Multiweight Optimization in Optimal Bounding Ellipsoid Algorithms

D. Joachim† and J.R. Deller, Jr.‡†

†CORRESPONDING AUTHOR Tulane University; Dept. Electrical Engineering & Computer Science; 211 Stanley Thomas Hall; 6823 St. Charles Ave.; New Orleans, LA 70118 USA; email: joachimd@tulane.edu; Tel: +1-504-865-5785; Fax: +1-504-862-3293

‡ Michigan State University; Dept. Electrical & Computer Engineering; 2120 Engineering Bldg.; East Lansing, MI 48824-1226 USA; email: deller@egr.msu.edu; Tel: +1-517-353-8840; Fax: +1-517-353-1980

Running title: Multiweight Optimization in Optimal Bounding Ellipsoid Algorithms

Keywords: System identification, Set-membership identification

Abstract.

Optimal Bounding Ellipsoid (OBE) algorithms offer an attractive alternative to traditional least squares methods for identification and filtering problems involving affine-in-parameters signal and system models. The benefits – including low computational efficiency, superior tracking ability, and selective updating that permits processor multi-tasking – are enhanced by multiweight (MW) optimization in which the data history is considered in determining update times and optimal weights on the observations. MW optimization for OBE algorithms is introduced, and an example MW-OBE algorithm implementation is developed around the recent quasi-OBE algorithm. Solution existence and uniqueness are discussed, and simulation studies are used to illustrate performance benefits.
In existing OBE algorithms, optimization takes place pointwise in time; all conditions (in particular, previous weights) extant at the time of optimization remain fixed. The globally optimal, causal solution at time $n$ would diminish the solution set in light of all observations in the interval $[1,n-1]$. Given a system of order $m$, a MW-OBE algorithm “revisits” $K$ past weights when the observation set at time $n$ is deemed innovative, so that the ellipsoid is optimally diminished with respect to the current and past $K$ observations, conditioned upon information known at the $n - K - 1$st update. The number of revisited past weights, $K$, may be time-varying and must be less than the system order $m$. The corresponding past data vectors are not required to be sequential. With respect to pointwise weighting, simultaneous optimization over multiple weights offers greater flexibility in shaping the hyperellipsoidal set, thus providing a better fit to the exact feasibility set.

This paper begins with the formulation of a general MW-OBE algorithm. The general algorithm is then optimized in the framework of the quasi-OBE (QOBE) [16]. Simulation results then demonstrate the potential of the developed MW-QOBE algorithms.

II. BACKGROUND

A. OBE algorithms

OBE algorithms are used to identify affine-in-parameters models of the form

$$y_n = \theta^*_s x_n + \varepsilon_{ns}$$

in which $\theta_s \in \mathbb{R}^m$ is the unknown “true” parameter vector to be identified; $\{x_n\}$ is a sequence of measurable vectors of dimension $m$; and $\{\varepsilon_{ns}\}$ is a “true” but unknown error sequence. It is assumed that the error has a known pointwise energy bound given by (1). The “true” model is posed only for analysis purposes and provides the background from which actual parameter vector estimates are derived. Given data on times $t \in [1,n]$, an exact feasibility set, say $\mathcal{P}_n$, of estimates for $\theta_s$ whose elements are consistent with these bounds is given by (see Fig. 1)

$$\mathcal{P}_n = \bigcap_{t=1}^{n} \mathcal{H}_t,$$

with $\mathcal{H}_t \overset{\text{def}}{=} \{ \theta : \varepsilon^2_{ts} = |y_t - \theta^T x_t|^2 < \gamma^2_t \}$.

OBE algorithms work with an hyperellipsoidal set, say $\mathcal{E}_n$, that is guaranteed to contain $\mathcal{P}_n$, hence $\theta_s$. The observations are scrutinized with respect to their ability to “shrink” $\mathcal{E}_n$, hence to more tightly bound $\mathcal{P}_n$. At time $n$, the hyperellipsoid is given by (e.g., [5])

$$\mathcal{E}_n \overset{\text{def}}{=} \{ \theta : (\theta - \theta_n)^T \mathcal{C}_n (\theta - \theta_n) \leq \kappa_n \}$$

in which $\mathcal{C}_n$ is a weighted covariance matrix of the observations,

$$\mathcal{C}_n = \sum_{t=1}^{n} q_{t,n} x_t x_t^T,$$

$k_n$ is the scalar

$$k_n = \theta^T \mathcal{C}_n \theta_n + \sum_{t=1}^{n} q_{t,n} (\gamma^2_t - y^2_t),$$

and $\theta_n$, the center of $\mathcal{E}_n$, is a weighted least-square-error estimator of $\theta_s$ at time $n$,

$$\theta_n = \mathcal{P}_n \mathcal{C}_n, \quad \text{with} \ \mathcal{P}_n \overset{\text{def}}{=} \mathcal{C}_n^{-1} \quad \text{and} \quad \mathcal{C}_n \overset{\text{def}}{=} \sum_{t=1}^{n} q_{t,n} y_t x_t.$$
and the effort of updating can be avoided at this time. Depending on the properties of the sequence \( \{ \varepsilon_{n,k} \} \), OBE algorithms update 10 percent of the time or less.

### B. Motivation for a new algorithm

All published OBE algorithms can be manipulated into the formal framework described in Section II-A, provided that time-varying (“n-dependent”) weight sequences are used as in (5)–(7) [5]. In all published algorithms, however, this time dependence (if any) has a simple structure arising from either a generalized “forgetting factor” (e.g., c.f. [6] and [7]), or some heuristic measures to induce adaptation (e.g., [17]). In no reported case is there an attempt to reoptimize any of the weights applied prior to time \( n \) in light of new measurements \( x_n \) and \( y_n \). That is, all optimization in existing OBE algorithms can be accomplished by manipulation of the current weight only except for possible inherent scaling of past weights. This can be inferred directly from the work in [5]. The globally optimal, but computationally infeasible causal solution at time \( n \) would optimize all weights \( \{ q_{t,n} \}_{t=1}^n \), in light of all known information \( \{ (x_t, y_t) \}_{t=1}^n \).

The MW-OBE algorithm can approach the performance of a globally-optimized algorithm at each time, by “revisiting” \( K \geq 0 \) past weights when data at time \( n \) are received and deemed innovative. The revision of past weights is made by additive adjustments to existing weight values, subject to the constraints that revised weights remain nonnegative and that the number of revised weights be less than the system order.

### III. OBE WITH MULTIPLE WEIGHT OPTIMIZATION

#### A. Introduction

This section introduces the general class of MW-OBE algorithms and presents background formulation for the MW-QOBE and MW-SM-WRLS algorithms.

At each time \( n \), conventional OBE algorithms update the previous covariance matrix \( C_{n-1} \), by incorporating a weighted outer product of the current data vector, if this vector is deemed innovative (e.g., [5]). This process, although efficient, may result in large ellipsoid volumes [10], [18], often due to the shape and size of the underlying polytope \( P_n \) [19]. However, when adequate persistency of excitation and infinite visitation conditions are present [18], pointwise reduction in ellipsoid volume may be improved by a joint weight optimization [20], [21].

---

**TABLE I**

**THE MW-OBE ALGORITHM.**

**I. Initialization:**
1. \( \theta_K = 0 \), \( \kappa_K = 1 \) and \( P_K = \frac{1}{\mu} I \), \( \mu \) small.
2. \( q_K = 0 \).

**II. Recursion:**
For \( n = K + 1, K + 2, \ldots \)

- **R:** Form \( X_n, y_n \), and \( \gamma_n \),
  \[ e_{n,n-1} = y_n - X_n^T \theta_{n-1} \]
  \[ R_n = P_{n-1} X_n \text{ and } G_{n,n-1} = X_n^T R_n \]
  If \( X_n \) is NOT “innovative,” goto R.
  - determine “optimal” weight vector \( \lambda_n \),
  \[ K_n = R_n [ \Lambda_{n-1}^{-1} + G_{n,n-1}^{-1} ]^{-1} \]
  \[ P_n = P_{n-1} - K_n R_n \]
  \[ \theta_n = \theta_{n-1} + K_n e_{n,n-1} \]
  \[ \kappa_n = \kappa_{n-1} + 1 \]
  \[ \gamma_n^T \Lambda_n \gamma_n - e_{n,n-1}^T [ \Lambda_{n-1}^{-1} + G_{n,n-1}^{-1} ]^{-1} e_{n,n-1} \]

- Next \( n \), goto R.

MW-OBE optimization begins by isolating the last \( K + 1 \) outer products in the covariance matrix (5)

\[ C_n = \sum_{t=1}^{n-K} q_{t,n} x_t x_t^T + \sum_{t=n-K}^{n} q_{t,n} x_t x_t^T \quad (8) \]

over which the block of weights \( \{ q_{t,n} \}_{t=n-K}^n \) is to be reoptimized at time \( n \). To express (8) in matrix form, let

\[ X_n = [ x_{n-K} \ldots x_n ] \]

and \( q_{t,n} \) the vector containing the weights applied to the vectors in \( X_n \) at time \( t \),

\[ q_n = [ q_{n-K,n} \ldots q_{n,n} ] \]

In these terms, (8) is written

\[ C_n = \sum_{t=1}^{n-K} q_{t,n} x_t x_t^T + X_n Q_{n,n} X_n^T \quad (11) \]

where \( Q_{n,n} = D(q_{n,n}) \), the diagonal matrix with the elements of vector \( q_{n,n} \) along the diagonal. Let \( \lambda_n \) represent the adjustments to the weights applied to observation matrix \( X_n \) at time \( n \), with corresponding diagonal matrix \( \Lambda_n = D(\lambda_n) \). The weight adjustments represent the difference between the \( a \) priori and newly computed weights, with the most recent “adjustment” being a modification to a zero (by definition) weight. The recursive expression (11) in terms of the weight adjustments becomes

\[ C_n = C_{n-1} + X_n \Lambda_n X_n^T \quad (12) \]

The (composite) weights at any time must be nonnegative, \( q_{n,n} = \lambda_n + q_{n,n-1} > 0 \). The block weight
assignment strategy is similar to sliding window over the sequence of data vectors, where the weights assigned to a data vector vary with time \( n \), but only vary during the time interval \([n,n+K]\). Accordingly, the time-varying weights are computed at time \( n \) by

\[
q_{t,n} = \begin{cases} 
\sum_{i=t}^{t+K} \lambda_{t,i}, & 0 \leq t < n - K \\
\sum_{i=t}^{n} \lambda_{t,i}, & n - K \leq t \leq n 
\end{cases},
\]

(13)

with the constraint \( q_{t,n} \geq 0 \) for any \( t \) and \( n \). The weight adjustments \( \lambda_{t,n} \) are zero for \( t \not\in [n-K \leq t \leq n] \).

This formulation allows the covariance matrix to be updated at each \( n \) with a set of weight adjustments to the present and past \( K \) observation vectors. In the next section, the general MW-OBE recursions are derived, beginning with (12).

\section{B. Recursions}

As in conventional OBE algorithms, the MW-OBE algorithm recursively computes the inverse covariance matrix, \( P_n = C_n^{-1} \), the ellipsoid center, \( \theta_n \), and the scalar \( \kappa_n \). A recursion for \( P_n \) is obtained by applying the matrix-inversion lemma [22] to the updating equation (12), yielding

\[
P_n = P_{n-1} - P_{n-1}X_nH_n^{-1}\Lambda_nX_n^TP_{n-1}^{-1},
\]

(14)

in which \( H_n \) is obtained as \( I + \Lambda_nG_{n|n-1} \) and \( G_{n|n-1} = X_n^TP_{n-1}X_n \). The existence of the matrix \( H_n^{-1} \Lambda_n = [\Lambda_n^{-1} + G_{n|n-1}]^{-1} \) is contingent upon all positive weight adjustments to \( \Lambda_n \) and upon the invertibility of \( [\Lambda_n^{-1} + G_{n|n-1}]^{-1} \). In practice this constraint is satisfied by omitting data vectors with zero weights. Another constraint, the reason for which will become clear below, is that \( K \) be no larger than \( m \).

Derivations of recursions for the ellipsoid center \( \theta_n \) and \( \kappa_n \) [23] yield:

\[
\theta_n = \theta_{n-1} + P_{n-1}X_n\Lambda_n\varepsilon_{n|n-1}, \quad \text{and} \quad \kappa_n = \kappa_{n-1} + \gamma_n^T\Lambda_n\gamma_n - \varepsilon_{n|n-1}^T H_n^{-1} \Lambda_n \varepsilon_{n|n-1}
\]

(15)

where \( \varepsilon_{n|n-1} \) is as defined in Table I. Recursions (15) and (15) reduce to the conventional SM-WRLS recursion [10] when \( K = 0 \).

\section{C. A posteriori error vector and energy matrices}

The objectives of identifying a set of solutions that closely identifies the parameters of a system is achieved by seeking to reduce either the distance (in some sense) between the parameter estimate and the “true” parameter vector, or the ellipsoid size. The latter in effect brings the estimator to a closer neighborhood of the true parameter vector when \( E_n \) becomes “small.” The progress of OBE algorithms in achieving the given objective is observable in the sequence of error vectors \( \varepsilon_{n|n-1} \), and energy matrices \( G_{n|n-1} \). In the process of reducing the current ellipsoid size, OBE algorithms “re-map” the error to satisfy inequality (1). As a result, \( \varepsilon_{n|n} \) and \( G_{n|n} \) provide important insights into algorithm behavior. In this section we express these quantities in their a posteriori representations to better illustrate MW-OBE behavior.

To satisfy the error bound constraint (1), the a posteriori error vector \( \varepsilon_{n|n} \) must belong to the closed hyperbox \( \{u \in \mathbb{R}^{K+1} : |u| \leq \gamma_n \} \) (Fig. 2). The transformation which maps \( \varepsilon_{n|n-1} \) to this hyper-box at time \( n \) is found by expressing \( \varepsilon_{n|n} \) in terms of \( \varepsilon_{n|n-1} 
\]

\[
\varepsilon_{n|n} = y_n - X_n^T\theta_n = y_n - X_n^T\theta_{n-1} - G_{n|n}\Lambda_n\varepsilon_{n|n-1} = \varepsilon_{n|n-1} - G_{n|n}\Lambda_n\varepsilon_{n|n-1} = [I - G_{n|n}\Lambda_n] \varepsilon_{n|n-1} = H_n^{-T} \varepsilon_{n|n-1}.
\]

(16)

From (16) we note that the mapping of error vector \( \varepsilon_{n|n-1} \) to its a posteriori image \( \varepsilon_{n|n} \) is achieved through the transformation \( H_n^{-T} \). In light of (16), the expression for the scalar \( \kappa_n \) takes a simpler form

\[
\kappa_n = \kappa_{n-1} + \gamma_n^T\Lambda_n\gamma_n - \varepsilon_{n|n-1}^T H_n^{-1} \Lambda_n \varepsilon_{n|n-1}.
\]

(17)

or, using \( \circ \) to represent the Hadamard product [24],

\[
\kappa_n = \kappa_{n-1} + \lambda_n^T (\gamma_n \circ \gamma_n - \varepsilon_{n|n-1} \circ \varepsilon_{n|n}).
\]

(18)

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig2.png}
\caption{Rectangular error bound constraint (1) (dashed line) when \( K = 1 \) and \( m = 2 \). An error \( \varepsilon_{n|n-1} \) originally violating the bound condition, in light of new observations, is mapped to a location inside the rectangle \( \{u : |u| \leq \gamma \} \) in its a posteriori \( \varepsilon_{n|n} \) form.}
\end{figure}

The weighted energy matrix (divided by \( \kappa_n \)) represents the projection of the data vectors in \( X_n \) on the current ellipsoid axes. The a posteriori matrix \( G_{n|n} \)
is found by substituting the right side of (14) into

\[ X_n^T P_n X_n, \]

yielding

\[ G_{n|n} = H_n^T G_{n|n-1}. \]  

Since \( G_{n|n} \) is symmetric, (19) is also equal to

\[ G_{n|n-1} H_n^{-1}. \]  

Thus, the matrix \( H_n^{-1} \) may be expressed as

\[ H_n^{-1} = (G_{n|n-1})^{-1} (G_{n|n}). \]  

Since \( P_{n-1} \) is positive definite, \( a^T G_{n|n-1} a = X_n a^T P_{n-1} (a X_n^T) \geq 0 \) for all \( a \in \mathbb{R}^{K+1} \), or \( G_{n|n-1} \) is positive semi-definite [22]. The same is deduced for \( G_{n|n} \). Further neither matrix has a zero eigenvalue because each is non-singular when the algorithm is updating.

D. Ellipsoid volume

The volume of ellipsoid \( E_n \) is proportional to the determinant of \( \kappa_n P_n \) and therefore the ratio

\[ \frac{\det (\kappa_n P_n)}{\det (\kappa_{n-1} P_{n-1})} = \left( \frac{\kappa_n}{\kappa_{n-1}} \right)^m \frac{\det (P_n)}{\det (P_{n-1})} \]  

(21)

represents an appropriate measure for the change in volume [10]. In light of (20), the ratio of determinants in (21) is equivalent to

\[ \det (H_n) = \frac{\det (P_{n-1})}{\det (P_n)}. \]  

(22)

Consequently (21) may be written as

\[ \frac{\det (\kappa_{n-1} P_{n-1})}{\det (\kappa_n P_n)} = \det (H_n) \left( \frac{\kappa_{n-1}}{\kappa_n} \right)^m. \]  

(23)

Results (22) and (23) will facilitate future developments.

E. Algorithm

MW-OBE algorithms (Table I) are similar to their simpler counterparts in that they check the observation matrix for innovation at each iteration. If the observation is deemed useful, a new inverse covariance matrix and ellipsoid center are computed. An advantage of MW-OBE over conventional OBE algorithms is the flexibility in selection and number of past observations to revisit. Updating the past sequentially-numbered \( K \) weights often results in reconsideration of previously-rejected observation vectors which typically remain uninformative. MW-OBE does not specify which past data are to be revisited. Observations taken in the past are more likely to be innovative.

F. Computational costs

The following discussion is restricted to the computational costs of the general form of the MW-OBE algorithm. These recursions, common to all MW-OBE algorithms, are only performed when a data matrix is deemed informative. MW-OBE algorithms compute the error vector \( \varepsilon_{n|n-1} \) and the energy matrix \( G_{n|n-1} \) in checking for innovation in the observation matrix. Although these costs are included in the following considerations, costs due to particular optimization methods (actual innovation check and weight generation) are excluded.

Detailed per update cost is shown in Table II. When \( \tilde{K} = 1 (K = 0) \), corresponding to the conventional OBE algorithm, this cost becomes \( 3/2(m^2 + 3m + 4) \), comparable to that of the algorithm described in [6] at \( O(m^2) \). This \( O(m^2) \) performance is maintained for \( m \gg \tilde{K} \). When \( K \approx m \), the per-update cost increases to \( O(m^4) \).

In order to implement real-time identification, we limit simulations to the re-visitiation of one or two past weights. In so doing, the computational cost are kept within a “reasonable” range (see Table III and Fig. 3) and with small values of \( K \), the inversion of matrix \( [A_n^{-1} + G_{n|n-1}]^{-1} \) is relatively inexpensive. The expression for computational cost shown in Table II includes the computations of \( \varepsilon_{n|n-1} \) and \( G_{n|n-1} \), which are used in the innovation check.

![Fig. 3. Computational costs (number of multiplications) of a single MW-OBE iteration for \( K = 0, 1, 2 \) and 3 for system orders \( m = 2, 4, 8, 12 \) and 16. \( K = 0 \) represents the conventional OBE algorithm case.](image)

OBE algorithms update the parameter estimator by using a small percentage of observation vectors, typically 5 - 10% [6], [10], [17], [25]. MW-OBE algorithms use...
TABLE II
Detailed computational costs of MW-OBE algorithms per update. The computation of $\varepsilon_{n|n-1}$ and $G_{n|n-1}$ are included although they are usually considered part of the innovation check. For simplicity, the notation $\tilde{K}$ is used to denote $K + 1$.

<table>
<thead>
<tr>
<th>Expression</th>
<th>No. Multiplies</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_n = P_n X_n$</td>
<td>$m\tilde{K}$</td>
</tr>
<tr>
<td>$G_{n</td>
<td>n-1}$</td>
</tr>
<tr>
<td>$K_n = R_n(A\gamma_n^{-1} + G_{n</td>
<td>n-1})^{-1}$</td>
</tr>
<tr>
<td>$P_n = P_{n-1} - K_n R_n^\prime$</td>
<td>$(\tilde{K}^3 + m\tilde{K}^2 + \tilde{K})$</td>
</tr>
<tr>
<td>$\theta_n = \theta_{n-1} + K_n \varepsilon_{n</td>
<td>n-1}$</td>
</tr>
<tr>
<td>$\kappa_n$</td>
<td>$m\tilde{K}$</td>
</tr>
<tr>
<td>Total</td>
<td>$(3m\tilde{K}/2)(m + \tilde{K} + 2)$</td>
</tr>
<tr>
<td></td>
<td>$+ \tilde{K}(\tilde{K}^2 + \tilde{K} + 4)$</td>
</tr>
</tbody>
</table>

TABLE III
Computational costs of MW-OBE algorithms per update for $K = 0, 1, 2$ and $3$ as functions of the model order $m$.

<table>
<thead>
<tr>
<th>$K$</th>
<th>No. Multiplies per Update</th>
<th>For typical $m$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$m = 10$</td>
</tr>
<tr>
<td>0</td>
<td>$(3/2)(m^2 + 3m + 4)$</td>
<td>201</td>
</tr>
<tr>
<td>1</td>
<td>$3m^2 + 12m + 20$</td>
<td>440</td>
</tr>
<tr>
<td>2</td>
<td>$3/2(3m^2 + 15m + 32)$</td>
<td>723</td>
</tr>
<tr>
<td>3</td>
<td>$6(m^2 + 6m + 16)$</td>
<td>1056</td>
</tr>
</tbody>
</table>

an even smaller number of observation vectors [21], [26] thereby compensating for the higher per-update computational cost. This trade-off makes MW-OBE algorithms attractive in applications where per-point convergence an important factor.

The costs in Table II exploit symmetries in the computations. Additional reduction in cost is achieved by noting that $P_n$ may be evaluated in a total of $\tilde{K}(\tilde{K}^2 + \tilde{K}m + m^2)$ [24], [27] multiplications using $LU$ decomposition of $H_n^{-1}A_n$. This technique also simplifies the computation of $\kappa_n$.

G. Optimization strategies

The matrix $P_n^{-1}/\kappa_n$ defines the hyperellipsoid at time $n$ [see (4)]. The measure det($\kappa_n P_n$) is proportional to the square of the volume of the ellipsoid and is most often minimized as the OBE optimization criterion. Minimization of the trace of $\{\kappa_n P_n\}$ is also a meaningful measure of size (e.g., [5], [7]). Minimization of the parameter $\kappa_n$, first suggested in [28], had been controversial with respect to its interpretability [5] until the recent development of the QOBE algorithm [15], [16]. QOBE, which minimizes $\kappa_n$ in conjunction with a specific weighting strategy, provides interesting interpretations of this optimization process. Because of the relative algebraic simplicity of the algorithm, we use the QOBE-like approach of $\kappa_n$ minimization to develop a specific instance of the MW-OBE algorithm, MW-QOBE.

IV. MULTI-WEIGHT QUASI-OBE ALGORITHM

MW-QOBE is a MW-OBE algorithm based on minimizing the scalar $\kappa_n$ with respect to the present and past $K$ weights if the new data admit further reduction. We compare this novel approach to the QOBE (quasi-OBE) algorithm [15], [29]–[31]. We prove the uniqueness of the optimal solution for general $K \leq m$, and experimentally study the case of $K = 1$.

A. $\kappa_n$ minimization

Recent study of the QOBE algorithm has shown the merit of minimizing the scalar $\kappa_n$ [15], [16]. This simple yet efficient algorithm offers good convergence of the parameter estimator to the true parameter vector. When the prediction error $\varepsilon_{n|n-1}$ generated by the current parameter estimator $\theta_{n-1}$ falls outside the error constraint (1), QOBE generates a new parameter estimate $\theta_n$ which re-maps the prediction error to (exactly) one of the bounding hyperplanes $|\varepsilon_{n|n}| = \gamma_n$. QOBE was developed from an “OBE” point of view but the condition for data acceptance ($|\varepsilon_{n|n-1}| < \gamma_n$) and the process of mapping $\varepsilon_{n|n-1}$ into $\varepsilon_{n|n}$ can be “decoupled” from the ellipsoid, a departure from other OBE algorithms (and reason for the name “QOBE”). This independence from the ellipsoid makes QOBE particularly interesting in time-varying applications due to its robustness to a “true” parameter $\theta_n$ moving outside the ellipsoid. In this section we develop a specific MW-OBE algorithm the minimization of $\kappa_n$ with respect to the weight vector $\lambda_n$ as the optimization criterion.

The following theorem provides the optimal weight vector $\lambda_n$ at time $n$ which minimizes $\kappa_n$ over the present and past $K$ weight adjustments.

Theorem 1: The scalar $\kappa_n$ is minimized by the weight adjustments

$$\lambda_n = (G_{n|n-1}S_n\Lambda_n)^{-1}(\varepsilon_{n|n-1} - S_n\gamma_n)$$

where $S_n$ is a diagonal matrix with diagonal elements $\pm 1$.

Proof: In minimizing $\kappa_n$ (15), we encounter the term $\varepsilon_{n|n-1}^T H_n^{-1} \Lambda_n \varepsilon_{n|n-1}$, a quadratic expression in $\varepsilon_{n|n-1}$ which involves the inverse matrix $H_n^{-1} \Lambda_n = [\Lambda_n^{-1} + G_{n|n-1}]^{-1}$. Let $\lambda_n(i)$ represent the $i$th element of the vector $\lambda_n$ at time $n$. $\lambda_n(i)$ is equivalent to $\Lambda_n(i, i)$ since
\[ \Lambda_n = D(\lambda_n). \] Differentiating the term \( H_n^{-1} \Lambda_n \) with respect to an arbitrary (scalar) weight \( \lambda_n(i) \), we obtain
\[ \frac{\partial (H_n^{-1} \Lambda_n)}{\partial \lambda_n(i)} = \frac{\partial H_n^{-1}}{\partial \lambda_n(i)} \Lambda_n + H_n^{-1} \frac{\partial \Lambda_n}{\partial \lambda_n(i)} \]
\[ = H_n^{-1} \frac{\partial \Lambda_n}{\partial \lambda_n(i)} (H_n^{-1})^T. \] (24)
Equating (24) to the zero vector yields
\[ \varepsilon^T_{n|n-1} \frac{\partial (H_n^{-1} \Lambda_n)}{\partial \lambda_n(i)} \varepsilon_{n|n-1} = \varepsilon^T_{n|n-1} H_n^{-1}(*,i) [H_n^{-1}(*,i)]^T \varepsilon_{n|n-1} \] (25)
where the column vector \( H_n^{-1}(*,i) \) is the \( i \)th column of matrix \( H_n^{-1} \). By incorporating (25) in the differentiation of (15) with respect to \( \lambda_n(i) \) [20] we obtain
\[ \frac{\partial \kappa_n}{\partial \lambda_n} = \Gamma_n^2 - [D(H_n^T \varepsilon_{n|n-1})]^2, \] (26)
where \( \frac{\partial \kappa_n}{\partial \lambda_n} \) is the diagonal matrix whose \( i \)th diagonal element is \( \frac{\partial \kappa_n}{\partial \lambda_n(i)} \) and \( \Gamma_n = D(\gamma_n) \).
Using \( |A| \) to denote matrix with \((i,j)\) element \( |A(i,j)|\), the minimizing equation becomes
\[ |H_n^T \varepsilon_{n|n-1} - \gamma_n| = 0. \] (27)
To solve (27), we define the column vector of signs, \( \pm 1 \), \( s_n \equiv \text{sign}(H_n^T \varepsilon_{n|n-1}) \). Using
\[ |H_n^T \varepsilon_{n|n-1}| = (H_n^T \varepsilon_{n|n-1}) \circ s_n \] (28)
in (27), and multiplying on the left by \( H_n^T \), the optimal weight vector is found to be the solution in \( \lambda_n \) of
\[ \varepsilon_{n|n-1} = (\gamma_n \circ s_n) + G_{n|n-1} \Lambda_n (\gamma_n \circ s_n). \] (29)
After some manipulation, we obtain
\[ \lambda_n = \left[ G_{n|n-1}^{-1} (\varepsilon_{n|n-1} - \gamma_n \circ s_n) \right] \circ (g_n \circ s_n) \] (30)
where \( g_n \) denotes the column vector of the diagonal elements of \( \Gamma_n^{-1} \). If \( S_n \equiv D(s_n) \), then (30) can be also expressed as
\[ \lambda_n = (G_{n|n-1} S_n \Gamma_n)^{-1} (\varepsilon_{n|n-1} - S_n \gamma_n). \] (31)
We verify that the solution point is a minimum by demonstrating a positive determinant of the Hessian, the matrix whose \((i,j)\) element is
\[ \frac{\partial^2 \kappa_n}{\partial \lambda_n(i) \partial \lambda_n(j)} \] (32)
alternatively represented by the (symmetric) matrix whose \( i \)th column is
\[ \frac{\partial}{\partial \lambda_n(i)} \left( \frac{\partial \kappa_n}{\partial \lambda_n} \right). \] (33)
The determinant of the Hessian matrix
\[ \frac{\partial^2 \kappa_n}{\partial \lambda_n(i) \partial \lambda_n(j)} \] becomes (see Appendix A)
\[ \frac{2K+1}{\det(H_n)} \det(S_n \Gamma_n) \det(G_{n|n-1}). \] (34)
A necessary and sufficient condition for a non-negative Hessian determinant (34) is \( \det(H_n) \geq 0 \). This condition is satisfied with valid weight adjustments, made evident by expressing \( H_n \) in its \textit{a posteriori} representation (20).

**Remark:** Setting \( K = 0 \) reduces this expression to the optimal scalar weight in the QOBE algorithm [15], [16].

**Corollary 1:** Optimizing \( \kappa_n \) over several weights at each \( n \) results in a non-increasing sequence \( \{\kappa_n\} \).

**Proof:** Substituting (27) into (15) we obtain
\[ \kappa_n = -(\varepsilon_{n|n-1} - S_n \gamma_n)^T G_{n|n-1}^{-1} (\varepsilon_{n|n-1} - S_n \gamma_n) + \kappa_{n-1} \] (35)
where \( G_{n|n-1}^{-1} \) is positive semi-definite. Therefore, \( \kappa_n \) is a non-increasing function when evaluated at \( \lambda_n \) of (31).
Equation (35) provides an algebraic proof of Corollary 1 along with the amount of decrease in \( \kappa_n \) at each step. To simply prove that \( \kappa_n \leq \kappa_{n-1} \) [using weights (31)], we note that \( \kappa_n^{\lambda_n=0} = \kappa_{n-1} \) -.

**B. Existence and uniqueness of an optimal solution**

By pre-determining the absence of an optimal solution through a computationally inexpensive test, an OBE algorithm avoids laborious computations that ultimately yield zero weights (indicating no optimal solution). The QOBE algorithm is particularly attractive due to verify its simple test for innovation, namely \( |\varepsilon_{n|n-1}| > \gamma_n \).

In the previous section, the MW-QOBE algorithm optimal weights are derived as functions of a sign vector \( s_n \). This \((K + 1) \times 1\) vector is formed from a set of \( 2K+1 \) possible permutations of \( \pm 1 \) elements, with the constraint (see Section III) that the weights at time \( n \) be positive,
\[ q_{n,n} = \lambda_n + q_{n,n-1} > 0. \] (36)
An efficient test for new information in the MW-OBE observation matrix is not yet found. However, we report a necessary condition for the existence of a solution in Lemma 1. This necessary condition serves to reduce the solution search.
1) Existence: The following lemma represents the principal algebraic result to later derive a necessary (not sufficient) condition for existence of the weight adjustment vector \( \lambda_n \).

**Lemma 1:** Let \( A \) represent an \( n \times n \) symmetric positive definite matrix and \( b \) an \( n \times 1 \) vector with \( b(i) > 0 \), for all \( i \in [1, n] \). A sign vector \( s \) with corresponding diagonal matrix \( S = D(s) \) satisfies the vector inequality \( S(b - As) > 0 \) only if
\[
 s^T b > s^T As > 0.
\]

**Proof:** Multiply each side of the vector inequality by a vector of ones of appropriate dimension. \( \blacksquare \)

**Lemma 2:** A weight adjustment vector \( \lambda_n \) (31) satisfying (36) exists only for sign vectors such that
\[
 s_n^T \Gamma_n G_n^{-1}[\varepsilon_n|n-1| - 1] \Gamma_n s_n > 0.
\] (37)

**Proof:** We may assume without loss of generality that \( \lambda_n > 0 \). Let \( s_b = \text{sign}(\Gamma_n G_n^{-1}|\varepsilon_n|n-1|) \), \( S_b = \mathcal{D}(s_b) \), \( S = S_n S_b \), \( b = \Gamma_n G_n^{-1}|\varepsilon_n|n-1| \) and \( A = S_b \Gamma_n G_n^{-1}[\varepsilon_n] \Gamma_n S_b \). Apply Lemma 1 to refine the search for a valid sign vector, hence, a weight adjustments vector.

**Remark:** When re-optimizing a single past weight \((K = 1)\), the sign vector \( S_n \) is either \( S_n = I \), or
\[
 s_n = \text{sign}[b(1) - b(2)] \begin{bmatrix} 1 \\ -1 \end{bmatrix}
\] (38)

Equation (38) and the fact that \( S_n \neq -I \) are direct consequences of Lemma 2. \( \blacksquare \)

2) Uniqueness: Next we prove that at most one weight adjustment vector \( \lambda \) satisfies (31) with constraint (36).

**Lemma 3:** Let \( A \) represent an \( n \times n \) symmetric positive definite matrix and let \( u \) and \( v \) be two \( n \times 1 \) vectors with \( v(i) > 0 \), for all \( i \in [1, n] \). There exists at most one sign vector with corresponding diagonal matrix \( S = D(s) \) such that
\[
 SA(u - Sv) > 0.
\] (39)

**Proof:** See Appendix B. \( \blacksquare \)

**Theorem 2 (Uniqueness):** At most one weight adjustment vector \( \lambda_n \) solving (31) also satisfies (36).

**Proof:** Let \( A = G_n^{-1}|n-1|, u = \varepsilon_n|n-1| \) and \( v = \gamma_n \). Apply Lemma 3. \( \blacksquare \)

C. Incremental gain

The merit of MW-QOBE with respect to QOBE is its ability to further reduce both \( \kappa_n \) and the ellipsoid volume at each update. In this section we explore these improvements.

**Theorem 3:** Let \( \lambda_{n,K} \) and \( \lambda_{n,K-1} \) be the weights which optimize (minimize) \( \kappa_{n,K} \) and \( \kappa_{n,K-1} \) in light of \( K \) and \( K - 1 \) observations respectively. Then \( \kappa_{n,K}(\lambda_{n,K}) < \kappa_{n,K-1}(\lambda_{n,K-1}) \).

**Proof:** The energy matrix at time \( n \) when optimizing over \( K \) weights may be partitioned in the following manner:
\[
 G_{n|n-1,K} = \begin{bmatrix} g_{n,K} & g_{n,K}^T \\ g_{n,K} & G_{n|n-1,K-1} \end{bmatrix}
\] (40)

with \( g_{n,K} = x_{n,K}^T P_{n-1} x_{n-K} \) and \( g_{n,K}^T = x_{n,K}^T P_{n-1} X_{n,K-1} \).

Define the scalar \( \Delta_{n,K} = g_{n,K} - g_{n,K}^T G_{n|n-1,K-1}^{-1} g_{n,K} \) and \( g_{n,K}^T = x_{n,K}^T P_{n-1} X_{n,K-1} \).

This phenomenon is illustrated in Fig. 4 for the case \( K = 1 \). At each iteration, the MW-QOBE algorithm attempts to map the error vector \( \varepsilon_{n|n-1} \) to the unique \( S_n \gamma_n \) vector (one of \( 2^{K+1} \)) which satisfies condition (36) through the transformation (27). This condition imposes a stringent requirement on the acceptance of observation vectors and as a result provides a more selective screening process.

D. Geometric interpretation the \( \kappa_n \) minimization

Geometric interpretation of the MW-QOBE algorithm provides insight into its behavior. The QOBE algorithm (MW-QOBE with \( K = 0 \)) maps the absolute value of the \( a \ posteriori \) error, \( |\varepsilon_{n|n-1}| \), to the bound \( \gamma_n \) [16]. A similar behavior is indicated in the MW-QOBE algorithm by incorporating (27) into (16) to obtain
\[
 \gamma_n \circ \gamma_n - \varepsilon_{n|n} \circ \varepsilon_{n|n} = 0.
\] (41)

This reveals that the MW-QOBE algorithm maps the component-wise absolute value of \( a \ posteriori \) error vector \( \varepsilon_{n|n} \) to the error vector bound \( \gamma_n \) by requiring
\[
 \gamma_n = |\varepsilon_{n|n}|.
\] (42)
induces further decrease in per update algorithm is significantly worse than that than that of QOBE. However, these selective algorithms tend to incorporate so few data that, even with the additional burden at times of update, the net result when frequent non-updates are factored-in renders MW-OBE comparable to QOBE (remaining $O(\tau n)$). Further, there is empirical evidence that reoptimization may result in a reduction greater better than 10% in updates over conventional OBE optimization.

**F. Computational cost**

The basic computational costs are described in Section III-F. The initial data acceptance check remains the same as in the QOBE algorithm. If a QOBE optimal weight exists, the algorithm attempts an optimization over the past $K$ weights (36) which requires the computation of $G_{n|n-1}^{-1}$ ($K^3$ floating-point operations (flops)) as well as its multiplication by a vector ($K^2$ flops) (31). Note that the computation of $\det(G_{n|n-1})$ is not necessary to check inequality (36) since $\det(G_{n|n-1}) > 0$ and therefore does not change this inequality.

The computational cost is decreased with increased $K$ due the reduction in observations deemed innovative.

**V. ILLUSTRATIVE EXAMPLES**

**A. Decrease in ellipsoid volume**

To illustrate the effect of $K$ in MW-QOBE algorithms, consider the AR(2) system,

$$y_n = -0.10y_{n-1} - 0.56y_{n-2} + \varepsilon_{n*}$$

(44)

where $\varepsilon_{n*}$ is uniformly distributed over the interval $(-1, +1)$. We use MW-QOBE with $K = 0$ (QOBE) and MW-QOBE ($K = 1$) to identify 100 observations of this AR(2) system. The QOBE algorithm found 21 points useful in the optimization process compared with

---

**TABLE IV**

**THE MW-QOBE ALGORITHM.**

| I. Initialization: | Same as Table I. |
| II. Recursion: | Same as Table I |
| If $|\varepsilon_{n|n-1}| > \gamma_n$ | Form $X_n$ from present and select past $K$ data vectors along with corresponding $y_n$ and $\gamma_n$ vectors. |
| $\varepsilon_{n|n-1} = y_n - X_n^T \theta_{n-1}$ | $R_n = P_{n-1} X_n$ and $G_{0|n-1} = X_n^T R_n$ |
| $R_n = P_{n-1} X_n$ and $G_{0|n-1} = X_n^T R_n$ | If there exists a sign matrix $S_n$ (as in 31) such that $S_n G_{n|n-1}^{-1}(\varepsilon_{n|n-1} - S_n \gamma_n) > 0$, then $\lambda_n = (G_{0|n-1} S_n \gamma_n)^{-1}(\varepsilon_{n|n-1} - S_n \gamma_n)$ |
| Otherwise, next $n$. | |

Next $n$.
five used by MW-QOBE with $K = 1$. Figure 5 shows the ellipsoids generated by the two techniques at times $n = 8$ and $n = 13$ with corresponding polytopes, ellipsoid centers and true parameters. The MW-QOBE ellipsoids show improved alignment with the major axis of the polytope, and reduced volumes.

We note that QOBE does not focus on decreasing volume. In fact, it more-or-less “ignores” the ellipsoid altogether in the attempt to minimize $\kappa_n$ which inherently decreases the ellipsoid volume.

### B. Weight assignments

In this section, we illustrate the weight assignment mechanism in the MW-QOBE algorithms and their data screening behavior. Consider the AR(3) system,

$$y_n = 0.49y_{n-1} + 0.61y_{n-2} + 0.58y_{n-3} + \varepsilon_n,$$  

(45)

where $\varepsilon_n$ is biased and found in the interval $[-0.5, 1]$. Table V shows the first few weight assignments with a constant error bound $\gamma = 1$. As expected, the QOBE algorithm uses the largest number of observations in the optimization process, followed by MW-QOBE ($K = 1$) and MW-QOBE ($K = 2$). At time $n = 17$ ($K = 2$), the non-zero weight $[9.34 \times 10^{-4}, 1.904 \times 10^{-3}, 4.148 \times 10^{-3}]^T$ is assigned to observations previously ignored. Therefore, observation vectors $x_{15}$ and $x_{16}$ became relevant in light of the new observation vector $x_{17}$. We also observe at times $n = 15, 16$ and 17 ($K = 1$) the compounding effects of weight adjustment as consecutive observation blocks are used. Between times 15 and 17, the weight applied to observation vector $x_{15}$ increased from $6.242 \times 10^{-3}$ to $6.551 \times 10^{-3}$ and the one applied to $x_{16}$ decreased from $1.906 \times 10^{-3}$ to $5.443 \times 10^{-4}$. This example illustrates the importance of allowing constrained negative adjustments to past data vectors which may not convey as much information as previously computed in light of a current observation.

The MW-QOBE ($K = 2$) only used nine observation vectors in identifying a 500-point sequence output of (45), compared with the 29 and 57 needed by MW-QOBE ($K = 1$) and QOBE, respectively. Using this reduced number of observations, the MW-QOBE ($K = 2$) was able to identify system (45) in a comparable amount time to that required by smaller $K$, as seen in Fig. 6a. The associated volume and $\kappa_n$ plots are shown in Fig. 6b and 7b. The *a posteriori* error in Fig. 7a shows its mapping to the error bound (when an observation is accepted).

### VI. CONCLUSION

The MW-OBE algorithm has been presented and optimized in the framework of the QOBE algorithm. By jointly updating past weight assignments, the MW-OBE shows potential convergence and feasible set volume reduction benefits. Simulation studies have revealed enhanced ability to minimize $\kappa_n$ with increasing $K$. Improved volume reduction also results with increasing $K$. Reoptimization of weights results in an increased computational cost per update, but the overall per-point complexity remains at (or below) $O(m)$ with infrequent updates. The number of updates is decreased dramatically with increasing $K$ due to a conservative aggregate check for innovation.

### APPENDIX

#### A. Derivation of the Hessian matrix

**Proof:** The derivative of (31) with respect to a scalar weight $\lambda_n(i)$ is expressed as

$$\frac{\partial}{\partial \lambda_n(i)} \left( \frac{\partial \kappa_n}{\partial \lambda_n} \right) = -2 \left( H_n^{-T} \varepsilon_{n|n-1} \right) \circ \left[ \frac{\partial}{\partial \lambda_n(i)} \left( H_n^{-T} \varepsilon_{n|n-1} \right) \right]$$

(46)

which, by incorporating (24), becomes

$$\frac{\partial}{\partial \lambda_n(i)} \left( \frac{\partial \kappa_n}{\partial \lambda_n} \right) = -2 \left( H_n^{-T} \varepsilon_{n|n-1} \right) \circ \left[ -H_n^{-T} \frac{\partial H_n^{-T}}{\partial \lambda_n(i)} H_n^{-T} \varepsilon_{n|n-1} \right].$$

Evaluating (48) at any root, say $\lambda^*_n$, of the first derivative (31) and by substituting $H_n^{-T} \varepsilon_{n|n-1} = S_n \gamma_n$ into (46), we obtain the columns of the Hessian matrix

$$\frac{\partial}{\partial \lambda_n(i)} \left( \frac{\partial \kappa_n}{\partial \lambda_n} \right) \lambda^*_n = 2S_n \gamma_n \circ H_n^{-T} \frac{\partial H_n^{-T}}{\partial \lambda_n(i)} S_n \gamma_n$$

$$= 2S_n \gamma_n \circ G_{n|n-1} H_n^{-1} \frac{\partial \lambda_n}{\partial \lambda_n(i)} S_n \gamma_n.$$  

(49)

The determinant of the Hessian is then

$$\det \left[ \frac{\partial^2 \kappa_n}{\partial \lambda_n(1) \partial \lambda_n}, \ldots, \frac{\partial^2 \kappa_n}{\partial \lambda_n(K+1) \partial \lambda_n} \right]$$

$$= \frac{1}{\det (H_n)} \det^2 (S_n \Gamma_n) \det (G_{n|n-1}).$$

#### B. Proof of Lemma 3

**Proof:** Let $s_1 \neq s_2$ be sign vectors with corresponding diagonal matrices $S_1$ and $S_2$ both satisfying inequality (39). Without loss of generality, we assume that the mismatched elements of $S_1$ and $S_2$ are consecutively
arranged in the top left quadrants as $S_{1a}$ and $S_{2a}$, since we may re-order the basis and preserve the positive definitiveness of $A$. The partitioned matrices (including $A$ appropriately partitioned) are

$$
S_1 = \begin{bmatrix}
S_{1a} & 0 \\
0 & S_{1b}
\end{bmatrix}, \quad S_2 = \begin{bmatrix}
-S_{1a} & 0 \\
0 & S_{1b}
\end{bmatrix},
$$

$$
A = \begin{bmatrix}
A_a & A_c \\
A_c & A_b
\end{bmatrix}.
$$

We now add the inequalities corresponding to $S_1$ and $S_2$,

$$
0 < (S_1 + S_2)Au - (S_1AS_1 + S_2AS_2)v
$$

and incorporate (50), to obtain the element-wise strictly negative matrix

$$
0 < \begin{bmatrix}
0 & -S_{1a} \\
S_{1b} & 0
\end{bmatrix} \begin{bmatrix}
A_a & A_c \\
A_c & A_b
\end{bmatrix} \begin{bmatrix}
u_a \\
v_b
\end{bmatrix}
$$

or,

$$
0 < \frac{-S_{1a}A_aS_{1a}v_a}{S_{1b}(A_a^Tu_a + A_bu_b) - S_{1a}A_bS_{1b}v_b}
$$

(51)

where $[u_a | u_b]^T$ and $[v_a | v_b]^T$ are appropriately partitioned vectors $u$ and $v$. Multiplying each side of the top partition inequality by the vector $v_a > 0$ maintains the inequality, therefore

$$
0 < -v_a^T S_{1a}A_aS_{1a}v_a,
$$

a contradiction since $A_a$ is positive definite. Hence a sign vector $s$ satisfying (39) is unique.

REFERENCES


(a) Time \( n = 8 \).

(b) Time \( n = 13 \).

Fig. 5. OBE ellipsoids resulting from the system identification of AR(2) system \( y_n = -0.10y_{n-1} - 0.56y_{n-2} + \varepsilon_n \) by QOBE (dashed line) and MW-QOBE (solid line, \( K=1 \)) at times \( n = 8 \) and 13 (limited data example). The star (\( * \)) represents the “true” parameter and the circles (\( \circ \)) the central estimators (superimposed). The underlying polytopes (exact feasible sets) are also shown.
TABLE V
FIRST 30 WEIGHTS ASSIGNED BY QOBE AND MW-QOBE (K = 1, 2) ALGORITHMS IN THE IDENTIFICATION OF THE AR(3) SYSTEM y_n = 0.49y_{n−1} + 0.61y_{n−2} + 0.58y_{n−3} + \varepsilon_n*, where the “true” measurement error sequence \{\varepsilon_n*\} is uniformly distributed over the interval (−1, +1). The QOBE and MW-QOBE (K = 1, 2) algorithms selected 32, 10, 3 observations, respectively, from a total of 100. The “-” (dash) signifies a weight of zero at the indicated time “n”.

<table>
<thead>
<tr>
<th>n</th>
<th>QOBE</th>
<th>MW-QOBE (K=1)</th>
<th>MW-QOBE (K=2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(q_{n,n})</td>
<td>(q_{n,n})</td>
<td>(q_{n,n})</td>
</tr>
<tr>
<td>5</td>
<td>1.837</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>6.203</td>
<td>12.002</td>
<td>7.117</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>12.002</td>
</tr>
<tr>
<td>9</td>
<td>50.115</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>15.073</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>16.917</td>
<td>62.422</td>
<td>25.913</td>
</tr>
<tr>
<td>16</td>
<td>20.392</td>
<td>19.061</td>
<td>65.508</td>
</tr>
<tr>
<td>18</td>
<td>2.025</td>
<td>-</td>
<td>21.261</td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>0.624</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>1.798</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>-</td>
<td>15.458</td>
<td>10.908</td>
</tr>
<tr>
<td>26</td>
<td>-</td>
<td>15.458</td>
<td>-</td>
</tr>
<tr>
<td>28</td>
<td>-</td>
<td>13.325</td>
<td>1.661</td>
</tr>
<tr>
<td>29</td>
<td>1.546</td>
<td>-</td>
<td>13.325</td>
</tr>
<tr>
<td>30</td>
<td>0.475</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>32</td>
<td>0.612</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>34</td>
<td>0.059</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>36</td>
<td>0.034</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>39</td>
<td>0.357</td>
<td>3.191</td>
<td>4.099</td>
</tr>
<tr>
<td>40</td>
<td>0.515</td>
<td>1.301</td>
<td>3.764</td>
</tr>
<tr>
<td>41</td>
<td>-</td>
<td>-</td>
<td>1.301</td>
</tr>
<tr>
<td>44</td>
<td>2.757</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>46</td>
<td>3.933</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>47</td>
<td>0.685</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>54</td>
<td>0.009</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>57</td>
<td>0.146</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>61</td>
<td>0.219</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>62</td>
<td>0.200</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Fig. 6. Convergence of the parameter estimator \(\theta_n(2)\) to the “true” parameter \(\theta_n(2) = 0.61\) and associated ellipsoid volume in the system identification of \(y_n = 0.49y_{n−1} + 0.61y_{n−2} + 0.58y_{n−3} + \varepsilon_n*\) by QOBE and MW-QOBE (K=1,2).
Fig. 7. $\kappa_n$ and a posteriori error $\varepsilon_n|n$ associated with the identification of the system of Fig. 6. A constant error bound value of 1 was utilized.