x}_{t-N+1|t-N} \). With the definitions

\[
M \triangleq \begin{bmatrix} \hat{x}_{t-N+1|t}^T & \hat{w}_{t-N+1|t}^T & \cdots & \hat{w}_{t-N+1|t}^T \end{bmatrix}^T
\]

\[
X \triangleq \begin{bmatrix} \bar{x}_{t-N+1|t-N}^T & \bar{y}_{t-N+1|t-N}^T \end{bmatrix}^T
\]

\[
Y \triangleq \begin{bmatrix} y_{t-N+1}^T & \cdots & y_t^T \end{bmatrix}^T
\]

\[
U \triangleq \begin{bmatrix} u_{t-N+1}^T & \cdots & u_{t-1}^T \end{bmatrix}^T
\]

(24)
we can express the above MHE problem in the mp-QP form of eqn. (16). The corresponding matrices $H$ and $F$ are derived in Appendix B.

In the above, we have implicitly assumed that the upper and lower bounds for the inequality constraints in eqn. (7) are fixed, and therefore do not require real-time adjustment. If real-time adjustment is desired, the values for these bounds can be included in the parameter vector $X$.

The real-time system requires the implementation of the polytopes and affine expressions. In addition, eqn. (15) is required to propagate the state mean. Storage arrays are required to save $(N-1)n_u$ previous input values, $(N-1)n_y$ previous output values and $(N-1)n$ previous state values.

The algorithm is summarized in

4 Example

We consider the following discrete-time system from literature (Rao & Rawlings, 2002):

$$
\begin{align*}
    x_{k+1} &= \begin{bmatrix} 0.9962 & 0.1949 \\ -0.1949 & 0.3815 \end{bmatrix} x_k + \begin{bmatrix} 0.03393 \\ 0.1949 \end{bmatrix} w_k \\
    y_k &= [1 \ -3] x_k + v_k
\end{align*}
$$

(25)

where $k = 1, 2, \ldots$; \{\(v_k\)\} is a sequence of independent, zero-mean, normally distributed random variables with variance 0.01; and \(w_k = |z_k|\) where \{\(z_k\)\} is a sequence of independent, zero-mean, normally distributed random variables with unit variance. Eqns. (25) and (26), are the state space realization of the continuous-time system $G(s) = (-3s + 1)/(s^2 + 3s + 1)$ sampled with a zero-order hold and sampling time of 0.3.

The initial state, \(x_1\), is normally distributed with zero mean and variance 1. Note that for this example, without loss of generality, there are no deterministic inputs, \(u_k\). Given the covariances of \(v_k\) and \(w_k\), the constrained estimation problem is designed with \(Q = 1\), \(R = 0.01\).

To capture additional knowledge of the random sequence \{\(w_k\)\}, the inequality constraint

$$
    w_k \geq 0
$$

(27)

is added.

For the mp-QP implementation, a horizon of $N = 3$ is chosen, which translates to $N-1 = 2$ inequality constraints (i.e., \(\hat{w}_{t-1,y} \geq 0\) and \(\hat{w}_{t-2,y} \geq 0\)). Expressing these constraints in the form of eqn. (16) trivially leads to the following matrices:
The steady state Kalman filter covariance matrix, calculated as the solution of the Riccatti equation, eqn. (14), at steady state is

\[
P = \begin{bmatrix}
0.0649 & 0.0023 \\
0.0023 & 0.0386 \\
\end{bmatrix}
\] (29)

The matrices \( H, F \) in eqn. (16), calculated from the expressions given in Appendix B, are:

\[
H = \begin{bmatrix}
345.3 & -230.8 & -55.10 & -48.47 \\
-230.8 & 521.2 & 26.51 & 1.491 \\
-55.10 & 26.51 & 17.03 & 3.620 \\
-48.47 & 1.491 & 3.620 & 16.17 \\
\end{bmatrix}
\]

\[
F = \begin{bmatrix}
15.44 & -0.9287 & 0 & 0 \\
-0.9287 & 25.96 & 0 & 0 \\
50.00 & -150.0 & 0 & 0 \\
79.04 & -47.48 & -27.54 & 0 \\
88.00 & -2.708 & -6.572 & -27.54 \\
\end{bmatrix}
\] (30)

The mp-QP results for this example are shown in Table 2. Note that the affine optimal expressions for \( Z \), eqn. (19), do not contain a constant term, which is due to the zero lower bounds in the inequality constraints, eqn. (27). Initial values \( \hat{x}_{11} = \hat{x}_{22} = [0 \ 0]^T \) are used to initialize both the simulation and the MHE.

Although the mean value of the sequence \( \{w_k\} \) is clearly non-zero (the actual value is \( 2/\sqrt{2\pi} \)), we deliberately set its value to 0, to test the robustness of the algorithm. For comparison purposes, we also calculated the Kalman filter predictions assuming the same statistics. We should point out that this is different from the comparison performed in (Rao & Rawlings, 2002), who used the correct mean for MHE and a zero mean for the Kalman filter. We believe our comparison is fairer because it tests the ability of both approaches to compensate for the incorrect mean.

The simulation results are shown in Figure 1 and Figure 2. The Kalman filter predictions are significantly biased for both state estimates, as well as for the output prediction, whereas the MHE estimates show much smaller bias and better accuracy. We see that the Kalman filter is significantly impacted by the incorrect statistics for the random variable \( w_k \). With the non-negativity constraint on \( w_k \), eqn. (27), the MHE is able to compensate partially for the incorrect value of the mean. Out of the 200 time
steps in the simulation, only 52 involve unconstrained state estimates, while in the remaining steps at least one constraint is activated, as shown in Figure 3.

We see that the MHE state estimates exhibit higher variance compared to actual. The variance can be reduced by increasing the moving horizon. Increasing the horizon to $N = 5$, as shown in Figure 4, reduces the variance. For $N = 5$, according to eqn. (23), there are at most 16 polytopes that must be characterized off-line, which is not a significant increase in computational load.

5 Conclusion and Discussion

The Kalman filter has been widely hailed as one of the most important engineering discoveries of the 20th century. Its use in many industries, most notably aerospace and automotive, is of paramount importance. One of the features that have made the Kalman filter so attractive, is its ease of design and implementation in terms of simple, stable recursive computations. In situations where the performance of the Kalman filter may not be adequate because of inequality constraints or strong nonlinearities, MHE can provide a solution. However, MHE requires the continual solution of an optimization problem in real time, thus eliminating one of the main attractions of Kalman filtering, namely simple recursive computation. The approach proposed in this paper re-introduces computational efficiency to constrained MHE, by replacing real-time optimization with function evaluation that is similar in spirit and not much more complicated than standard Kalman filtering. This will be beneficial for state estimation problems where constraints are important, but processing limitations prevent the real-time implementation of an optimizer. We illustrated our approach through a computer simulation example that reinforced two key advantages of MHE, namely the ability to include knowledge of inequality constraints in the estimation, and the ability to compensate for incorrect choices in the statistics. We expect that many current implementations of the Kalman filter that involve ad hoc handling of constraints would benefit from our approach. Possible applications areas include aeronautics, automotive, biomonitoring, and drug delivery, among others.

Additional study would further illustrate, refine, and expand our approach.

The problem of infeasibility can be considered for the real-time system, if infeasibility of the optimization problem is a possibility, although it was not an issue for the example considered in this paper. Two possible ways of dealing with infeasibility include failure modes, or through the use of soft constraints, as is frequently done with MPC.

We considered only linear models in this work. Moving horizon estimation with nonlinear models is an active research area, as is research into parametric techniques for nonlinear optimization. Future investigations should consider parametric moving horizon estimation with nonlinear models, starting with certain classes of nonlinear models. A challenge for the nonlinear case is that, it will not, in general, be possible to assume a constant state covariance matrix.

Given the close connection between estimation and control, we believe it would be useful to also consider the parametric solution of the control and estimation problems together.
References


Appendix A. Derivation of $\lambda_{A}$ and $Z$ in eqn. (21)

Assuming that the active constraints in matrix $D$ are linearly independent, the first order Karush-Kuhn-Tucker (KKT) conditions for the mp-QP problem, eqn. (16), are given by

\[
\begin{align*}
HZ + D^T \lambda &= 0 \\
\lambda_i (D^{(i)} Z - T^{(i)} - S^{(i)} X) &= 0, \quad i = 1, \ldots, q \\
\lambda &\geq 0
\end{align*}
\]  

(A.1)  
(A.2)  
(A.3)

where superscript $(i)$ represents row number, and $q$ is the number of constraints. From eqn. (A.1) we get

\[
Z = -H^{-1} D^T \lambda = H^{-1} D_A^T \lambda_A
\]  

(A.4)

Let $\lambda_A \neq 0$ and $\lambda_{\bar{A}} = 0$ represent the Lagrange multipliers corresponding to the active and inactive constraints, respectively. Then, for the active constraints, eqn. (A.2) yields

\[
D_A Z - T_A - S_A X = 0
\]  

(A.5)

and

where $D_A, T_A, S_A$ are corresponding matrices. Substituting eqn. (A.4) into eqn. (A.5) leads to

\[
\lambda_A = -(D_A H^{-1} D_A^T)^{-1} (T_A + S_A X)
\]  

(A.6)

Substituting (A.6) into (A.4) yields

\[
Z = H^{-1} D_A^T (D_A H^{-1} D_A^T)^{-1} (T_A + S_A X)
\]  

(A.7)

We see from eqn. (A.6) and eqn. (A.7) that $\lambda_A$ and $Z$ are affine functions of $X$.
Appendix B. Conversion of the MHE problem to the mp-QP form

In this appendix we derive the equations for the mp-QP form of the MHE problem

\[ \phi_t = \frac{1}{2} M^T H M + X^T F M + X^T K X \]  

(A.8)

For completeness we have included in the above expression the resulting constant term, \( X^T K X \), which may be removed since it does not affect the optimization. We assume a constant state covariance matrix, \( P \).

We begin by substituting the output equation, eqn. (11), into the moving horizon objective function, eqn. (4), and expanding the last term. After rearrangement we obtain

\[
\phi_t = \sum_{k=t-N+1}^{t} \hat{x}_k^T C^T R^{-1} C \hat{x}_k + \sum_{k=t-N+1}^{t-1} \hat{w}_k^T Q^{-1} \hat{w}_k + \hat{x}_{t-N+1|^t}^T P^{-1} \hat{x}_{t-N+1|^t}
\]

quadratic terms

\[-2 \sum_{k=t-N+1}^{t} y_k^T R^{-1} C \hat{x}_k - 2 \hat{x}_{t-N+1|^t}^T P^{-1} \hat{x}_{t-N+1|^t} \]

linear terms

\[+ \hat{x}_{t-N+1|^t}^T P^{-1} \hat{x}_{t-N+1|^t} \sum_{k=t-N+1}^{t} y_k^T R^{-1} y_k \]

constant terms

Letting \( \overline{R} = C^T R^{-1} C \) and substituting the state equation, eqn. (10), into the first summation yields:

\[
\sum_{k=t-N+1}^{t} x_k^T \overline{R} x_k = \hat{x}_{t-N+1|^t}^T \overline{R} \hat{x}_{t-N+1|^t} + (Az + Bu_t + G \hat{w}_t)^T \overline{R} (Az + Bu_t + G \hat{w}_t) + \ldots + (A^{N-1} z + A^{N-2} (Bu_{t-N+1} + G \hat{w}_{t-N+1}) + \ldots + Bu_{t-N+1} + G \hat{w}_{t-N+1})^T \overline{R} (A^{N-1} z + A^{N-2} (Bu_{t-N+1} + G \hat{w}_{t-N+1}) + \ldots + Bu_{t-N+1} + G \hat{w}_{t-N+1})
\]

\[= \hat{x}_{t-N+1|^t}^T \overline{R} \hat{x}_{t-N+1|^t} + \hat{x}_{t-N+1|^t}^T A^T \overline{R} A \hat{x}_{t-N+1|^t} + \ldots + \hat{x}_{t-N+1|^t}^T A^{(N-1)^T} \overline{R} A^{N-1} \hat{x}_{t-N+1|^t} + \hat{w}_{t-N+1|^t}^T (G^T \overline{R} G + G^T A^{(N-2)^T} \overline{R} A^{N-2} G) \hat{w}_{t-N+1|^t} + \ldots + \hat{w}_{t-N+1|^t}^T G^T A^{(N-2)^T} \overline{R} G \hat{w}_{t-N+1|^t} \]

\[+ \hat{w}_{t-N+2|^t}^T (G^T \overline{R} G + G^T A^{(N-3)^T} \overline{R} A^{N-3} G) \hat{w}_{t-N+2|^t} + \ldots + \hat{w}_{t-N+2|^t}^T G^T A^{(N-3)^T} \overline{R} G \hat{w}_{t-N+2|^t} \]

\[+ \ldots + \hat{w}_{t-2|^t}^T G^T A^{(N-3)^T} \overline{R} G \hat{w}_{t-2|^t} + \hat{w}_{t-2|^t}^T G^T \overline{R} G \hat{w}_{t-2|^t} + \ldots + \hat{w}_{t-2|^t}^T G^T \overline{R} G \hat{w}_{t-2|^t} \]

\[\hat{w}_{t-1|^t}^T H \hat{w}_{t-1|^t} \]
Substituting the state equation into the fourth term of eqn. (A.9), and expanding we obtain

\[
\begin{aligned}
&+ u_1^T (B^T \bar{R}B + D^T A^T \bar{R}CAB + \cdots + B^T A^{(N-2)T} \bar{R}A^{(N-2)} B)u_1 \\
&+ 2u_1^T (B^T A^T \bar{R}B + B^T A^{2T} \bar{R}AB + \cdots + B^T A^{(N-2)T} \bar{R}A^{(N-3)} B)u_2 + \cdots + 2u_1^T B^T A^{(N-2)T} \bar{R}Bu_{N-1} \\
&+ u_2^T (B^T \bar{R}B + B^T A^T \bar{R}AB + \cdots + B^T A^{(N-3)T} \bar{R}A^{(N-3)} B)u_2 \\
&+ 2u_2^T (B^T A^T \bar{R}B + B^T A^{2T} \bar{R}AB + \cdots + B^T A^{(N-3)T} \bar{R}A^{(N-4)} B)u_3 + \cdots + 2u_2^T B^T A^{(N-3)T} \bar{R}Bu_{N-1} \\
&+ \cdots + 2u_{N-2}^T D^T A^T \bar{R}Bu_{N-2} + 2u_{N-2}^T D^T A^T \bar{R}Bu_{N-1} + u_{N-2}^T B^T \bar{R}Bu_{N-1} \\
&\quad \triangleq U^TH_{uu}U \\
&+ 2x_{r-N+1y}^T (A^T \bar{R}G + A^{2T} \bar{R}AGB + \cdots + A^{(N-1)T} \bar{R}A^{(N-2)} G)\hat{w}_{r-N+1y} \\
&+ 2x_{r-N+1y}^T (A^{2T} \bar{R}G + A^T \bar{R}AG + \cdots + A^{(N-1)T} \bar{R}A^{(N-3)} G)\hat{w}_{r-N+2y} + \cdots + 2x_{r-N+1y}^T A^{(N-1)T} \bar{R}G\hat{w}_{r-N+1y} \\
&\quad \triangleq 2x_{r-N+1y}^T H_{uw}W \\
&+ 2u_1^T (B^T \bar{R}A + B^T A^T \bar{R}A^2 + \cdots + B^T A^{(N-2)T} \bar{R}A^{(N-1)} )\hat{x}_{r-N+1y} \\
&+ 2u_2^T (B^T \bar{R}A^2 + B^T A^T \bar{R}A^3 + \cdots + B^T A^{(N-3)T} \bar{R}A^{(N-1)} )\hat{x}_{r-N+1y} + \cdots + 2u_{N-1}^T B^T \bar{R}A^{N-1} \hat{x}_{r-N+1y} \\
&\quad \triangleq 2U^TF_{ex}\hat{x}_{r-N+1y} \\
&+ 2u_1^T (B^T \bar{R}G + B^T A^T \bar{R}AG + \cdots + B^T A^{(N-2)T} \bar{R}A^{(N-2)} G)\hat{w}_{r-N+1y} \\
&+ 2u_1^T (B^T \bar{R}G + B^T A^T \bar{R}G + \cdots + B^T A^{(N-2)T} \bar{R}A^{(N-3)} G)\hat{w}_{r-N+2y} + \cdots + 2u_1^T B^T A^{(N-2)T} \bar{R}G\hat{w}_{r-N+1y} \\
&+ 2u_2^T (B^T \bar{R}AG + \cdots + B^T A^{(N-3)T} \bar{R}A^{(N-2)} G)\hat{w}_{r-N+1y} \\
&+ 2u_2^T (B^T \bar{R}AG + B^T A^T \bar{R}AG + \cdots + B^T A^{(N-3)T} \bar{R}A^{(N-3)} G)\hat{w}_{r-N+2y} + \cdots + 2u_2^T B^T A^{(N-3)T} \bar{R}G\hat{w}_{r-N+1y} \\
&+ \cdots + 2u_{N-2}^T B^T \bar{R}AG\hat{w}_{r-2y} + 2u_{N-2}^T B^T \bar{R}G\hat{w}_{r-2y} + 2u_{N-1}^T B^T \bar{R}G\hat{w}_{r-1y} \\
&\quad \triangleq 2U^TF_{uw}W 
\end{aligned}
\]
\[
\sum_{k=t-N+1}^{t} y_j^T R^{-1} C \hat{x}_j = y_{t-N+1}^T R^{-1} C \hat{x}_{t-N+1|t} + y_{t-N+2}^T R^{-1} C A \hat{x}_{t-N+1|t} + \ldots + y_t^T R^{-1} C (A^{N-1} \hat{x}_{t-N+1|t} + A^{N-2} B u_{t-N+1} + A^{N-2} G \hat{w}_{t-N+1|t} + \ldots + B u_{t-1|t} + G \hat{w}_{t-1|t})
\]
\[
= y_{t-N+1}^T R^{-1} C \hat{x}_{t-N+1|t} + y_{t-N+2}^T R^{-1} C A \hat{x}_{t-N+1|t} + \ldots + y_t^T R^{-1} C A^{N-1} \hat{x}_{t-N+1|t}
\]
\[
+ y_{t-N+1}^T R^{-1} C D u_1 + \ldots + y_{t}^T R^{-1} (A^{N-2} B u_{t-N+1} + B u_{t-1})
\]
\[
+ \frac{y_t^T F_{yw} G \hat{w}_{t-N+1|t} + \ldots + y_t^T R^{-1} (A^{N-2} G \hat{w}_{t-N+1|t} + \ldots + G \hat{w}_{t-1})}{\frac{\delta y^T F_{yw}}{}}
\]

(A.11)

Substituting Eqns. (A.10) and (A.11) into (A.9), and using the definitions of the vectors \( M, X, W, U, Y \) from section 3, the MHE problem can be expressed in multiparametric form, eqn. (A.8), with the associated matrices defined as follows:

\[
H = \frac{1}{2}(\hat{H} + \hat{H}^T) \quad \text{(to make symmetric)}
\]

(A.12)

\[
\hat{H} = \begin{bmatrix}
H_{zz} + \bar{R} + P^{-1} & H_{zw} \\
H_{zw}^T & H_{ww} + \hat{Q}_{N-1}
\end{bmatrix}
\]

\[
F = \begin{bmatrix}
P^{-1} & 0 \\
F_{yz} & F_{yw} \\
F_{uz} & F_{uw}
\end{bmatrix}
\]

(A.13)

\[
K = \begin{bmatrix}
P^{-1} & 0 & 0 \\
0 & \bar{\hat{R}}_N & F_{yu} \\
0 & 0 & H_{uu}
\end{bmatrix}
\]

(A.14)

where:

\[
H_{zz} = \hat{A}^T R_{N-1} \hat{A}, \quad H_{zw} = \hat{A}^T R_{N-1} \hat{G}, \quad H_{ww} = \hat{G}^T R_{N-1} \hat{G},
\]

\[
F_{uz} = \hat{B}^T \bar{\hat{R}}_{N-1} \hat{A}, \quad H_{uu} = \hat{B}^T \bar{\hat{R}}_{N-1} \hat{B}
\]
\[ \hat{A} = \begin{bmatrix} A \\
A^2 \\
\vdots \\
A^{N-1} \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} B & 0 & \cdots & 0 \\
AB & B & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
A^{N-2}B & A^{N-3}B & \cdots & B \end{bmatrix}, \quad \hat{G} = \begin{bmatrix} G & 0 & \cdots & 0 \\
AG & G & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
A^{N-2}BG & A^{N-3}G & \cdots & G \end{bmatrix} \]

\[ \hat{Q}_N = \begin{bmatrix} Q^{-1} \\
\vdots \\
Q^{-1} \end{bmatrix}, \quad \hat{R}_N = \begin{bmatrix} R^{-1} \\
\vdots \\
R^{-1} \end{bmatrix}, \quad \hat{R}_N = \begin{bmatrix} \bar{R} \\
\vdots \\
\bar{R} \end{bmatrix} \]

\[ F_{yz} = \hat{R}_N, \quad F_{yw} = \hat{R}_N \begin{bmatrix} C \\
CA \\
\vdots \\
CA^{N-1} \end{bmatrix}, \quad F_{yw} = \hat{R}_N \begin{bmatrix} 0 & 0 & \cdots & 0 \\
CG & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
CA^{N-2}G & CA^{N-3}G & \cdots & CG \end{bmatrix} \]

\[ F_{yw} = \hat{R}_N \begin{bmatrix} 0 & 0 & \cdots & 0 \\
CB & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
CA^{N-2}B & CA^{N-3}B & \cdots & CB \end{bmatrix} \]

\[ F_{zw} = \begin{bmatrix} F_{zw}^{(1)} \\
\vdots \\
F_{zw}^{(N-1)} \end{bmatrix}, \text{where superscript } \{i\} \text{ denotes row } i \]

For \( N \leq 3 \),

\[ F_{zw}^{(i)} = \begin{bmatrix} B & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
A^{N-1}B & A^{N-2}B & \cdots & A^{N-1}B \end{bmatrix} \begin{bmatrix} R^{-1} \end{bmatrix} \begin{bmatrix} A^{i-1}G & \cdots & G & 0 & \cdots & 0 & 0 \\
A^{i-2}G & \cdots & AG & G & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
A^{N-2}G & \cdots & \cdots & \cdots & \cdots & \cdots & AG & G \end{bmatrix} \]

For \( N > 3 \),

\[ F_{zw}^{(i)} = \begin{bmatrix} B & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
A^{N-1}B & A^{N-2}B & \cdots & A^{N-1}B \end{bmatrix} \begin{bmatrix} R^{-1} \end{bmatrix} \begin{bmatrix} A^{i-1}G & \cdots & G & 0 & \cdots & 0 & 0 \\
A^{i-2}G & \cdots & AG & G & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
A^{N-2}G & \cdots & \cdots & \cdots & \cdots & \cdots & AG & G \end{bmatrix} \]
\[
F_{uv}^{(i)} = \begin{cases}
\begin{bmatrix}
B
\end{bmatrix}^T R^{-1} \begin{bmatrix}
G & G & 0 & \cdots & 0
\end{bmatrix}, & i = 1
\end{cases}
\]
\[
\begin{cases}
\begin{bmatrix}
A^{N-2}B
\end{bmatrix}^T R^{-1} \begin{bmatrix}
A^{N-2}G & A^{N-3}G & \cdots & AG & G
\end{bmatrix}, & i = 1
\end{cases}
\]
\[
\begin{cases}
\begin{bmatrix}
B
\end{bmatrix}^T R^{-1} \begin{bmatrix}
A^{i-1}G & \cdots & G & 0 & \cdots & 0 & 0
\end{bmatrix}, & i = 2, \ldots, N-1
\end{cases}
\]
\[
\begin{cases}
\begin{bmatrix}
A^{N-i-1}B
\end{bmatrix}^T R^{-1} \begin{bmatrix}
A^{N-i-2}G & \cdots & \cdots & \cdots & AG & G
\end{bmatrix}, & i = 2, \ldots, N-1
\end{cases}
\]
### Table 1. Summary of mp-MHE algorithm

**Off-line calculations**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>Based on system matrices $A, B, G,$ and $C$ (eqns. (10) and (11)) and covariances $Q, R,$ and $P$, calculate matrices for the quadratic program (eqn. (16)) according to equations in Appendix B and those describing the constraints.</td>
</tr>
<tr>
<td>Step 1</td>
<td>Select a value of the parameter vector $X$. Solve the resulting quadratic program (eqn. (18)).</td>
</tr>
<tr>
<td>Step 2</td>
<td>Determine the affine relationships describing $Z_{opt}$ near $X$ (eqn. (19)). Express optimum in terms of $M_{opt}$ via eqn. (17): $M_{opt} = Z_{opt} - H^{-1}F^{T}X$.</td>
</tr>
<tr>
<td>Step 3</td>
<td>Characterize the critical region (linear polytope) over which the affine relationship holds (eqn. (20)).</td>
</tr>
<tr>
<td>Step 4</td>
<td>Reverse one of the constraints in the set of inequalities (eqn. (18)) and solve the feasibility linear program (eqn. (20)); if not feasible, consider reversing another constraint.</td>
</tr>
<tr>
<td>Step 5</td>
<td>Go to step 1 unless all polytopes (step 3) have been characterized.</td>
</tr>
</tbody>
</table>

**On-line calculations**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>Access current value of $y_{j}$; retrieve past values $u_{t-1}$ and $z_{t-N+1:t-N}$; update $\bar{x}_{t-N+1:t-N}$ via eqn. (15); update parameter vector $X$ (eqn (24)).</td>
</tr>
<tr>
<td>Step 2</td>
<td>Select appropriate polytope (eqn. (20)) and corresponding affine expression to calculate $M_{opt}$ (eqns. (19) and (17)).</td>
</tr>
<tr>
<td>Step 3</td>
<td>From $M_{opt}$ calculate the trajectory of state estimates from $t-N+1$ through $t$ using the model (eqn. (10)). The current state estimate is $\hat{x}_{t+N}^{mh}$.</td>
</tr>
<tr>
<td>Step 4</td>
<td>Update storage arrays associated with past inputs, output, and past moving horizon estimates, for use at time $t+1$.</td>
</tr>
</tbody>
</table>
Table 2. Multiparametric solution to numerical example in section 4

<table>
<thead>
<tr>
<th>Characterization of polytope for $X \preceq \begin{bmatrix} \tilde{x}<em>{1-2y-3}^1 &amp; \tilde{x}</em>{1-2y-3}^2 &amp; y_{1-2} &amp; y_{1-1} &amp; y_{1} \end{bmatrix}^T$</th>
<th>Formula for calculation of optimal $\begin{bmatrix} \tilde{x}<em>{1-2y}^T &amp; \tilde{w}</em>{1-2y} &amp; \tilde{w}_{1-1y} \end{bmatrix}^T$</th>
</tr>
</thead>
</table>
| Polytope 1:  
$\begin{bmatrix} -1.158 & -0.5319 & -0.6527 & 1.521 & -0.3379 \\ -1.306 & -0.6620 & -0.0357 & -0.6616 & 1.356 \end{bmatrix} X \preceq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  
$X = \begin{bmatrix} 0.5137 & 0.2572 & 0.0509 & 0.1012 & 0.1565 \\ 0.1676 & 0.1125 & -0.3062 & 0.0298 & 0.0520 \\ 1.158 & 0.5319 & 0.6527 & -1.521 & 0.3379 \\ 1.306 & 0.6620 & 0.0357 & 0.6616 & -1.356 \end{bmatrix}$ |
| Polytope 2 ($\tilde{w}_{1-ly} = 0$):  
$\begin{bmatrix} 5.604 & 2.575 & 3.160 & -7.366 & 1.636 \\ -0.2625 & -0.1827 & 0.5525 & -2.033 & 1.661 \end{bmatrix} X \preceq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  
$X = \begin{bmatrix} 0.0860 & 0.0607 & -0.1903 & 0.6632 & 0.0317 \\ 0.0372 & 0.0526 & -0.3797 & 0.2011 & 0.0139 \end{bmatrix}$ |
| Polytope 3 ($\tilde{w}_{1-2y} = 0$):  
$\begin{bmatrix} 5.163 & 2.618 & 0.1413 & 2.616 & -5.364 \\ -0.1967 & -0.0447 & -0.6264 & 2.008 & -1.336 \end{bmatrix} X \preceq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  
$X = \begin{bmatrix} 0.0686 & 0.0315 & 0.0387 & -0.1244 & 0.6190 \\ 0.0198 & 0.0376 & -0.3103 & -0.0450 & 0.2055 \\ 0.1967 & 0.0447 & 0.6264 & -2.008 & 1.336 \end{bmatrix}$ |
| Polytope 4 ($\tilde{w}_{1-ly} = \tilde{w}_{1-2y} = 0$):  
$\begin{bmatrix} 2.827 & 0.6425 & 9.004 & -28.86 & 19.21 \\ 3.082 & 2.145 & -6.485 & 23.86 & -19.50 \end{bmatrix} X \preceq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  
$X = \begin{bmatrix} 0.0432 & 0.0232 & -0.0771 & 0.2467 & 0.3721 \\ 0.0136 & 0.0362 & -0.3300 & 0.0184 & 0.1633 \end{bmatrix}$ |
Figure 1. Comparison of state predictions for MHE ( $N = 3$ ) and Kalman filter
Figure 2. Comparison of output predictions for MHE ($N = 3$) and Kalman filter
Figure 3. Active polytopes in implementation of mp-QP. Numbers correspond to Table 2, i.e. 1=no constraint active, 2=constraint on $\hat{w}_{t-1}$ active, 3=2=constraint on $\hat{w}_{t-2}$ active, 4=2=constraint on both $\hat{w}_{t-1}$ and $\hat{w}_{t-2}$ active.
Figure 4. Comparison of state predictions for MHE ($N = 5$) and Kalman filter estimators