51st IEEE Conference on Decision and Control, Maui, HI, December 2012

Moving Horizon Estimation for Staged QP Problems

Eric Chu, Arezou Keshavarz, Dimitry Gorinevsky, and Stephen Boyd

***Abstract*— This paper considers moving horizon estimation (MHE) approach to solution of staged quadratic programming (QP) problems. Using an insight into the constrained solution structure for the growing horizon, we develop a very accurate iterative update of the arrival cost in the MHE solution. The update uses a quadratic approximation of the arrival cost and information about the previously active or inactive constraints. In the absence of constraints, the update is the familiar Kalman filter in information form. In the presence of the constraints, the update requires solving a sequence of linear systems with varying size. The proposed MHE update provides very good performance in numerical examples. This includes problems with ℓ_1 regularization where optimal estimation allows us to perform online segmentation of streaming data.**

I. Introduction

A. Problem statement

Given a series of quadratic forms $g_0(u)$, $g_t(u,v)$, where $u, v \in \mathbf{R}^n$ and t is an integer index, consider a series of time staged quadratic programming (QP) optimization problems for increasing T

$$\begin{array}{ll} \text{minimize} & g_0(z_0) + \sum_{t=1}^T g_t(z_{t-1}, z_t) \\ \text{subject to} & F^{\rm eq} z_t = G^{\rm eq} z_{t-1} + h^{\rm eq} \\ & F^{\rm ineq} z_t \preceq G^{\rm ineq} z_{t-1} + h^{\rm ineq} \\ & t = 1, \ldots, T, \end{array} \quad (1)$$

where $z_t \in \mathbf{R}^n$ are decision variables; the matrices $F^{\rm eq}$, $F^{\rm ineq}$ have sizes compatible with z_t and the vectors $h^{\rm eq}$, $h^{\rm ineq}$; and $\preceq$ denotes component-wise $\leq$ for two vectors. The optimization interval $t \in [1,\ T]$ is further called *full horizon*. Our goal is to provide a sequence of optimal solutions $Z^*(T) = (z^\star_{0|T}, \ldots, z^\star_{t|T})$ as T increases.

The stage costs g_t and the initial cost function, g_0, are convex quadratic functions that can be written in the form

$$g_t(u,v) = \begin{bmatrix} u \\ v \end{bmatrix}^T \begin{bmatrix} R_t & Q_t \\ Q_t^T & M_t \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} - 2 \begin{bmatrix} s_t \\ r_t \end{bmatrix}^T \begin{bmatrix} u \\ v \end{bmatrix}, \quad (2)$$

$$g_0(v) = v^T P_0 v - 2 q_0^T v, \quad (3)$$

where matrices R_t, Q_t, M_t, $P_0 \in \mathbf{R}^{n,n}$ provide the convexity.

The motivation for considering the QP formulation and nonlinear estimation problem examples that lead to such formulation are presented below. In the absence of constraints, the time staged QP is a slight generalization of the optimal linear Gaussian estimation problem. In linear Gaussian estimation

$$g_t(z_{t-1}, z_t) = \frac{1}{2}||y_t - C_t z_t||^2_{\Phi_t} + \frac{1}{2}||z_t - A_t z_{t-1}||^2_{\Psi_t}, \quad (4)$$

The authors are with the Information Systems Laboratory, Electrical Engineering Department, Stanford University, Stanford, CA 94305-9510, USA. Email: {echu508,arezou,gorin,boyd}@stanford.edu.

where y_t is the observed data; z_t is the estimated state; A_t and C_t are the state update and the observation matrices, respectively; we use the notation $\|x\|_Q^2 = x^T Q x$, and Φ_t, Ψ_t are the inverse covariance matrices of the observation and process noise, respectively. As T grows, the solution for problem (4) can be recursively computed using the well-known Kalman filter approach.

This paper pursues a moving horizon estimation (MHE) approach to solve the time-staged QP (1). For each T the MHE solves the problem

$$\begin{array}{ll} \text{minimize} & V_{T-N}(z_{T-N}) + \sum_{t=T-N+1}^T g_t(z_{t-1}, z_t) \\ \text{subject to} & F^{\rm eq} z_t = G^{\rm eq} z_{t-1} + h^{\rm eq} \\ & F^{\rm ineq} z_t \preceq G^{\rm ineq} z_{t-1} + h^{\rm ineq} \\ & t = T-N+1, \ldots, T, \end{array} \quad (5)$$

where N is fixed, and the time interval $t \in [T-N+1,\ T]$ is called the receding horizon of length N. The arrival cost $V_{T-N}(z_{T-N})$ in (5) is computed through a recursive update. The arrival cost update is the key contribution of this paper and is described below. Unlike the recursive Kalman filter solution, which is exact, the MHE solution is approximate, because the arrival cost update is approximate. This paper proposes a very accurate approximation of the arrival cost.

B. Motivation

Time-staged QP problems of the form (1) arise in many nonlinear estimation applications, such as robotics, image processing, process monitoring, navigation, and other filtering and trend estimation applications. This QP formulation includes state and observation constraints, which makes it more versatile than the standard Kalman filter.

An important source of the time-staged QP formulations are problems related to compressed sensing or robust estimation. Such problem impose penalty functions to shape the error residuals. Common examples include the Huber and ℓ_1 penalty function. One well-known example is total-variation denoising in which an ℓ_1 penalty is used to promote sparsity in labeling data segments. We consider the total-variation denoising example in Section III-B of this paper.

Time-staged QPs also arise in order-restricted inference problems and bioequivalence problems in drug testing.

For increasing T, these problems eventually require truncating the size of the QP problem solved. The MHE offers a systematic approach to such streaming data processing.

A brief discussion of some earlier work on MHE is presented in the next subsection. It is worth noting that many (perhaps even most) examples in the published MHE work consider QP formulations that can be described by (1).

C. Previous work

Some of the important earlier work on MHE includes [1], [2], [3], [4], where the MHE problems and solutions are established as the (estimation) dual to receding horizon control, or *model predictive control* (MPC). In [5] and later work, the stability of MPC is established by introducing a quadratic approximation for the cost-to-go at the terminal state of the receding horizon. As a dual problem, the MHE ability to track the full horizon solution (stability) requires adding a quadratic arrival cost penalty at the beginning of the receding horizon. For MPC, this line of reasoning is well discussed in [6], where further references can be found.

For MHE, a quadratic approximation of the arrival cost leads to an *extended Kalman filter* (EKF) update for the arrival cost. Such an update is used in much of the existing work in MHE. In the presence of constraints, however, the quadratic approximation of the arrival cost used in EKF might be inaccurate. This is especially true if the inequality constraints switch between being active and inactive at the optimal solution. As a result, tuning the EKF update in MHE is more of an art than science. It is well recognized that improperly tuned arrival cost update can render the MHE update unstable [7].

This paper proposes a new approach to approximating the arrival cost that is aware of the constraints changing from active to inactive as the optimization horizon moves. The earlier paper [8] discussed a special case of the MHE approach proposed in this paper in application to isotonic regression estimation. The MHE arrival cost update in [8] uses knowledge about the constraints being active or inactive based on the previous MHE update. This paper develops an extension of this idea to a general class of QP-representable MHE problems.

D. Contribution of the paper

This paper establishes a class of staged QP problems (1)–(3). This class includes many MHE problems in particular the problems with ℓ_1 regularization where the optimal MHE estimation allows us to perform on-line segmentation of streaming data. Two examples of such problems are considered in Section III.

The contribution of the paper is twofold: first, we introduce the approximation hypothesis that the optimal solution of a time-staged QP problem deep inside the horizon (*i.e.*, far back enough in the filter history) is insensitive to the last stage of the problem. In particular, the active set of the inequality constraints becomes fixed after a particular time as the problem horizon grows. This hypothesis holds for most problems in practice.

Second, we introduce an iterative update of the arrival cost in the MHE solution. The iterative update uses a quadratic approximation of the arrival cost and active/inactive constraint information from the previous step. In the absence of the constraints, the update becomes the familiar Kalman filter in information form. In the presence of the constraints, the update requires solving linear systems of varying size.

This paper demonstrates that our approximation hypothesis holds well in practice and the proposed arrival cost update leads to good estimates in our numerical examples. A theoretical proof of our approximation hypothesis is not included in this paper; note that stability proofs for MPC appear well after MPC's adoption.

II. Moving horizon estimation

A. Exact arrival cost function

Since minimizing the full problem is equivalent to minimizing over a partially minimized problem, the full horizon problem (1)–(3) can be presented in the form (5). This requires using an arrival cost $V_{T-N}(z_{T-N}) = V^\star_{T-N}(z_{T-N})$ obtained by partial minimization with respect to the decision variables $z_0, z_1, \ldots, z_{T-N-1}$.

$$\begin{array}{ll} V^\star_{T-N}(z_{T-N}) = & \\ \min\limits_{z_0,z_1,\ldots,z_{T-N-1}} & \left(g_0(z_0) + \sum_{t=1}^{t=T-N} g_t(z_{t-1}, z_t)\right) \\ \text{subject to} & F^{\mathrm{eq}} z_t = G^{\mathrm{eq}} z_{t-1} + h^{\mathrm{eq}} \\ & F^{\mathrm{ineq}} z_t \preceq G^{\mathrm{ineq}} z_{t-1} + h^{\mathrm{ineq}} \\ & t = 1, \ldots, T-N \end{array} \tag{6}$$

We will further call $V^\star_{T-N}(z_{T-N})$ in (6) the exact arrival cost at $T-N$ to distinguish it from the approximate arrival cost introduced below.

A recursive update for the exact arrival cost from $t-1$ to t can be obtained by extending the partial minimization to the next decision variable z_t. This yields the Bellman equation

$$\begin{array}{rl} V^\star_t(u) = \min\limits_v & V^\star_{t-1}(v) + g_t(u, v) \\ \text{subject to} & F^{\mathrm{eq}} v = G^{\mathrm{eq}} v + h^{\mathrm{eq}} \\ & F^{\mathrm{ineq}} v \preceq G^{\mathrm{ineq}} v + h^{\mathrm{ineq}}, \end{array} \tag{7}$$

where u corresponds to z_t and v to z_{t-1}.

B. Arrival cost and active constraints

Because of the inequality constraints in (6), the exact arrival cost $V^\star_t(u)$ is a complicated piecewise quadratic function of its argument. Transitions between the quadratic segments corresponds to the inequality constraints in (1) switching between being active and inactive. Calculating the arrival cost function in (6) involves solving an inequality constrained QP in the decision variables $z_0, z_1, \ldots, z_{T-N-1}$. As the number $T-N$ of decision variables grows, in most cases we can no longer effectively calculate $V^\star_{T-N}(u)$ in (6) for all $u \in \mathbf{R}^n$.

The proposed approach is to use quadratic approximation of the arrival cost in the vicinity of the optimal solution $Z^*(T) = (z^\star_{0|T}, \ldots, z^\star_{t|T})$. If $Z^*(T)$ were known ahead of time, then we could replace the affine inequality constraints in (6) with equality constraints whenever we know that the constraint is tight for $Z^\star(T)$. When the inequality constraint is not tight for the solution $Z^\star(T)$, we could drop the constraint.

For the constraints in (6), consider the index set $\mathcal{A}_{t|T}$ of the inequality constraints that are active at step t for the solution $Z^\star(T)$ of the full-horizon problem with the horizon T. The active constraint set $\mathcal{A}_{t|T}$ can be described as

$$\mathcal{A}_{t|T} = \{i \mid F_i^{\text{ineq}} z_t^\star = G_i^{\text{ineq}} z_{t-1}^\star + h_i^{\text{ineq}}\}, \qquad (8)$$

for $t = 1, \ldots, T$. Note that the the set of active constraints $\mathcal{A}_{t|T}$ depends on the horizon T. As T grows, the set $\mathcal{A}_{t|T}$ of the constraints active at time step t may change.

By replacing the constrains in (6) with the affine equality constraints, the arrival cost becomes a solution of equality constrained QP - a quadratic function. Unfortunately, finding the set of active constraints $\mathcal{A}_{t|T}$ exactly requires solving the full-horizon inequality constrained QP in the first place.

We will next discuss how a quadratic approximation of the arrival cost function and the approximate set of the active constraints $\mathcal{A}_{t|T}$ can be computed recursively without the need to solve the full horizon problem.

C. Approximation hypothesis

The MHE update relies on representing the full horizon problem (1) in the form (5), where N is fixed and the horizon T is growing. In what follows, we assume that N is fixed and $T > N$. For each T, the arrival cost is computed based on the MHE solution and the arrival cost that have been obtained at the previous step, for $T - 1$. After that, the new MHE solution to the N-step QP problem is computed.

The proposed MHE update is based on the following approximation hypothesis: for N large enough and $t \le T - N$, $\mathcal{A}_{t|T} = \mathcal{A}_{t|T+d}$ for any $d > 0$. This approximation implies that, as T grows, the active constraint set for $t \le T - N$ does not change and that $\mathcal{A}_{t|T} = \mathcal{A}_t$, where $\mathcal{A}_t$ is independent of T for T large enough.

Assuming that $\mathcal{A}_t$ is known, the combined equality and inequality constraints in (6) for $t = 1, \ldots, T - N$ can be changed to the equality constraints

$$F_t z_t = G_t z_{t-1} + h_t, \qquad (9)$$

where the matrices F_t, G_t, and h_t are defined as

$$F_t = \begin{bmatrix} F^{\text{eq}} \\ F_{\mathcal{A}_t}^{\text{ineq}} \end{bmatrix}, \; G_t = \begin{bmatrix} G^{\text{eq}} \\ G_{\mathcal{A}_t}^{\text{ineq}} \end{bmatrix}, \; h_t = \begin{bmatrix} h^{\text{eq}} \\ h_{\mathcal{A}_t}^{\text{ineq}} \end{bmatrix}, \qquad (10)$$

and $F_{\mathcal{A}_t}$ denotes the sub-matrix of F formed from the rows F_i for all $i \in \mathcal{A}_t$.

The approximation hypothesis is a stylized fact that is not exactly true, but appears to hold in practice for reasonably large N. As is shown below, it allows us to obtain very accurate MHE approximations of the exact full-horizon solution.

D. Update of the approximate arrival cost

Assuming that the introduced approximation hypothesis holds and $\mathcal{A}_t$ is known, the approximate arrival cost V_t can be found via an approximate Bellman iteration. By replacing the constraints in (7) with the equality constraints (9), we get the approximate Bellman update as

$$\begin{aligned} V_t(v) = \min\nolimits_u \quad & V_{t-1}(u) + g_t(u, v) \\ \text{subject to} \quad & F_{t-1} v = G_{t-1} u + h_{t-1} \end{aligned} \qquad (11)$$

At each horizon T, the proposed MHE approach solves the problem (5), where the (approximate) arrival cost V_{T-N} is updated according to (11). The arrival cost update (11) depends on the active constraint set $\mathcal{A}_{T-N}$ through (10). At each horizon T, we take $\mathcal{A}_{T-N} = \mathcal{A}_{T-N|T} = \mathcal{A}_{T-N|T-1}$, the active set for the solution that was computed at horizon $T - 1$. Such update of the active constraint set follows the approximation hypothesis described in Section II-C.

Unlike the exact Bellman update (7) that has inequality constraints, the approximate update (11) has affine equality constraints only. Starting with $V_0(v) = g_0(v)$, the approximate arrival cost function obtained in such update is always quadratic and can be presented in the form

$$V_t(v) = v^T P_t v - 2 q_t^T v + \alpha, \qquad (12)$$

where matrix P_t and vector q_t fully define the approximate arrival cost; the scalar constant α has no impact on the optimization problem.

The approximate arrival cost update (11) can be formulated as an update for P_t, q_t in (12). Consider the partial minimization problem for the approximate Bellman iteration (11), where V_{t-1} is given by (12) and $g_t(u, v)$ is given by (2). The Karush-Kuhn-Tucker (KKT) conditions for optimality yield

$$\begin{bmatrix} P_{t-1} + R_t & Q_t & F_t^T \\ Q_t^T & M_t & -G_t^T \\ F_t & -G_t & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ \mu \end{bmatrix} = \begin{bmatrix} q_{t-1} + s_t \\ r_t \\ -h_t \end{bmatrix},$$

Eliminating the variables u and μ from these conditions yields a quadratic form in v that describes $V_t(v)$ (12). The expressions for P_t and q_t can be extracted from this quadratic form. The obtained recursive update for P_t and q_t has the form

$$P_t = M_t - \begin{bmatrix} Q_t^T \\ F_t \end{bmatrix}^T \begin{bmatrix} P_{t-1} + R_t & -G_t^T \\ -G_t & 0 \end{bmatrix}^{-1} \begin{bmatrix} Q_t^T \\ F_t \end{bmatrix}$$

$$q_t = r_t - \begin{bmatrix} Q_t^T \\ F_t \end{bmatrix}^T \begin{bmatrix} P_{t-1} + R_t & -G_t^T \\ -G_t & 0 \end{bmatrix}^{-1} \begin{bmatrix} q_{t-1} + s_t \\ -h_t \end{bmatrix}.$$

This, along with (10) and the active constraint set $\mathcal{A}_{T-N}$ update, fully specifies the arrival cost update (11). In contrast to the familiar Kalman filter, the matrices updated and inverted in our equations have varying size. The size of the matrices F_t, G_t, and h_t changes depending on the number of the inequality constrains that are active at time t.

E. MHE update summary

To recap, the proposed MHE approach approximately solves the time-staged QP problem (1)–(3). At each horizon T, it provides a solution $Z^{\text{mhe}}(T)$ that approximates the optimal solution $Z^\star(T)$. For $T \le N$, we solve the full problem (1). For $T > N$, the N-step QP problem (5) is solved with the approximate arrival cost function $V_{T-N}(z_{T-N})$ of the form (12), where matrices P_{t-N} and q_{t-N} are defined recursively as described in Section II-D. The solution to the QP problem (5) can be obtained using any suitable QP solver.

For $T > N$ the N-step solution to (5) defines the last N decision vectors $z_{t|T}^{\text{mhe}}$ in $Z^{\text{mhe}}(T) =$

$(z^{\rm mhe}_{0|T}, z^{\rm mhe}_{1|T}, \ldots, z^{\rm mhe}_{T|T})$. The decision vectors for $t < T-N$ are taken from $Z^{\rm mhe}(T-1)$ such that $z^{\rm mhe}_{t|T} = z^{\rm mhe}_{t|T-1}$.

III. Examples

A. Kalman filter

Consider a linear Gaussian estimation problem for a dynamical system

$$\begin{aligned} x_{t+1} &= A_t x_t + w_t \\ y_t &= C_t x_t + v_t, \end{aligned}$$

where x_t is the time-dependent state vector, y_t is the observation vector, and A_t and C_t are matrices of appropriate sizes. The noise vectors w_t and v_t are assumed to be Gaussian with distribution $\mathcal{N}(0, \Psi_t^{-1})$ and $\mathcal{N}(0, \Phi_t^{-1})$, respectively. The initial state x_0 is assumed to have distribution $\mathcal{N}(P^{-1}q, P^{-1})$.

The problem of maximum likelihood estimation of the state z_t, given the observations y_t has the form (1) with no constraints, where $g_t(z_{t-1}, z_t)$ is described by (4), and can be presented in the form (2) with

$$M_t = \Psi_t + C_t^T \Phi_t C_t, \qquad Q_t = -\Psi_t^T A_t,$$
$$R_t = A_t^T \Psi_t A_t, \qquad r_t = C_t^T \Phi_t y_t, \qquad s_t = 0.$$

In this example, there are no constraints and all approximations are exact. The update (7) for the (exact) arrival cost function $V_t^\star$ can be performed by the recursive iteration

$$\begin{aligned} P_t &= \Psi_t + C_t^T \Phi_t C_t - \Psi_t A_t (P_{t-1} + A_t^T \Psi_t A_t)^{-1} A_t^T \Psi_t \\ q_t &= C_t^T \Phi_t y_t + \Psi_t A_t (P_{t-1} + A_t^T \Psi_t A_t)^{-1} q_{t-1}. \end{aligned}$$

In this case, our approximation hypothesis holds exactly (since there are no constraints), and any choice of N in the MHE formulation (5) with the exact arrival cost will yield the same solution. These recursive estimation equations provide information form of the Kalman filter, where P_t is the inverse covariance matrix (information matrix). The standard Kalman filter corresponds to the update with $N = 0$, and provides the last point of the optimal full-horizon solution as $z^\star_{T|T} = P_T^{-1} q_T$.

B. Total variation denoising

Consider the estimation problem

$$\begin{array}{ll} \text{minimize} & x_0^T P x_0 - 2q^T x_0 + \\ & \sum_{t=1}^T \lambda \|x_t - x_{t-1}\|_1 + \\ & \sum_{t=0}^T \|y_t - C x_t\|_\Phi^2, \end{array} \qquad (13)$$

where $x_t \in \mathbf{R}^n$, $t = 0, \ldots, T$ are the decision variables, $y_t \in \mathbf{R}^m$, $t = 0, \ldots, T$ are the raw data, $C \in \mathbf{R}^{m\times n}$ is the observation matrix, and $\Phi \in \mathbf{R}^{m\times m}$ is a positive definite weight matrix. The ℓ_1-norm is used to induce sparsity of changes in the solution x_t and force segmentation of the solution into several intervals with constant x_t. Problem (13) is known as total-variation denoising and is important for many applications.

To implement the MHE update of Section II-E, we formulate an equivalent problem

$$\begin{array}{ll} \text{minimize} & x_0^T P x_0 - 2q^T x_0 + \sum_{t=1}^T \lambda \mathbf{1}^T a_t + \\ & \sum_{t=0}^T (y_t - C x_t)^T \Phi (y_t - C x_t) \\ \text{subject to} & -a_t \preceq x_t - x_{t-1} \\ & \quad a_t \succeq x_t - x_{t-1} \\ & \quad t = 1, \ldots, T. \end{array}$$

Using decision variable vectors $z_t = (x_t, a_t)$, we can express this as the time-staged QP problem

$$\begin{array}{ll} \text{minimize} & z_0^T P_0 z_0 - 2q_0^T z_0 + \sum_{t=1}^T g_t(z_{t-1}, z_t) \\ \text{subject to} & F^{\rm ineq} z_t \preceq G^{\rm ineq} z_{t-1}, \\ & t = 1, \ldots, T, \end{array}$$

where

$$P_0 = \begin{bmatrix} P + C^T \Phi C & 0 \\ 0 & 0 \end{bmatrix}, \qquad q_0 = \begin{bmatrix} q + C^T \Phi y_0 \\ 0 \end{bmatrix},$$

$$F^{\rm ineq} = \begin{bmatrix} -I & -I \\ I & -I \end{bmatrix}, \qquad G^{\rm ineq} = \begin{bmatrix} -I & 0 \\ I & 0 \end{bmatrix}.$$

The cost function g_t has the form (2) with

$$M_t = \begin{bmatrix} C^T \Phi C & 0 \\ 0 & 0 \end{bmatrix}, \qquad Q_t = R_t = 0,$$

$$r_t = \begin{bmatrix} C^T \Phi y_t \\ -(\lambda/2)\mathbf{1} \end{bmatrix}, \qquad s_t = 0.$$

Using this representation, we solve an instance of problem (13) with scalar state x_t and scalar data y_t, with $C = 1$, $\Phi = 1$, $P = 0$, $q = 0$, and $T = 200$. We use the moving horizon of $N = 20$ and $\lambda = 20$. The time series y_t, $t = 0, \ldots, T$ was obtained by adding random noise to a piecewise constant signal. We implement the MHE update of Section II-E to solve this total variation denoising problem.

Figure 1 shows the raw data y_t as dots and the state x_t in the full-horizon solution $Z^\star(T)$ for $T = 200$ as a piece-wise constant dashed line. The uneven solid line in the plot corresponds to the x-component of the final state $z^{\rm mhe}_{t|t}$ estimated at each step t of the MHE update. In this plot, the MHE solution overlaps with the full horizon solution, we have $z^{\rm mhe}_{t|t} = z^\star_{t|t}$ within the accuracy of the computations. The vertical lines in Figure 1 delineate change points where some of the constrains change between active and inactive.

C. ℓ_1 trend filtering

We now consider the ℓ_1 trend filtering problem described in [9]. The problem can be formulated as

$$\begin{array}{ll} \text{minimize} & x_0^T P x_0 - 2q^T x_0 + \\ & \sum_{t=2}^T \lambda |x_t - 2x_{t-1} + x_{t-2}| + \\ & \sum_{t=0}^T (y_t - x_t)^2, \end{array} \qquad (14)$$

where x_t is a scalar decision variable and time series y_t, $t = 0, \ldots, T$ represents the data to be filtered. This problem places an ℓ_1 penalty on the second difference $x_t - 2x_{t-1} + x_{t-2}$ to make it sparse. The solutions of (14) follow a piecewise linear trend with sparse kink points. The segmentation of the solution into affine intervals (straight

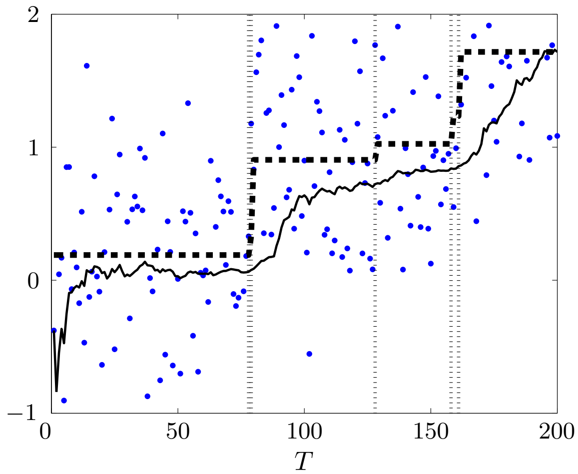

Fig. 1. MHE filtering results for total variation denoising

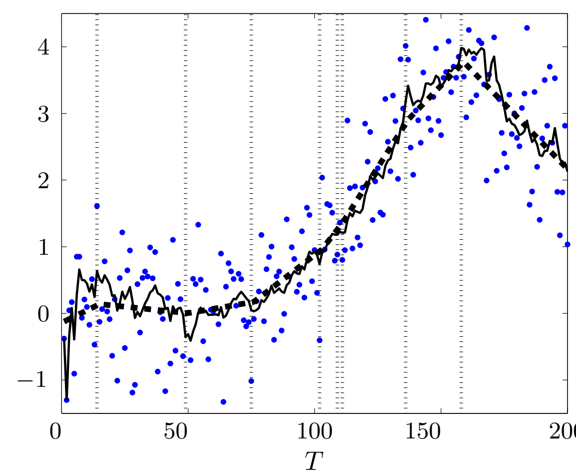

Fig. 2. MHE filtering results for ℓ_1 filter.

line segments) facilitates data interpretation, compression, and forecasting.

Similar to the previous subsection, the problem can be transformed into a time-staged QP. An equivalent form of problem (14) is

$$\begin{array}{ll} \text{minimize} & x_0^T P x_0 - 2q^T x_0 + \sum_{t=2}^T \lambda a_t + \\ & \sum_{t=0}^T (y_t - x_t)^2 \\ \text{subject to} & \quad |x_t - 2x_{t-1} + x_{t-2}| \leq a_t \\ & \quad t = 2, \ldots, T. \end{array}$$

With the decision variable $z_t = (x_t, x_{t-1}, a_t)$, we can express the problem as

$$\begin{array}{ll} \text{minimize} & z_0^T P_0 z_0 - 2q_0^T z_0 + \sum_{t=1}^T g_t(z_{t-1}, z_t) \\ \text{subject to} & F^{\text{ineq}} z_t \preceq G^{\text{ineq}} z_{t-1} \\ & F^{\text{eq}} z_t = G^{\text{eq}} z_{t-1}, \quad t = 2, \ldots, T, \end{array}$$

where

$$P_0 = \begin{bmatrix} P+1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad q_0 = \begin{bmatrix} q + y_0 \\ 0 \\ 0 \end{bmatrix}$$

$$F^{\text{ineq}} = \begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix}, \qquad G^{\text{ineq}} = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix},$$

$$F^{\text{eq}} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}, \qquad G^{\text{eq}} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

The representation (2) for the cost function g_t is given by

$$M_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad Q_t = R_t = 0,$$

$$r_t = \begin{bmatrix} y_t \\ 0 \\ -(\lambda/2) \end{bmatrix}, \qquad s_t = 0.$$

Figure 2 shows the result of applying the MHE algorithm of Section II-E to problem (14) with $\lambda = 50$, $P = 0$, $q = 0$. The data y_t, $t = 0, \ldots, T$ is shown as dots. The solid line shows the optimal filtering solution $z^\star_{t|t}$. In this case, the MHE solution $z^{\text{mhe}}_{t|t} = z^\star_{t|t}$ (within the accuracy of the computations) for all $t = 0, \ldots, 200$. The vertical lines in Figure 2 delineate the change points.

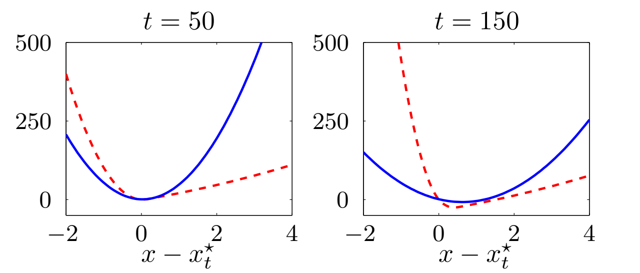

Fig. 3. The true $V^\star_t(x)$ (dashed) and the approximate $V_t(x)$ (solid) arrival costs.

IV. Accuracy of Approximation

A. Accuracy of arrival cost approximation

Consider the total-variation denoising problem (13) discussed as Example 2 in Section III-B. This is a one-dimensional example with the scalar state x_t, scalar observation y_t, $C = 1$, $\Phi = 1$, and $\lambda = 20$. In this example, an MHE update with window of $N = 20$ is applied to data time series of the length $T = 200$. We calculate the exact arrival cost function $V^\star_t(\cdot)$, $t = 0, \ldots, 200$ by gridding, and compare it to the approximate arrival cost functions, obtained using our method. Figure 3 compares the exact arrival cost function $V^\star_t(x)$ and the proposed quadratic approximation of the arrival cost function $V_t(x)$ used in the MHE. The plots show these two value functions for $t = 50$ (left) and $t = 150$ (right). On the x-axis, we plot the deviation of the state variable from the exact full horizon solution. The y-axis is the arrival cost function value offset by a constant (which is inconsequential for the optimization problem). We observe that the exact arrival cost function $V^\star_t$ and the approximate arrival cost function V_t match very well in the vicinity of the operating point $x = 0$.

B. Accuracy of segmentation

Consider the ℓ_1 trend filtering problem (14), discussed in Section III-C. In this formulation, the proposed MHE update algorithm provides a piecewise linear segmentation of the underlying trend from streaming data. The segmentation is

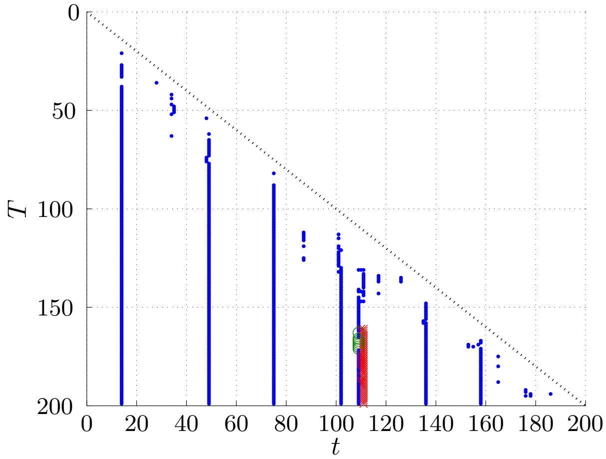

Fig. 4. Segmentation for ℓ_1 trend filtering: a (blue) dot indicates a change point where our MHE solution agrees with the full horizon solution, a (red) 'x' represents false negatives where our MHE solution failed to detect a change point, and a (green) 'o' represents false positives where our MHE solution erroneously detected a change point.

described by the change points (kinks) in the solution. Using the proposed MHE update, we are able to obtain *online segmentation* of streaming data as the problem size T grows out of bounds for a batch solution.

Figure 4 shows the change points detected using the proposed MHE algorithm for the same simulation results as shown in Figure 2. The change points were detected as points where $|x_t - 2x_{t-1} + x_{t-2}| > \epsilon$, where ϵ is defined by the accuracy of the QP optimization solution.

The dots show the detected change points. The points where the full-horizon solution $Z^\star(T)$ and our MHE algorithm do not agree are indicated with an 'x' (false negative) or an 'o' (false positive). The horizontal axis represents the horizon length T, while the vertical axis shows time t inside the full horizon $[1, T]$. The dotted diagonal represents the causality boundary: the solution is obtained for $t \leq T$. The MHE filtering solution in Figure 4 does match the full horizon solution extremely accurately. The differing change points have the solution that is very close to the constraint, close to the QP solver accuracy limit.

C. Accuracy of optimal solution approximation

Checking the accuracy of the arrival cost approximation in Section IV-A needs to be complemented by verification of optimal solution approximation. Furthermore, if decision vector z_t has dimension larger than 2, visualizing the value function is not practical. We compare the performance of the MHE with proposed arrival cost approximation against the performance of an estimator that solves the full-horizon problem at each step.

We again consider the total variation denoising example (13), but we use a larger problem dimension, with x_t and y_t having size of $n = m = 5$, $C = I$, $\Phi = I$, and $\lambda = 20$. We consider an MHE update with window of $N = 50$ applied to the data time series y_t of the length $T = 200$. Figure 5 shows the results. For each t, we plot all $n = 5$

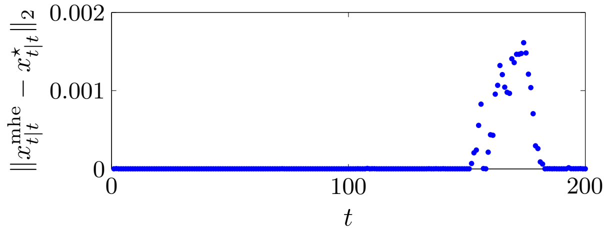

Fig. 5. Comparison of the error between the full horizon estimates ($x^\star_{t|t}$) and the moving horizon estimates ($x^{\mathrm{mhe}}_{t|t}$).

entries of the estimate obtained. The last state estimated with our MHE update $x^{\mathrm{mhe}}_{t|t}$ is plotted along the horizontal axis. The estimation error $x^{\mathrm{mhe}}_{t|t} - x^\star_{t|t}$ is shown on the vertical axis. The small error observed is on the order of accuracy of the QP solver. Furthermore, the numerical QP solution is more accurate for the MHE solution that requires to solve the problem of small size (50 steps) compared to the full-horizon solution (200 steps). This result means that the MHE performance with the proposed approximation of the arrival cost is practically identical to the full-horizon estimator, at least in this example.

V. Conclusion

This paper has presented a moving horizon estimator (MHE) for staged QP problems suitable for streaming data processing. The MHE update is based on the approximation assumption that the active constraint set remains fixed far enough in the past from the last time stage. This assumption allows us to derive an explicit Bellman update for the arrival cost. The update is a generalization of Kalman filter update; with a difference that at each step the dimensions of the updated matrices change depending on the number of active inequality constraints.

References

[1] H. Michalska and D. Mayne, "Moving horizon observers and observer-based control," *IEEE Transactions on Automatic Control*, no. 6, pp. 995–1006, Jun. 1995.

[2] P. Moraal and J. Grizzle, "Observer design for nonlinear systems with discrete-time measurements," *IEEE Transactions on Automatic Control*, no. 3, pp. 395–404, Mar. 1995.

[3] C. Rao, J. Rawlings, and D. Mayne, "Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations," *IEEE Transactions on Automatic Control*, no. 2, pp. 246–258, 2003.

[4] G. C. Goodwin, M. M. Seron, and J. A. D. Dona, *Constrained Control and Estimation - An Optimisation Approach.* Springer Verlag, 2005.

[5] H. Michalska and D. Mayne, "Receding horizon control of nonlinear systems," *IEEE Transactions on Automatic Control*, no. 5, pp. 814–824, May 1990.

[6] J. Primbs and V. Nevistic, "A new approach to stability analysis of finite receding horizon control without end constraint," *IEEE Transactions on Automatic Control*, no. 8, pp. 1507–1512, Aug. 2000.

[7] J. H. Lee and N. L. Ricker, "Extended kalman filter based nonlinear model predictive control," *Industrial and Engineering Chemistry Research*, vol. 33, no. 6, pp. 1530–1541, 1994.

[8] D. Gorinevsky, "Efficient filtering using monotonic walk model," in *American Control Conference*, Jun. 2008, pp. 2816–2821.

[9] S. J. Kim, K. Koh, S. Boyd, and D. Gorinevsky, "ℓ_1 trend filtering," *SIAM Review*, no. 2, pp. 339–360, 2009.