An Algorithm of Smooth Output Reconstruction for Linear Systems with Quantized Measurements

∗Hongzhong Zhu(The University of Tokyo) Hiroshi Fujimoto (The University of Tokyo)

Abstract– This paper proposes an approximate algorithm for solving an l1- minimization problem used for reconstructing the system output. Firstly, the equality constraints are eliminated and the problem is reduced to be an optimization problem with only linear inequality constraints. Then, the non-differentiability of the objective function is relaxed using a regularization parameter. Finally, the solution is determined by taking advantage of the characteristic of constraints. The algorithm can remarkably reduce the calculation time compared with conventional solvers.

Key Words: quantization, polynomial fitting approach, ℓ_1 minimization, convex optimization

1 Introduction

Quantization in I/O signals is an inherent feature in many control systems including digital systems, networked ones, low resolution sensor/actuator systems, and so on. In some cases, the quantization error is substantially small compared to system noise and the desired control accuracy. However, this is not always the case in various systems such as the computer storage systems, NC machine tools, industrial robots and ultra-precision positioning systems where the required accuracy is nano order. In these systems, the quantization error caused by low-resolution sensors could significantly degrade the control performance and may cause limit cycle oscillations [1].

In order to cope with the quantization effects, various methods were proposed in the literature to reconstruct the real system output. One strategy is to utilize the system model information. For instance, some observer-based methods were proposed to estimate the system state, and a reconstruction output can be obtained by utilizing the estimated state and the model information [4, 5]. Hirata *et al.* proposed a method to estimate the quantization error directly via the least square method in the presence of constant disturbance [6]. However, since these methods strongly depend on the system models, the accuracy of the reconstructed output could be degraded drastically when the models are not precise.

There is another line of research on reconstructing the real system output from the noise-corrupted output by curve fitting methods. For instance, polynomial filtering approaches were proposed to recover the non-uniformly sampled signals [7], and position signals obtained from incremental encoders [8]. The methods can work well if the output signal can be locally approximated by low-degree polynomials. However, if this is not the case, higher-degree polynomials are required. Hence, it would be difficult to adopt the methods.

Based on the above observation, Zhu *et al.* proposed a method to reconstruct the position and estimate the velocity of motion systems based on low-resolution encoders [2]. The methods transformed the reconstruction problem into a convex optimization problem combining a curve fitting approach with the observer-based ones. The methods can achieve the smooth position/velocity estimation with high precision even if the encoder resolution is extremely low. However, the conventional solvers, such as the barrier method, may take too much processing time to solve the optimization problem in real time.

The purpose of this paper is to develop an algorithm for efficiently solving the optimization problem proposed in [2]. Firstly, the equality constraints are eliminated so that the optimization problem is reduced to be a minimization problem with linear inequality constraints. Then, the non-differentiability of the objective function is relaxed by applying the approximation $\|x\|_1 \approx \sum_{i=1}^{n} \sqrt{x_i^2 + \mu}\ (x \in \mathbb{R}^n)$ where $\mu > 0$ is a regularization parameter. Finally, the optimal solution is achieved by seeking for the shortest distance between the solution of the non-constraint problem and the feasible set.

This paper is organized as follows. In Section 2, the system model is introduced, and the problem setting is described. Section 2.2 presents the polynomial fitting approach based on convex optimization that consists of the ℓ_1-norm regularization method and some constraint conditions. A numerical example is given in Section 4 to illustrate the advantages of the proposed approach. In Sections 5, the proposed algorithm is implemented in DSP, and its effectiveness is demonstrated through experiments using a high-precision positioning stage. The conclusions are summarized in Section 6.

2 Problem formulation

2.1 System description

Consider the linear time-invariant SISO system with quantized output given by

$$\boldsymbol{x}[k+1] = \boldsymbol{A}\boldsymbol{x}[k] + \boldsymbol{B}u[k] + w[k], \tag{1}$$
$$y[k] = \boldsymbol{C}\boldsymbol{x}[k], \tag{2}$$
$$y_q[k] = \boldsymbol{Q}(y[k]), \tag{3}$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $y \in \mathbb{R}$, and $y_q \in \mathbb{R}$ are the system state, the system output, the corrupted output and the quantized output, respectively. $v \in \mathbb{R}$ is the measurement noise and $w \in \mathbb{R}$ is the unknown input disturbance. $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ are constant system matrices of appropriate dimensions. $\boldsymbol{Q}(\cdot)$ denotes the quantization. For simplicity, w is assumed to be white noise in this study. If this is not the case, augmented system can be constructed by properly modeled the disturbance [2].

The function $\boldsymbol{Q}(\cdot)$ in (3) is assumed as the uniform quantization defined by

$$\boldsymbol{Q}(y) = i \cdot \Delta, \quad y \in ((i-0.5)\Delta\ ,\ (i+0.5)\Delta] \tag{4}$$

where $i \in \mathbb{Z}$, $\Delta > 0$ denotes the quantization step. The quantizer (4) could represent low-resolution op-

tical encoders or analog-to-digital converters in real systems, where Δ is referred to as the resolution.

The difference between the system output y and the measured quantized output y_q, denoted by $\xi := y - y_q$, is bounded by

$$|\xi| \leq \frac{\Delta}{2}. \tag{5}$$

2.2 Output reconstruction

In this subsection, the output reconstruction approach proposed in [2] is introduced. Then, a conventional algorithm for solving the optimization problem is presented.

2.2.1 Polynomial fitting formulation

A polynomial for fitting $p+1$ quantized measurements ($\{y_q[k-i]\}_{i=0,1,\cdots,p}$) is considered. Here, k denotes the current time instant. In order to do so, first, choose the time interval $[a \;\; b]$ in advance, and introduce the virtual time indices $\{\tau_i\}_{i=0,1,\cdots,p}$, which are defined by

$$\tau_0 = a, \tau_1 = a + h, \cdots, \tau_i = a + ih, \cdots, \tau_p = b,$$

where $h = (b-a)/p$, to equally divide $[a \;\; b]$. Then the data $(\tau_i, y_q[k-p+i])_{i=0,1,\cdots,p}$ are fitted in Euclidean space with the polynomial:

$$g_k(t) = \alpha_0 + \alpha_1\tau + \alpha_2\tau^2 + \cdots + \alpha_m\tau^m, \tag{6}$$

where $\tau \in [a \;\; b]$, $\alpha_0,\ \alpha_1,\ \cdots\ \alpha_m$ are the polynomial coefficients, and m is the degree of the polynomial. Without loss of generality, m is assumed to be $m \leq p+1$. The fitting problem is formulated by

$$\min_{\boldsymbol{\alpha}} : \frac{1}{2}\left\| \begin{bmatrix} g_k(\tau_0) - y_q[k-p] \\ \vdots \\ g_k(\tau_i) - y_q[k-p+i] \\ \vdots \\ g_k(\tau_p) - y_q[k] \end{bmatrix} \right\|_2^2 + \eta\|\boldsymbol{\alpha}\|_1, \tag{7}$$

with variable $\boldsymbol{\alpha} := [\alpha_0 \;\; \alpha_1 \;\; \cdots \;\; \alpha_m]^T$. Here, η is the weighting factor. This is an ℓ_1-norm regularization problem, and can be expressed by

$$\min_{\boldsymbol{\alpha}} : \quad \frac{1}{2}\|\boldsymbol{T\alpha} - \boldsymbol{\beta}\|_2^2 + \eta\|\boldsymbol{\alpha}\|_1, \tag{8}$$

where $\boldsymbol{T} = \left[\tau_i^j\right]$ and $\boldsymbol{\beta} = \left[y_q[k-p+i]\right]$ ($i = 0, 1, \cdots, p$ and $j = 0, 1, \cdots, m$).

Note that (8) becomes a normal least-square polynomial fitting problem if $\eta = 0$. In this way, the unnecessary terms of the polynomial (6) can be removed automatically if a high-degree polynomial is used to fit a simple signal. In practical situation, m can be relative large so that simple signals as well as complex signals could be properly fitted. According to the recommendation in [2], $[a \;\; b]$ is chosen as $[-1 \;\; 1]$ in this study.

2.2.2 Moving horizon manner

Once the convex optimization problem (8) is solved, the polynomial (6) is determined. Hence, the quantity $\bar{y}[k]$ is given by

$$\bar{y}[k] := g_k(\tau_p) = \boldsymbol{\tau}_p^T\boldsymbol{\alpha}, \tag{9}$$

where $\boldsymbol{\tau}_p = [1 \;\; \tau_p \;\; \tau_p^2 \;\; \cdots \;\; \tau_p^m]^T$, which is regarded as the reconstruction value of the true output $y[k]$. When the time instant k is updated, say, from k to $k+1$, a new quantized measurement $y_q[k+1]$ is sampled. Then, a new polynomial $g_{k+1}(\tau)$ can be determined by fitting the new data $\{y_q[k-p+1], y_q[k-p+2], \cdots, y_q[k+1]\}$. The updated quantity $\bar{y}[k+1]$ is calculated by

$$\bar{y}[k+1] = g_{k+1}(\tau_p).$$

Figure 1 shows the updating strategy. As shown above, $\bar{y}[k]$ is determined from the data $\{y_q[k-p], y_q[k-p+2], \cdots, y_q[k]\}$ in the moving horizon manner. In the case of $k \leq p$, $y_q[k-i]$ ($i = k, k+1, \cdots, p$) is set as $y_q[k-i] := 0$.

2.2.3 Constraint conditions

In order to improve the reconstruction accuracy, some constraint conditions are added in the optimization problem (8) by taking advantage of the model information.

First, since the discrepancy between the system output $y[k]$ and the corresponding quantized output $y_q[k]$ is bounded by (5), the quantity $\bar{y}[k]$ should be bounded by

$$|\bar{y}[k] - y_q[k]| \leq \frac{\Delta}{2}. \tag{10}$$

According to (6) and (9), this condition can be expressed by

$$\left|\boldsymbol{\tau}_p^T\boldsymbol{\alpha} - y_q[k]\right| \leq \frac{\Delta}{2}. \tag{11}$$

Note that the left hand side of (11) is convex with respect to $\boldsymbol{\alpha}$.

Second, the fitting curve $g_k(t)$ (at time instant k) should take $\bar{y}[k\text{-}1]$ into account as well as the quantized output sequence $y_q[i]$ ($i = k, k-1, \cdots, k-p$). In addition, the system information (including the disturbance model) should be taken into account. Let $\hat{y}[k]$ be the output of the observer for the system (1)$\sim$(3). Namely,

$$\hat{\boldsymbol{x}}[k+1] = \boldsymbol{A}\hat{\boldsymbol{x}}[k] + \boldsymbol{B}u[k] + \boldsymbol{L}(\bar{y}[k] - \hat{y}[k]), \tag{12}$$

$$\hat{y}[k] = \boldsymbol{C}\hat{\boldsymbol{x}}[k], \tag{13}$$

where $\boldsymbol{L} \in \mathbb{R}^{n\times 1}$ is the observer gain to stabilize $\boldsymbol{A} - \boldsymbol{LC}$. Then, the following two constraints are added.

$$g_k(\tau_{p-1}) = \hat{y}[k-1], \tag{14}$$

$$\dot{g}_k(\tau_{p-\frac{1}{2}}) = \frac{1}{h}(\hat{y}[k] - \hat{y}[k-1]). \tag{15}$$

where $\tau_{p-\frac{1}{2}}$ is the mid-point in the internal $[\tau_{p-1},\ \tau_p]$, and $\dot{g}_k(\tau)$ is the first derivative of the polynomial (6), which is given by

$$\dot{g}_k(\tau) = \alpha_1 + 2\alpha_2\tau + \cdots + m\alpha_m\tau^{m-1}. \tag{16}$$

Equation (14) indicates that $\hat{y}[k-1]$ and $\bar{y}[k]$ belong to $g_k(\tau)$, and equation (15) is a slope condition for obtaining smooth reconstruction. Conditions (14) and (15) can be expressed by

$$\boldsymbol{\Gamma\alpha} = \boldsymbol{\gamma}, \tag{17}$$

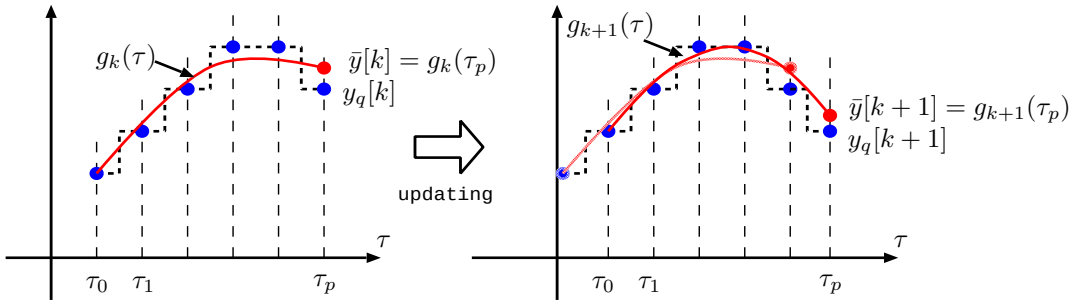

Fig. 1: Updating strategy. A new polynomial is calculated when the time instant is updated.

where

$$\boldsymbol{\Gamma} = \begin{bmatrix} 1 & \tau_{p-1} & \tau_{p-1}^2 & \cdots & \tau_{p-1}^m \\ 0 & 1 & 2 \times \tau_{p-\frac{1}{2}} & \cdots & m \times \tau_{p-\frac{1}{2}}^{m-1} \end{bmatrix},$$

$$\boldsymbol{\gamma} = \begin{bmatrix} \hat{y}[k-1] \\ \frac{1}{h}(\hat{y}[k] - \hat{y}[k-1]) \end{bmatrix}.$$

Note that Γ is full row rank. For the implementation of real-time calculation, $\hat{y}[k]$ is computed via

$$\hat{y}[k] = \boldsymbol{C}\big[\boldsymbol{A}\hat{\boldsymbol{x}}[k-1] + \boldsymbol{B}u[k-1] + \boldsymbol{L}\Big(\bar{y}[k-1] - \hat{y}[k-1]\Big)\big]. \quad (18)$$

2.2.4 System output reconstruction

The quantized output reconstruction algorithm is summarized as follows:

Algorithm for output reconstruction:

1. ***Choose*** *proper values for p, m, η; set initial value for $\boldsymbol{\beta}$ as $\boldsymbol{\beta} = 0$, assign the values for Δ;*
2. ***repeat***
   (a) *Sample the position measurement $y_q[k]$, update the vector $\boldsymbol{\beta}$;*
   (b) *solve the problem*

$$\min_{\boldsymbol{\alpha}}: \ \frac{1}{2}\|\boldsymbol{T}\boldsymbol{\alpha} - \boldsymbol{\beta}\|_2^2 + \eta\|\boldsymbol{\alpha}\|_1 \quad (19)$$
$$\text{subject to}: \ (11),\ (17)$$

   to obtain $\boldsymbol{\alpha}$;
   (c) *calculate $\bar{y}[k]$, $\hat{\boldsymbol{x}}[k]$ by* (9), (12) *and* (13);
   (d) *set $k \leftarrow k+1$.*

The optimization problem (19) is a convex optimization problem and can be solved efficiently. Note that the constraint conditions are independent with each other, the problem is therefore feasible if the polynomial (6) has a degree not less than 2 ($m \geq 2$).

The optimization problem (19) can be rewritten by

$$\min_{\boldsymbol{\alpha},\boldsymbol{\rho}}: \ \frac{1}{2}\|\boldsymbol{T}\boldsymbol{\alpha} - \boldsymbol{\beta}\|_2^2 + \eta\sum_{i=1}^{m+1}\boldsymbol{\rho}_i, \quad (20)$$
$$\text{subject to}:$$
$$\boldsymbol{\Gamma}\boldsymbol{\alpha} = \boldsymbol{\gamma}$$
$$\big|\boldsymbol{\tau}_p^T\boldsymbol{\alpha} - y_q[k]\big| \leq \tfrac{\Delta}{2}$$
$$-\boldsymbol{\rho}_i \leq \boldsymbol{\alpha}_i \leq \boldsymbol{\rho}_i, \quad i = 1, \cdots, m+1.$$

The minimization problem (20) is a quadratic program (QP), and a considerable number of approaches have been proposed in the recent years including interior-point methods [9]. However, solving the problem in real-time for field applications, such as DSP, is a remaining issue [2]. In the next section, a novel approach is proposed to solve the problem efficiently.

3 Algorithm

3.1 Eliminating Equality Constraints

Suppose $\boldsymbol{F} \in \mathbb{R}^{(m+1)\times(m-1)}$ is a matrix whose range is the nullspace of $\boldsymbol{\Gamma}$ ($\boldsymbol{\Gamma}\boldsymbol{F} = \boldsymbol{0}$) and $\boldsymbol{\alpha}_0$ is a particular solution of (17), the the affine feasible set is given by

$$\{\boldsymbol{\alpha} \,|\, \boldsymbol{\Gamma}\boldsymbol{\alpha} = \boldsymbol{\gamma}\} = \{\boldsymbol{F}\boldsymbol{z} + \boldsymbol{\alpha}_0 \,|\, \boldsymbol{z} \in \mathbb{R}^{m-1}\}. \quad (21)$$

The eliminated optimization problem of (19) becomes

$$\min_{\boldsymbol{z}}: \ \frac{1}{2}\|\boldsymbol{M}\boldsymbol{z} + \boldsymbol{\zeta}\|_2^2 + \eta\|\boldsymbol{F}\boldsymbol{z} + \boldsymbol{\alpha}_0\|_1, \quad (22)$$
$$\text{subject to}: \ |\boldsymbol{\lambda}^T\boldsymbol{z} + \xi| \leq \frac{\Delta}{2}$$

where $\boldsymbol{M} = \boldsymbol{T}\boldsymbol{F}$, $\boldsymbol{\zeta} = \boldsymbol{T}\boldsymbol{\alpha}_0 - \boldsymbol{\beta}$, $\boldsymbol{\lambda}^T = \boldsymbol{\tau}_p^T\boldsymbol{F}$ and $\xi = \boldsymbol{\tau}_p^T\boldsymbol{\alpha}_0 - y_q$.

3.2 Convex Relaxation

In order to overcome the non-differentiability of the second term in the objective function of (22), we consider the homotopic principle used in [3] which consists in approximating $\|\boldsymbol{\omega}\|_1$ by $\sum_{i=1}^{m+1}\sqrt{\boldsymbol{\omega}_i^2 + \mu}$, where $\mu > 0$ is a regularization parameter tending to zero. The objective function of the optimization problem (22) is expressed by

$$f(\boldsymbol{z}) = \frac{1}{2}\|\boldsymbol{M}\boldsymbol{z} + \boldsymbol{\zeta}\|_2^2 + \eta\sum_{i=1}^{m+1}\sqrt{(\boldsymbol{F}\boldsymbol{z} + \boldsymbol{\alpha}_0)_i^2 + \mu}, \quad (23)$$

where $(\cdot)_i$ denotes the i^{th} element. The function (23) is strictly convex since

$$\nabla^2 f(\boldsymbol{z}) = \boldsymbol{M}^T\boldsymbol{M} + \eta\sum_{i=1}^{m+1}\frac{\mu\boldsymbol{F}_i^T\boldsymbol{F}_i}{[(\boldsymbol{F}_i\boldsymbol{z} + \boldsymbol{\alpha}_{0_i})^2 + \mu]^{\frac{3}{2}}}$$

is positive definite. Here, $\boldsymbol{F}_i$ is the i^{th} row of $\boldsymbol{F}$ and $\boldsymbol{\alpha}_{0_i}$ is the i^{th} element of $\boldsymbol{\alpha}_0$.

3.3 Approximation algorithm

Note that *a)* the objective function of (22) is strictly convex and *b)* the feasible set is a polyhedral set bounded by two parallel hyperplanes, as shown in Fig. 2. In the figure, $\boldsymbol{\alpha}^*$ is the optimal solution

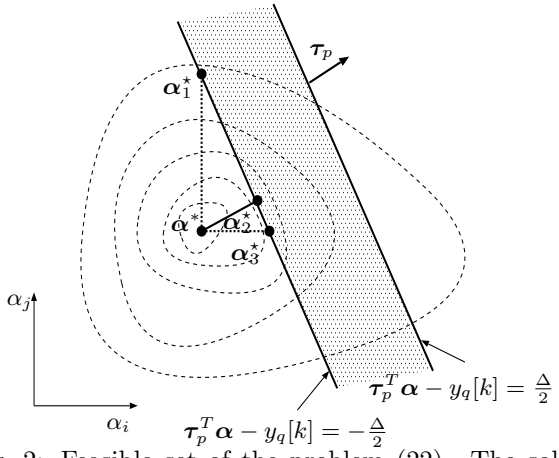

Fig. 2: Feasible set of the problem (22). The solution $\boldsymbol{\alpha}^*$ is the optimal solution without considering the inequality constraints. The solution $\boldsymbol{\alpha}_{1,2,3}^\star$ are the potential optimal solutions.

of (22) without inequality constraints (calculated by $\boldsymbol{\alpha}^* = \boldsymbol{F}\boldsymbol{z}^* + \boldsymbol{\alpha}_0$ where $\boldsymbol{z}^*$ is the solution of (22) without constraints), and $\boldsymbol{\alpha}_j^\star$ $(j = 1, 2, 3)$ are the potential optimal solutions explained in the following. $\boldsymbol{\alpha}_2^\star$ is the nearest point from $\boldsymbol{\alpha}^*$, and $\boldsymbol{\alpha}_{1,3}^\star$ are the solutions to maintain the sparse characteristics of the optimization problem (19). In the following, an algorithm is proposed to find out the solution $\boldsymbol{\alpha}_2^\star$ for the optimization problem (19).

Algorithm for solving the problem (22)

1. ***Solve*** *the problem (22) without considering the constraint conditions to obtain* $\boldsymbol{z}^*$*;*
2. ***Calculate*** $\boldsymbol{\alpha}^*$ *by* $\boldsymbol{\alpha}^* = \boldsymbol{F}\boldsymbol{z}^* + \boldsymbol{\alpha}_0$*;*
3. ***If*** $\boldsymbol{\alpha}^* \in \{\boldsymbol{\alpha} \,|\, |\boldsymbol{\tau}_p^T\boldsymbol{\alpha} - y_q[k]| \le \frac{\Delta}{2}\}$, ***then*** $\boldsymbol{\alpha}_2^\star = \boldsymbol{\alpha}^*$, ***else***
   Apply the orthogonal projection of $\boldsymbol{\alpha}^*$ *to the hyperplanes* $\boldsymbol{\tau}_p^T\boldsymbol{\alpha} - y_q[k] = -\frac{\Delta}{2}$ *and* $\boldsymbol{\tau}_p^T\boldsymbol{\alpha} - y_q[k] = \frac{\Delta}{2}$ *to obtain the pedal points* $\boldsymbol{\alpha}_1$ *and* $\boldsymbol{\alpha}_2$*,* ***then if*** $\|\boldsymbol{\alpha}^* - \boldsymbol{\alpha}_1\|_2 < \|\boldsymbol{\alpha}^* - \boldsymbol{\alpha}_2\|_2$, $\boldsymbol{\alpha}_2^\star = \boldsymbol{\alpha}_1$, ***else*** $\boldsymbol{\alpha}_2^\star = \boldsymbol{\alpha}_2$.

The concept is shown in Fig. 2. Step 1 in the algorithm is realized by solving the extremum of the objective function of (22), as shown in the following.

The gradient of $f(\boldsymbol{z})$ is

$$\nabla f(\boldsymbol{z}) = \left(\boldsymbol{M}^T\boldsymbol{M} + \eta \sum_{i=1}^{m+1} \frac{\boldsymbol{F}_i^T\boldsymbol{F}_i}{\sqrt{(\boldsymbol{F}_i\boldsymbol{z} + \boldsymbol{\alpha}_{0_i})^2 + \mu}}\right)\boldsymbol{z} + \boldsymbol{M}^T\boldsymbol{\zeta} + \eta \sum_{i=1}^{m+1} \frac{\boldsymbol{\alpha}_{0_i}\boldsymbol{F}_i^T}{\sqrt{(\boldsymbol{F}_i\boldsymbol{z} + \boldsymbol{\alpha}_{0_i})^2 + \mu}} \quad (24)$$

The optimal solution $\boldsymbol{z}^*$ is obtained by solving the following system of nonlinear equations

$$\left(\boldsymbol{M}^T\boldsymbol{M} + \eta \sum_{i=1}^{m+1} \frac{\boldsymbol{F}_i^T\boldsymbol{F}_i}{\sqrt{(\boldsymbol{F}_i\boldsymbol{z} + \boldsymbol{\alpha}_{0_i})^2 + \mu}}\right)\boldsymbol{z} = -\boldsymbol{M}^T\boldsymbol{\zeta} - \eta \sum_{i=1}^{m+1} \frac{\boldsymbol{\alpha}_{0_i}\boldsymbol{F}_i^T}{\sqrt{(\boldsymbol{F}_i\boldsymbol{z} + \boldsymbol{\alpha}_{0_i})^2 + \mu}}. \quad (25)$$

Table 1: Comparison of different algorithms

solver	calculation time	average error e_r
CVX	46.1 min	$4.44e^{-6}$
fmincon	351 sec	$7.95e^{-6}$
proposed	6.19 sec	$4.06e^{-6}$

Since $\left(\boldsymbol{M}^T\boldsymbol{M} + \eta \sum_{i=1}^{m+1} \frac{\boldsymbol{F}_i^T\boldsymbol{F}_i}{\sqrt{(\boldsymbol{F}_i\boldsymbol{z} + \boldsymbol{\alpha}_{0_i})^2 + \mu}}\right)$ is positive definite, (25) can be cast as a fixed point problem of the form

$$\boldsymbol{z} = \boldsymbol{\phi}(\boldsymbol{z}), \quad (26)$$

where

$$\boldsymbol{\phi}(\boldsymbol{z}) = \left(\boldsymbol{M}^T\boldsymbol{M} + \eta \sum_{i=1}^{m+1} \frac{\boldsymbol{F}_i^T\boldsymbol{F}_i}{\sqrt{(\boldsymbol{F}_i\boldsymbol{z} + \boldsymbol{\alpha}_{0_i})^2 + \mu}}\right)^{-1} \times \left(-\boldsymbol{M}^T\boldsymbol{\zeta} - \eta \sum_{i=1}^{m+1} \frac{\boldsymbol{\alpha}_{0_i}\boldsymbol{F}_i^T}{\sqrt{(\boldsymbol{F}_i\boldsymbol{z} + \boldsymbol{\alpha}_{0_i})^2 + \mu}}\right).$$

In order to solve (26), we proceed by taking an initial solution $\boldsymbol{z}_-$, and iteratively, compute $\boldsymbol{z} = \boldsymbol{\phi}(\boldsymbol{z}_-)$ until two consecutive solutions $\boldsymbol{z}$ and $\boldsymbol{z}_-$ are sufficiently close.

The calculation of pedals in step 3 of the algorithm is obtained by

$$\boldsymbol{\alpha}_{1,2} = \boldsymbol{\alpha}^* + \frac{\delta - \boldsymbol{\tau}_p^T\boldsymbol{\alpha}^*}{\boldsymbol{\tau}_p^T\boldsymbol{\tau}_p}\boldsymbol{\tau}_p, \quad (27)$$

where $\delta = y_q[k] - \frac{\Delta}{2}$ for $\boldsymbol{\alpha}_1$ or $\delta = y_q[k] + \frac{\Delta}{2}$ for $\boldsymbol{\alpha}_2$.

4 Illustrative examples

In this section, one numerical example is given to illustrate the effectiveness of the proposed method. Consider the motion system given by

$$G(s) = \frac{26.5}{14.7s^2 + 24s}, \quad (28)$$

and the quantization step is set as $\Delta = 50$ μm. For implementation, the system is discretized by the sampling period $T_s = 1$ ms. In the setup, the order of the polynomial in (6) is set as $m = 6$, the number of quantized output used for fitting is $p + 1 = 30$, and the weight factor is $\eta = 7 \times 10^{-6}$. The observer gain $\boldsymbol{L}$ in (12) (13) is properly given as $[0.1083\ 2.8111]^T$ so that the bandwidth is 8 Hz.

The input u is set as $u = 0.3\sin(\pi t)$, and the initial output is 0.1 m. For comparison, a widely used software CVX [13] and the function "fmincon" in Matlab® using the algorithm of "sqp" are implemented. The reconstruction error used for evaluation are defined by

$$e_r := \bar{y} - y. \quad (29)$$

The comparison of the calculation time and the accuracy of reconstruction error evaluated via Root-Mean-Square (RMS) is shown in Table 1. The results show that the proposed algorithm can remarkably reduce the calculation time while maintaining the computing accuracy.

The numerical simulations for comparing

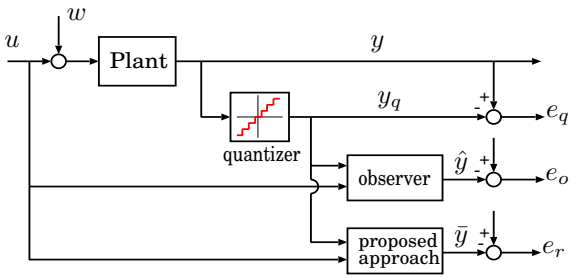

Fig. 3: Block diagram of the quantized system for simulation.

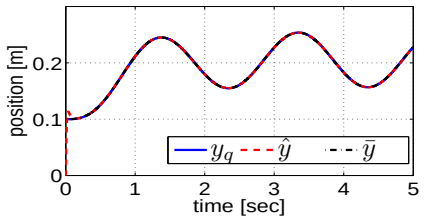

Fig. 4: Output comparison.

(a) quantized output y_q,

(b) an observer-based output estimate $\hat{y}$,

(c) the reconstructed output $\bar{y}$ using the proposed algorithm,

are also performed. The block diagram is shown in Figure 3. For comparison, the errors of different methods are defined:

$$e_q := y_q - y, \tag{30}$$
$$e_o := \hat{y} - y, \tag{31}$$

The numerical results are shown in Figures 4∼6. Figure 4 shows the outputs of y_q, $\hat{y}$ and $\bar{y}$. The reconstruction error e_r is much smaller compared to the quantization error e_q and the estimation error e_o. Figure 6 illustrates the effectiveness of the proposed approach clearly.

5 Experimental validation

In this section, the effectiveness of the proposed algorithm will be validated through experiments using a high precision stage.

5.1 Description of the experiment system

The high precision stage is shown in Figure 7. The stage is a simplified model of the scanner in exposure systems used for the fabrication of integrated circuits, which is required to achieve extremely precise motion control. Two linear motors located at the both sides of the carriage are applied to drive

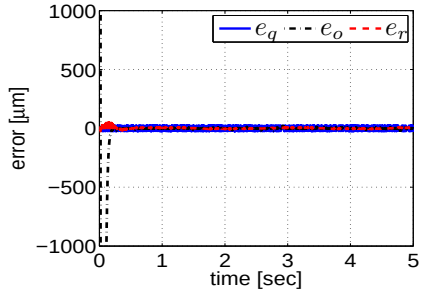

Fig. 5: Comparison of the errors.

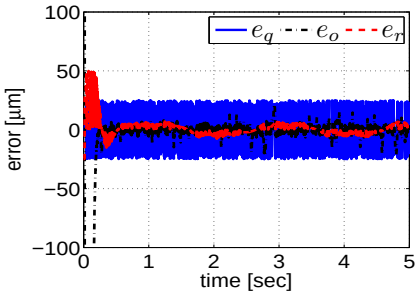

Fig. 6: Comparison of the errors (enlarged).

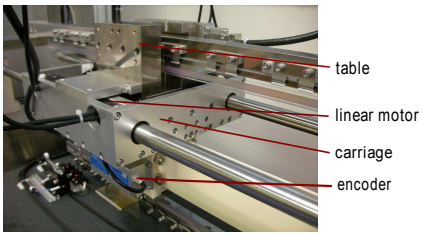

Fig. 7: Experimental setup.

the stage. An air guide system is introduced to reduce the friction between the stator and the slider of the motors. DSP(TMS320C6713, 225MHz) is used as the processor to implement the controllers and the proposed fitting approach. The nominal model is expressed as (28).

5.2 Control system

A two-degree-of-freedom controller is exploited to control the stage. The feed forward controller, designed by perfect tracking control (PTC) method [12], is the stable inverse system of the nominal plant so that the perfect tracking at every sampling instant can be guaranteed. The feedback controller is designed as the PID compensator given by

$$K = k_p + k_d \frac{s}{0.004s + 1} + k_i \frac{1}{s}, \tag{32}$$

where $k_p = 2065.6$, $k_i = 34624$, $k_d = 43.45$, the bandwidth of the close-loop system is 10 Hz. The control system is discretized with the sampling period $T_s = 1$ ms.

Though a linear encoder with the resolution of 1 nm is available for the position measurement in this system, we do not use it for control purpose. Instead, a software quantizer is introduced, whose resolution (i.e., quantization step) is $\Delta = 20$ μm in order to evaluate the effectiveness of the proposed method. The quantized output y_q (with the resolution 20 μm) is supposed to be used for control. The high-resolution encoder is used only for monitoring the performance precisely. In this way, the measurements from the linear encoder is regarded as the actual output y, and the signal from the software quantizer is treated as the quantized output y_q.

5.3 Implementation of the algorithm

According to Section 3.3, we implement the proposed algorithm to yield $\bar{y}$ based on u and y_q. The number of data used for polynomial fitting is set as $p + 1 = 15$, and the degree of polynomial is set as $m = 4$. The weight factor η is properly chosen as $\eta = 7 \times 10^{-6}$. The gain $\boldsymbol{L}$ of the state estimator is given as same as the gain in Section 4. By this

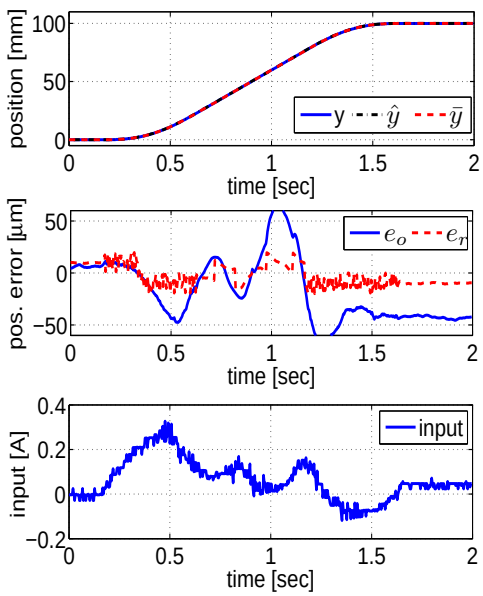

Fig. 8: Comparison of position estimation and position estimation errors.

setting, the average computational time of the algorithm is about 160 μs, which implies the problem can be solved safely during the sampling period $T_s = 1$ ms.

5.4 Experimental results

Figure 8 shows the comparison of position estimation between observer and reconstruction method using proposed algorithm. The upper graph shows the comparison of real output, estimated position by observer and reconstructed position $\bar{y}$. The middle graph shows the comparison of estimation error by observer and the reconstruction error by reconstruction method. And the lower graph shows the control input. Due to the friction and the control input is not zero even the stage is in motionless, the estimation/reconstruction errors cannot converge to zero when stage is in steady state. However, from the results, it is observed that the output reconstruction error is far small than the estimation error by observer.

6 Conclusion

This paper has presented an algorithm to reconstruct the smooth output based on the quantized measurement in the presence of system disturbances. By fitting the quantized measurements with polynomials in a moving horizon manner, a smooth signal is reconstructed via solving an ℓ_1 minimization problem with some equality and inequality constraints. Firstly, the equality constraints are eliminated so that the problem is reduced to be an optimization problem with linear inequality constraints only. Then, the non-differentiability of the objective function is relaxed using a regularization parameter. Its performance has been illustrated through numerical simulation. Furthermore, the proposed method has been implemented in a real high-precision linear stage through DSP, and its effectiveness is verified though the experiments using a high-precision system.

References

[1] S. Devasia, E. Eleftheriou, and M.S.O. Reza, "A Survey of Control Issues in Nanopositioning," *IEEE Trans. Control Syst. Technol.*, Vol. 15, No. 5, pp. 802–823 (2007).

[2] H. Zhu, H. Fujimoto, and T. Sugie, "Proposal of Position Reconstruction with Polynomial Fitting Approach for Precise Motion Control," *IFAC-Mechatronics*, Apr. 2013.

[3] C. Ramirez, and M. Argaez, "An ℓ_1 Minimization Algorithm for Non-Smooth Regularization in Image Processing," *Signal, Image and Video Processing*, pp. 1863–1703, Mar. 2013.

[4] D. Luong-Van, M. Tordon, and J. Katupitiya, "Covariance Profiling for an Adaptive Kalman filter to Suppress Sensor Quantization Effects," *Proc. 43rd IEEE Conf. on Decis. Control*, pp. 2680–2685, 2004.

[5] J. Zhang, and M. Fu, "A Reset State Estimator for Linear Systems to Suppress Sensor Quantization Effects," *Proc. 17th IFAC World Congress*, pp. 9254–9259,2008.

[6] M. Hirata, and T. Kidokoro, "Servo Performance Enhancement of Motion System via a Quantization Error Estimation Method–Introduction to Nanoscale Servo Control," *IEEE Trans. Ind. Electron.*, Vol. 56, No. 10, pp. 3817–3824, 2009.

[7] T.I. Laakso, A. Tarczynski, N.P. Murphy, and V. Valimaki, "Polynomial Filtering Approach to Reconstruction and Noise Reduction of Nonuniformly Sampled Signals," *Signal Processing*, Vol. 80, No. 4, pp. 567–575, 2000.

[8] R.J.E. Merry, M.J.G. van de Molengraft, and M. Steinbuch, "Velocity and acceleration estimation for optical incremental encoders," *Mechatronics*, Vol. 20, No. 1, pp. 20–26, 2010.

[9] S. Boyd, L. Vandenberghe. Convex Optimization, *Cambridge University Press*, 2004.

[10] W. Gautschi, "Optimally Conditioned Vandermonde Matrices," *Numer. Math.*, Vol. 24, No. 1, pp. 1-12, 1975.

[11] W.-W. Che, and G.-H. Yang, "H_∞ filter design for continuous-time systems with quantised signals," *International Journal of Systems Science*, DOI: 10.1080/00207721.2011.600473, 2004.

[12] H. Fujimoto, Y. Hori, and A. Kawamura, "Perfect tracking control based on multirate feed forward control with generalized sampling periods," *IEEE Trans. Ind. Electron.*, Vol. 48, No. 3, pp. 636–644, 2001.

[13] CVX: Matlab Software for Disciplined Convex Programming, *http://www.stanford.edu/∼boyd/software.html.*