A Study on Piecewise Linear Approximation in the $L_1$ Optimal Control Problem of Sampled-Data Systems

Jung Hoon Kim and Tomomichi Hagiwara

Abstract—This paper studies a mathematical basis of piecewise linear approximation in the $L_1$ optimal control problem of sampled-data systems, which aims at minimizing the $L_\infty$-induced norm of sampled-data systems. To take into account of intersample behavior, we first consider the lifted representation of sampled-data systems. The sampling interval $[0, h)$ is then divided into $M$ subintervals with an equal width by fast-lifting. Finally, the signals on each subinterval are approximated by linear functions through the introduction of two types of ‘linearizing operators.’ This method is called piecewise linear approximation of sampled-data systems. By using the arguments of preadjoint operators, we verify that this approximation techniques gives a theoretical basis for tackling the analysis and synthesis problems on the $L_1$ optimal control of sampled-data systems.

I. INTRODUCTION

Disturbance rejection problem is one of the main issues in control systems, and system norms are used to evaluate the effect of disturbances. Among various system norms, the $L_\infty$-induced (or $l_\infty$-induced) norm is used to deal with the bounded persistent disturbances, such as steps and sinusoids, which are often encountered in control systems. Because this norm corresponds to the $L_1$ (or $l_1$) norm of the impulse response of the system in the continuous-time (or discrete-time) case, the study associated with the treatment of the $L_\infty$-induced norm (or $l_\infty$-induced norm) has been named the $L_1$ (or $l_1$) problem. Some special cases of $L_1$ ($l_1$) control problem have been formulated in [1] but a general case was not dealt with. The general case of continuous-time $L_1$ problem was discussed in [2]–[4] while the case of discrete-time $l_1$ problem was dealt with in [5]–[8]. Stimulated by the success in these studies, the $L_1$ problem of sampled-data systems (with intersample behavior taken into account) has been studied in [9]–[11]. However, in contrast to the case of the $H_\infty$ problem of sampled-data systems [12]–[19], no exact solution has been obtained for the $L_1$ problem of sampled-data systems, for which only approximate methods have been provided. More precisely, in [9]–[11], a sampled-data system is approximately treated as a discrete-time system through the fast-sample/fast-hold (FSFH) approximation technique [20], and it is shown that the $l_\infty$-induced norm of the resulting discrete-time system converges to the $L_\infty$-induced norm of the original sampled-data system at the rate of $1/M$, as the FSFH approximation parameter $M$ tends to infinity.

As a significant advance over the conventional methods through the FSFH approximation, the present authors developed in [22] an extended method for the $L_1$ analysis problem of sampled-data systems by using ideas of fast-lifting [18] and piecewise linear approximation [21]. Fast-lifting also has an integer parameter $M$ as in the FSFH approximation technique. However, it is used only to subdivide the sampling interval $[0, h)$ into $M$ smaller pieces, while the conventional FSFH approximation technique takes $M$ equally spaced sampling points on the interval $[0, h)$; no information is hence lost as to signals on $[0, h)$ by fast-lifting. Through such treatment, a method for computing an upper bound and a lower bound of the $L_\infty$-induced norm of sampled-data systems is provided and it is shown that the gap between the upper and lower bounds converges to 0 at convergence rate $1/M^2$ for the fast-lifting parameter $M$.

Unfortunately, however, this method is restricted to analysis and cannot be used directly for synthesis. This is because this method computes the $L_1[0, h/M)$ norms of linear kernel functions associated with the continuous-time generalized plant and the discrete-time controller, and the structure of the way the controller parameters are involved in the kernel functions is complicated. In contrast, this paper aims at establishing a mathematical basis for piecewise linear approximation in the $L_1$ optimal controller synthesis problem of sampled-data systems. We first apply fast-lifting to divide the sampling interval $[0, h)$ into $M$ subintervals with an equal width. We then approximate the signals on the interval $[0, h/M)$ by linear functions, where we introduce two types of ‘linearizing operators’ for signals, one for the input signals and the other for the output signals. By applying the arguments of preadjoint operators, we develop a theoretical basis for piecewise linear approximation in the $L_1$ optimal controller synthesis problem, with convergence proof and derivation of associated error bounds.

The organization of this paper is as follows. In Section II, we give some mathematical preliminaries. We then review the lifted representation of sampled-data systems [12]–[14] to take into account of intersample behavior in Section III. In Section IV, we provide a piecewise linear approximation procedure of sampled-data systems. Section V provides the main arguments of this paper, in which we give an error analysis associate with piecewise linear approximation and show that it has an attractive property that can be a solid mathematical basis for dealing with the $L_1$ optimal controller synthesis problem of sampled-data systems. We present final remarks in Section VI.

II. MATHEMATICAL PRELIMINARIES

This section gives some mathematical preliminaries. We use the notations $\mathbb{N}$, $\mathbb{R}_\infty$, $\mathbb{R}_1^\nu$ and $\mathcal{B}(\cdot)$ to denote the set of positive integers, the Banach space of $\nu$-dimensional real vectors equipped with vector $\infty$-norm, the Banach space of $\nu$-dimensional real vectors equipped with vector $1$-norm, and the range of an operator, respectively.
The dual space of a Banach space \( X \), i.e., the space of all bounded linear functionals on \( X \), is denoted by \( X^* \).

Let \( X \) and \( Y \) be Banach spaces. For a linear operator \( T : X \to Y \), its adjoint [23]–[25] is denoted by \( T^* : Y^* \to X^* \), which satisfies by definition that

\[
\forall x \in X, \forall \phi \in Y^*, \quad \langle Tx, \phi \rangle = \langle x, T^* \phi \rangle
\]

where the notation \( \langle y, \phi \rangle \) means the value of the linear functional \( \phi \) at \( y \). For the given Banach spaces \( X \) and \( Y \), suppose that there exists unique Banach spaces, denoted by \( X_0 \) and \( Y_0 \), such that their dual spaces \( (X_0)^* \) and \( (Y_0)^* \) coincide with \( X \) and \( Y \), respectively. Then, if there exists an operator \( T_0 \) such that \( (T_0)^* = T \), then \( T_0 \) is called the preadjoint \([23]–[25]\) of \( T \) and \( T_0 \) can be easily seen that such an operator \( T_0 \), if it exists, is unique. It is a fact that \( ||T_0|| = ||T|| \), where \( ||T_0|| \) denotes the norm of \( T_0 \) induced from the norms on \( Y_0 \) and \( X_0 \), while \( ||T|| \) denotes the norm of \( T = (T_0)^* \) induced from the (dual) norms on \( (X_0)^* \) and \( (Y_0)^* \).

Not every operator has a preadjoint, but those operators we deal with in this paper do; it suffices to note that for \( X = (L_1(0, h))^m \) and \( X = \mathbb{R}_n^m \), a unique \( X_0 \) such that \((X_0)^* = X \) is \( X_0 = (L_1(0, h))^m \) and \( X_0 = \mathbb{R}_n^m \), respectively.

Regarding \((L_1(0, h))^m \), we sometimes drop \( m \) and slightly abuse a term for simplicity, especially when we refer to the induced norm of an operator; for an operator \( T : X \to Y \) with \( X \) being Banach space with norm \( || \cdot ||_X \) and \( || \cdot ||_Y \), respectively, we call \( ||T|| = \sup_{x \in X \{0\}} |Tx|/|x| \) the \( L_1(0, h) \)-induced norm of \( T \) if either \( X \) or \( Y \) is \((L_1(0, h))^m \). A similar convention applies when \( L_\infty(0, h) \) is replaced by a similar space.

The notation \( || \cdot || \) is used to mean either the \( L_\infty(0, h) \) norm of a vector function, i.e.,

\[
||f(\cdot)|| := \max \esssup_{0 \le t \le h} |f(t)|
\]

(or that with \( h \) replaced by \( h/M \) or \( \infty \), the \( L_\infty(0, h) \)-induced norm (or that with \( h/M \) or \( \infty \) instead of \( h \)) of an operator, or the \( \infty \)-norm of a matrix or a vector, whose distinction will be clear from the context. On the other hand, the notation \( || \cdot ||_1 \) is used to mean either the \( L_1(0, h) \) norm of a vector function, i.e.,

\[
||f(\cdot)||_1 := \sum_{i} \int_{0}^{h} |f_i(t)| dt
\]

(or that with \( h \) replaced by \( h/M \) or \( \infty \), the \( L_1(0, h) \)-induced norm (or that with \( h/M \) or \( \infty \) instead of \( h \)) of an operator, or the \( 1 \)-norm of a matrix or a vector, whose distinction will also be clear from the context.

For a Banach space \( X \), we identify the direct product \((X^m)^n \) with \( X^{mn} \) when we refer to the norm on the former. We also use the notation \( l_X \) to denote the space of all \( X \)-valued sequences, where \( X \) is some Banach space.

### III. LIFTED REPRESENTATION OF SAMPLED-DATA SYSTEMS

Let us consider the sampled-data system \( \Sigma_{SD} \) shown in Figure 1, where \( P \) denotes the continuous-time generalized plant, while \( \Psi \), \( \mathcal{H} \) and \( S \) denote the discrete-time controller, the zero-order hold and the ideal sampler, respectively, operating with sampling period \( h \) in a synchronous fashion.
and the operators
\[ B_1w = \int_0^h \exp(A(h - \theta))B_1w(\theta)d\theta : (L_\infty[0, h])^{n_w} \to \mathbb{R}^{n_1} \quad (5) \]
\[ (C_1x)(\theta) = C_1 \exp(A\theta)x : \mathbb{R}^{n_1} \to \cdots : (L_\infty[0, h_0])^{n_w} \to (L_\infty[0, h_0])^{n_z} \quad (21), (22) \]
(although the introduction of \( H_0 \) was only implicit there), described

Connecting \( \Psi \) to the above \( \hat{P} \) leads to the mapping between \( \{ \hat{w}_k \}^{\infty}_{k=0} \) and \( \{ \hat{z}_k \}^{\infty}_{k=0} \), which we denote by \( \mathcal{F}(\hat{P}, \Psi) \); it coincides with the lifted representation \( \mathcal{W}_h \mathcal{F}(P, H_s \Psi) \mathcal{W}_h^{-1} \) for the mapping \( \mathcal{F}(P, H_s \Psi S) \) between \( w \in L_\infty^{n_w} \) and \( z \in L_\infty^{n_z} \) in \( \Sigma_{SD} \). Since \( \mathcal{W}_h \) is norm-preserving, we see that the \( L_\infty \)-induced norm \( \| \mathcal{F}(P, H_s \Psi S) \| \) of \( \Sigma_{SD} \) coincides with the \( L_\infty^{n_z} \)-induced norm \( \| \mathcal{F}(P, \Psi) \| \).

Let us introduce the notation \( \mathbf{M}_1 := [C_1 \quad D_{12}] \), which can be described by
\[ \left( \begin{array}{c} x \\ u \end{array} \right)(\theta) = C_0 \exp(A_2\theta) \left( \begin{array}{c} x \\ u \end{array} \right) \quad (9) \]
where
\[ C_0 := [C_1 \quad D_{12}], \quad A_2 := \begin{bmatrix} A & B_2 \\ 0 & 0 \end{bmatrix} \quad (10) \]

Then, \( \mathcal{F}(\hat{P}, \Psi) \) admits the representation
\[ \mathcal{F}(\hat{P}, \Psi) = \mathbf{M}_1 \mathcal{F}(P_d, \Psi) \mathbf{B}_1 + \mathbf{D}_{11} \quad (11) \]
as in [11], where \( \mathcal{F}(P_d, \Psi) \) denotes the mapping between the discrete-time signals \( \eta_k \in \mathbb{R}^{n_\eta} \) and \( \zeta_k \in \mathbb{R}^{n_z+n_\eta} \) associated with the closed-loop system obtained by connecting \( \Psi \) to the (standard lifting-free) discrete-time plant
\[ \begin{align*}
\begin{cases}
x_{k+1} = A_dx_k + \eta_k + B_{2d}u_k \\
\zeta_k = \begin{bmatrix} I \\ 0 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ I \end{bmatrix} u_k \\
y_k = C_{2d}u_k
\end{cases}
\end{align*} \quad (12) \]

Remark 1: In (11), since the left hand side denotes a dynamical system in discrete-time with the lifted input \( \{ \hat{w}_k \}^{\infty}_{k=0} \) and output \( \{ \hat{z}_k \}^{\infty}_{k=0} \), the operator \( \mathbf{B}_1 \) on the right hand side acts on every \( \hat{w}_k \), and \( \hat{z}_k \) is associated with the output of \( \mathbf{M}_1 \) for every \( k \). Similarly for the interpretation of \( \mathbf{D}_{11} \). Similar conventions apply to the following arguments.

In this paper, we aim at approximating the operators \( \mathbf{B}_1, \mathbf{M}_1 \) and \( \mathbf{D}_{11} \) by using the idea of piecewise linear approximation [21], [22], and establishing a theoretical basis of piecewise linear approximation in the \( L_1 \) optimal control problem of sampled-data systems.

IV. PIECEWISE LINEAR APPROXIMATION OF SAMPLED-DATA SYSTEMS

This section is devoted to providing a method for piecewise linear approximation of sampled-data systems. As a key idea in applying the piecewise linear approximation method, we first review fast-lifting [19]. For \( M \in \mathbb{N} \) and \( h' := h/M \), fast-lifting (with the fast-lifting parameter \( M \)) is defined as the mapping from \( f \in (L_\infty[0, h])^\nu \) (or \( f \in (L_1[0, h])^\nu \)) to \( \hat{f} := [(f^{(1)})^T \cdots (f^{(M)})^T]^T \in (L_\infty[0, h'])^{M\nu} \) (or \( \hat{f} \in (L_1[0, h'])^{M\nu} \)) denoted by \( \hat{f} = L_M f \) (irrespective of the underlying space for \( f \), for notational simplicity), where
\[ f^{(i)}(\theta') := f((i-1)h' + \theta') \quad (0 \leq \theta' < h') \quad (13) \]
It is easy to see that \( L_M \) is norm-preserving. It readily follows from this property that
\[ \| \mathcal{F}(\hat{P}, \Psi) \| = \| L_M \mathcal{F}(\hat{P}, \Psi) L_M^{-1} \| \quad (14) \]
where the right-hand side means the \( L_\infty^{(0,h')}_0 \)-induced norm. We denote \( L_M \mathcal{F}(\hat{P}, \Psi) L_M^{-1} \) by \( \mathcal{F}_{M}(\hat{P}, \Psi) \), i.e.,
\[ \mathcal{F}_{M}(\hat{P}, \Psi) = L_M \mathbf{M}_1 \mathcal{F}(P_d, \Psi) \mathbf{B}_1 L_M^{-1} + L_M \mathbf{D}_{11} L_M^{-1} \quad (15) \]
and we call it the fast-lifted representation of \( \Sigma_{SD} \).

We are in a position to review the piecewise linear approximation treatment of the operators \( \mathbf{B}_1, \mathbf{M}_1 \) and \( \mathbf{D}_{11} \) in the above fast-lifted representation (15), which immediately leads us to piecewise linear approximation of \( \Sigma_{SD} \). Such treatment was developed in [21], [22] for analysis problems, but it is very hard to extend the method in [22] in such a way that the sampled-data \( L_1 \) controller synthesis problem can be dealt with. This paper employs essentially the same approximation of these operators but aims at establishing a mathematical basis for a discretization procedure of the continuous-time generalized plant for the sampled-data \( L_1 \) controller synthesis problem. Here, the sampled-data \( L_1 \) controller synthesis problem aims at minimizing the \( L_\infty \)-induced norm of \( \Sigma_{SD} \), and giving a mathematical basis implies to ensure that an appropriate discretization procedure of the generalized plant, together with the associated approximation error analysis, reduces the problem to an (almost) equivalent discrete-time \( l_1 \) control problem aiming at minimizing the \( l_\infty \)-induced norm. To describe the approximation treatment, we first introduce with \( h' = h/M \) the ‘linearizing’ operator \( \mathbf{J}'_1 : (L_\infty[0, h'])^{n_w} \to (L_\infty[0, h'])^{n_z} \) [21], [22] (by which we mean that \( \mathbf{J}'_1 w \) is always a linear function) described by
\[ (\mathbf{J}'_1 w)(\theta') = \int_0^{h'} f_0(\tau')w(\tau')d\tau' + \theta' \int_0^{h'} f_1(\tau')w(\tau')d\tau' \quad (16) \]
with the scalar-valued functions
\[ f_0(\tau') = -\frac{6}{(h')^2} \tau' + \frac{4}{h'}, \quad f_1(\tau') = \frac{12}{(h')^3} \tau' - \frac{6}{(h')^2} \quad (17) \]
\( \mathbf{J}'_1 \) is tailored to possess important properties in terms of some Taylor expansion arguments [21], [22], and further satisfies \( \mathbf{J}'_1 w = w \) for any linear function \( w \); such properties will be used in the proof of our main results. We further introduce another ‘linearizing’ operator \( \mathbf{H}'_1 : (L_\infty[0, h'])^{n_w} \to (L_\infty[0, h'])^{n_z} \), as well as the operator \( \mathbf{D}'_{11} : (L_\infty[0, h'])^{n_w} \to (L_\infty[0, h'])^{n_z} \) [21], [22] (although the introduction of \( \mathbf{H}'_1 \) was only implicit there), described
respectively by

\[
(H_1)z(\theta') = z(0) + \theta' \frac{dz(\tau')}{d\tau'} \bigg|_{\tau'=0} (0 \leq \theta' < h')
\]  

(18)

\[
(D_{p1}^j)w(\theta') = \int_0^{\theta'} C_1 B_1 w(\tau')d\tau' + D_{11} w(\theta') (0 \leq \theta' < h')
\]  

(19)

Obviously, \( H_1 \) is not an operator on \( L_\infty(0, h') \) but on its subspace of functions continuous and (right) differentiable at time 0. However, this issue causes no problem since \( H_1 \) is used for approximating \( M_1 \) and operates only on its output. On the other hand, approximation of \( B_1 \) should take into account that its input may be discontinuous, so that the other more involved linearizing operator \( J_0 \) is used. \( D_{p1}^j \) is used for approximating \( D_{11} \) as in [21], [22]. The details of such approximation treatment is as follows.

Following the ideas mentioned above, we consider replacing \( L_1 M_1 \) and \( B_1 L_{-1}^j \) in (15) with \( H_1 L_1 M_1 \) and \( B_1 L_{-1}^j J_1 ^\circ \), respectively, where \( (\cdot)^{\circ} \) denotes \( \text{diag}(\cdot, \ldots, \cdot) \) consisting of \( M \) copies of \( (\cdot) \). To facilitate such treatment, let us introduce the operators \( H_{M1} \) and \( J_{M1} \) by

\[
H_{M1} = H_1^{\circ} L_1 M_1 : (L_\infty(0, h'))^M \rightarrow (L_\infty(0, h'))^M
\]  

(20)

\[
J_{M1} = L_1^{\circ} J_1 ^\circ : (L_\infty(0, h'))^M \rightarrow (L_\infty(0, h'))^M
\]  

(21)

**Remark 2:** By introducing the matrix \( A_{2d}^M \) appropriately, \( B_1 L_{-1}^j \) can be described as \( A_{2d}^M B_1 \). Hence, approximating \( B_1 L_{-1}^j \) with \( B_1 L_{-1}^j J_1 ^\circ \) is equivalent to approximating \( B_1 \) with \( B_1 J_1 ^\circ \). Similarly, approximating \( L_1 M_1 \) with \( H_1 L_1 M_1 \) is equivalent to approximating \( M_1 \) with \( H_1 M_1 \).

Next, to facilitate the treatment of \( L_1 M_1 L_{-1}^j \) in (15), we introduce the operators \( B_1 \), \( M_1 \) and \( D_{11} \) respectively, with the horizon \([0, h')\) replaced by \([0, h']\), as well as the matrices

\[
A_{2d}^M = \exp(A_2 h') : \mathbb{R}^{n+n_u} \rightarrow \mathbb{R}^{n+n_u}
\]

\[
J = \begin{bmatrix}
I \\
0
\end{bmatrix} : \mathbb{R}^{n+n_u} \rightarrow \mathbb{R}^{n+n_u}
\]

Then, it is easy to see that \( L_1 M_1 L_{-1}^j \) is described by

\[
L_1 M_1 L_{-1}^j = M_1^r A_{M0} B_1^r + D_{11}
\]  

(22)

(see, e.g., [19]), where

\[
A_{M0} := \begin{bmatrix}
0 & 0 & \cdots & 0 \\
J & \cdots & \cdots & \cdots \\
\vdots & \cdots & \cdots & \cdots \\
(A_{2d}^M)^{M-2} J & \cdots & \cdots & 0
\end{bmatrix}
\]  

(23)

Applying once again the aforementioned ideas to (22), we further define the operator

\[
D_{M1} = H_1^{\circ} M_1^r A_{M0} B_1^r J_1 ^\circ + D_{p1}:
\]

\[
(L_\infty(0, h'))^M \rightarrow (L_\infty(0, h'))^M
\]  

(24)

What has been done up to now is that the input and output of \( M_1 F(P_d, \Psi) B_1 \) in (15) are approximated by piecewise linear functions, similar treatment has been done on the first term on the right hand side of (22), and the second term of (22) was approximated by \( D_{p1} \). The last treatment has followed a similar technique in [21], [22]. To summarize, we have introduced the following approximation of \( F_\infty(P, \Psi) \):

\[
F_{M1}(\hat{P}, \Psi) := H_{M1} M_1 F(P_d, \Psi) B_1 J_{M1} + D_{M1}
\]  

(25)

We call it piecewise linear approximation of the sampled-data system \( \Sigma_{BD} \), which alleviates the difficulty in the computation of \( \| F_\infty(P, \Psi) \| \).

**V. MAIN RESULTS**

This section is devoted to showing that the error in the piecewise linear approximation converges to 0 at the rate of \( 1/M^2 \) as \( M \rightarrow \infty \). To evaluate the error in the approximation of \( \| F_\infty(P, \Psi) \| = \| F_\infty(\hat{P}, \Psi) \| \) by \( \| F_{M1}(\hat{P}, \Psi) \| \), we first introduce

\[
F_{M1}^0(\hat{P}, \Psi) := H_{M1} M_1 F(P_d, \Psi) B_1 L_{-1}^j
\]

\[
= F_{M1}(\hat{P}, \Psi) - L_1 D_{11} L_{-1}^j
\]  

(26)

\[
F_{M1}(\hat{P}, \Psi) := H_{M1} M_1 F(P_d, \Psi) B_1 J_{M1}
\]

\[
= F_{M1}(\hat{P}, \Psi) - D_{M1}
\]  

(27)

which correspond to `finite-rank portions' of (15) and (25), respectively. Hence, we see that the evaluating \( J_{M1} - L_{-1}^j \) and \( H_{M1} - L_{M1} \) is important in the error analysis. The following lemma is relevant to such evaluation and plays a key role in our discussions.

**Lemma 1:** Suppose that \((A, B_1)\) is controllable, \((C_0, A_2)\) is observable, \( M \geq n + n_u \) and that \( h/M \) is a non-pathological sampling period with respect to \( A_2 \). Then, we have the following properties regarding the predajoints \( J_{M1} \) and \( B_1 \) and the operators \( H_{M1} \) and \( M_1 \).

a) There exists an operator \( (J_{M1})^{-L} : \mathbb{R}(J_{M1} B_1) \rightarrow (L_1(0, h'))^{n_u} \) and a constant \( K_B \) such that \( \| (J_{M1})^{-L} J_{M1} B_1 \| \leq K_B \) and

\[
\| (I - L_1 M_1 (J_{M1})^{-L}) (R(J_{M1} B_1))^{n_u} \| \leq K_B \frac{M^2}{M^2}
\]  

(28)

(b) There exists an operator \( (H_{M1})^{-L} : \mathbb{R}(H_{M1} M_1) \rightarrow (L_\infty(0, h'))^{M n_u} \) and a constant \( K_C \) such that \( \| (H_{M1})^{-L} H_{M1} (R(M_1))^{M n_u} \| \leq K_C \) and

\[
\| (I - L_1 M_1 (H_{M1})^{-L}) (R(H_{M1} M_1))^{M n_u} \| \leq K_C \frac{M^2}{M^2}
\]  

(29)

Remark 3: The two norms \( \| \cdot \|_1 \) and \( \| \cdot \| \) in Lemma 1 mean the \( L_1(0, h') \)-induced norm and the \( L_\infty(0, h') \)-induced norm, respectively. From the definition of the predajoint in Section II, \( J_{M1} : (L_1(0, h'))^{n_u} \rightarrow (L_1(0, h'))^{M n_u} \) is given by

\[
J_{M1} := J_{1*} L_M
\]  

(30)

where the predajoint \( J_{1*} \) is given by

\[
(J_{1*} w)(\theta') = f_0(\theta') \int_0^{h'} w(\tau')d\tau' + f_1(\theta') \int_0^{h'} \tau' w(\tau')d\tau'
\]  

(31)
Remark 4: If we note (30), it is not hard to see that the claim (28) can be roughly restated as the assertion that the input of $\mathcal{B}_1$ could be approximated by a piecewise linear function with $M$ segments, causing only slight affections on its output, because of the ‘low pass’ nature of the integral operator $\mathcal{B}_1$. The claim (29) also has a similar interpretation. Regarding rigorous proof of Lemma 1, we mostly follow a result concerned with the FSFH (or piecewise constant) approximation technique [20]. In [11], however, integral inequalities are used to establish the associated convergence rate (see (18), (19) and (20) in [11] for details). Since these integral inequalities cannot be used for establishing (28) and (29), we use instead a Taylor expansion technique for the proof of Lemma 1 in a similar way to [21], [22]. The details are given in Appendix.

Remark 5: The controllability and observability assumptions in Lemma 1 are only for the ease in the proof, and can in fact be removed. This is because we can always replace these pairs with controllable and observable ones, without changing the ranges $\mathcal{R}(\mathcal{B}_1)$ and $\mathcal{R}(\mathcal{M}_1)$.

We have the following important result from Lemma 1.

Proposition 1: Suppose that $M \geq n + n_u$ and $h/M$ is a non-pathological sampling period with respect to $A_2$. Then, there exists a constant $K_0$ independent of $\psi$, such that
\[
\|\mathcal{F}_M(\tilde{P}, \psi) - \mathcal{F}_0(\tilde{P}, \psi)\| \leq \frac{K_0}{M^2} \|\mathcal{F}_0(\tilde{P}, \psi)\|
\] (32)

Proof: We first deal with the approximation on the output side. Let us introduce the following ‘finite-rank portion’ of $\mathcal{F}(\tilde{P}, \psi)$:
\[
\mathcal{F}_0(\tilde{P}, \psi) := \mathcal{M}_1 \mathcal{F}(P_d, \psi) \mathcal{B}_1
\] (33)

From the second assertion of Lemma 1, we have
\[
\|\mathcal{H}_M \mathcal{F}_0(\tilde{P}, \psi) \mathcal{L}_M^{-1} - \mathcal{F}_M(\tilde{P}, \psi)\|
= \|\mathcal{H}_M \mathcal{F}(P_d, \psi) \mathcal{B}_1 \mathcal{L}_M^{-1} - \mathcal{L}_M \mathcal{M}_1 \mathcal{F}(P_d, \psi) \mathcal{B}_1 \mathcal{L}_M^{-1}\|
= \|(I - \mathcal{L}_M \mathcal{H}_M^{-1}) \mathcal{M}_1 \mathcal{F}(P_d, \psi) \mathcal{B}_1 \mathcal{L}_M^{-1}\|
\leq \frac{K_C}{M^2} \|\mathcal{H}_M \mathcal{F}_0(\tilde{P}, \psi) \mathcal{L}_M^{-1}\|
\] (34)

We next deal with the approximation on the input side. It follows from the first assertion of Lemma 1 that
\[
\|\mathcal{F}_M(\tilde{P}, \psi) - \mathcal{H}_M \mathcal{F}_0(\tilde{P}, \psi) \mathcal{L}_M^{-1}\|
= \|\mathcal{H}_M \mathcal{M}_1 \mathcal{F}(P_d, \psi) \mathcal{B}_1 \mathcal{L}_M^{-1} - \mathcal{H}_M \mathcal{M}_1 \mathcal{F}(P_d, \psi) \mathcal{B}_1 \mathcal{L}_M^{-1}\|
= \|(I - \mathcal{M}_1 \mathcal{H}_M^{-1}) \mathcal{H}_M \mathcal{M}_1 \mathcal{F}(P_d, \psi) \mathcal{B}_1 \mathcal{L}_M^{-1}\|
\leq \frac{K_B}{M^2} \|\mathcal{F}_0(\tilde{P}, \psi)\|
\] (35)

In particular, this implies the following inequality regarding the right hand side of (34).
\[
\|\mathcal{H}_M \mathcal{F}_0(\tilde{P}, \psi) \mathcal{L}_M^{-1}\| \leq \left(1 + \frac{K_B}{M^2}\right) \|\mathcal{F}_0(\tilde{P}, \psi)\|
\] (36)

Combining (34), (35) and (36) leads to (32) with $K_0 := K_C(1 + K_B) + K_B$. This completes the proof.

In view of (26) and (27), it is also important to evaluate $\mathcal{D}_M - \mathcal{L}_M \mathcal{D}_1 \mathcal{L}_M^{-1}$, for which we have the following result.

Lemma 2 ([21], [22]): The inequality
\[
\|\mathcal{D}_M - \mathcal{L}_M \mathcal{D}_1 \mathcal{L}_M^{-1}\| \leq \frac{K_1}{M^2}
\] (37)

holds with $K_1$ defined as
\[
K_1 := \frac{1}{2} \|C_1\| \cdot \|A\| \cdot \|B_1\| e^{2\|A\| h} (1 + e^{\|A\| h})
+ \frac{1}{2} \|C_1\| \cdot \|A\| \cdot \|B_1\| e^{2\|A\| h} (38)
\]

We are in a position to give the following main result on the error analysis of piecewise linear approximation.

Theorem 2: Suppose that $M \geq n + n_u$ and $h/M$ is a non-pathological sampling period with respect to $A_2$. Then,\[
\left(1 - \frac{K_0}{M^2}\right) \|\mathcal{F}_M(\tilde{P}, \psi)\| - \frac{K_1}{M^2} \|\mathcal{F}(\tilde{P}, \psi)\| \leq \left(1 + \frac{K_0}{M^2}\right) \|\mathcal{F}_M(\tilde{P}, \psi)\| + \frac{K_1}{M^2}
\] (39)

Proof: A key in the proof is to show that $\|\mathcal{F}_M(\tilde{P}, \psi)\| \leq \|\mathcal{F}_M(P, \psi)\|$. This inequality follows from the properties of $L_\infty[h, h']$ if we note that the infinite (Toeplitz) matrix representation of the input/output relation of $\mathcal{F}_0(\tilde{P}, \psi) = \mathcal{H}_M \mathcal{M}_1 \mathcal{F}(P_d, \psi) \mathcal{B}_1 \mathcal{J}_M$ is strictly block lower triangular (with respect to the partitioning associated with $\tilde{w}_k$ and $\tilde{z}_k$) because of the structure of $P_d$ (note that (12) has no direct feedthrough matrix between $\tilde{w}_k$ and $\tilde{z}_k$; this infinite matrix obviously has no overlap of nonzero entries with the infinite matrix representation of $\mathcal{F}_M(\tilde{P}, \psi) - \mathcal{F}_0(\tilde{P}, \psi) = \mathcal{D}_M$ (which is nothing but the infinite block diagonal matrix with all diagonal entries given by $\mathcal{D}_M$)).

It follows from Proposition 1 and Lemma 2 that
\[
\|\mathcal{F}_M(\tilde{P}, \psi)\| \leq \|\mathcal{F}_M(\tilde{P}, \psi)\| + \frac{K_0}{M^2} \|\mathcal{F}_0(\tilde{P}, \psi)\| + \frac{K_1}{M^2}
\] (40)

Since $\|\mathcal{F}_M(\tilde{P}, \psi)\| = \|\mathcal{F}(\tilde{P}, \psi)\|$, the assertion follows immediately.

There exists $M_0 \in \mathbb{N}$ such that $M \geq n + n_u$ and $h/M$ is a non-pathological sampling period for all $M \geq M_0$. Hence, (39) holds for all $M \geq M_0$, so that Theorem 2 gives a theoretical basis for such an $L_1$ optimal controller synthesis method for $\Sigma_{SD}$ that seeks for $\psi$ minimizing $\|\mathcal{F}_M(\tilde{P}, \psi)\|$ for a sufficiently large $M$. To see this, let\[
\gamma_{opt} := \inf \|\mathcal{F}(\tilde{P}, \psi)\|
\] (41)

and take an $M$. Suppose that $\psi_M$ is an $\varepsilon$-suboptimal controller with respect to $\|\mathcal{F}_M(\tilde{P}, \psi)\|$, i.e., $\|\mathcal{F}_M(\tilde{P}, \psi_M)\| \leq \gamma_{opt} + \varepsilon$. Then, $\|\mathcal{F}_M(\tilde{P}, \psi_M)\| > \gamma_{opt}$, and the desired $\varepsilon$-suboptimality is ensured.
\[ \gamma_M + \epsilon (\epsilon > 0), \text{ where } \gamma_M := \inf_{\Psi} \| F_{M1}(\hat{P}, \Psi) \|. \] It follows from the second inequality of (39) that

\[ \gamma_{opt} \leq \| F(\hat{P}, \Psi_M) \| \leq \left( 1 + \frac{K_0}{M^2} \right) \| F_{M1}(\hat{P}, \Psi_M) \| + \frac{K_1}{M^2} \]  

(42)

Let \( M_1 \in \mathbb{N} \) be the minimum such that \( M_1 \geq M_0 \) and \( M_1^2 > K_0 \). Then, for \( M \geq M_1 \), the first inequality of (39) implies that

\[ \left( 1 - \frac{K_0}{M^2} \right) \gamma_M - \frac{K_1}{M^2} \leq \gamma_{opt} \]  

(43)

and thus \( \gamma_M \leq (M^2 \gamma_{opt} + K_1)/(M^2 - K_0) \). Substituting this into (42), we see that

\[ \gamma_{opt} \leq \| F(\hat{P}, \Psi_M) \| \leq \gamma_{opt} + \epsilon + \frac{X}{M^2} \]  

(44)

where

\[ X := \frac{2K_0 \gamma_{opt}}{1 - K_0/M_1^2} + \frac{2K_1}{1 - K_0/M_1^2} + K_0 \epsilon \]  

(45)

Since the suboptimality measure \( \epsilon > 0 \) with respect to \( \| F_{M1}(\hat{P}, \Psi) \| \) can be taken arbitrarily small in designing \( \Psi \), taking a suboptimal \( \Psi_M \) with a sufficiently small \( \epsilon \) and letting \( M \) sufficiently large is ensured to lead to a method for \( L_1 \) optimal controller synthesis for \( \Sigma_{SD} \). By (44), we could say that the convergence of \( \Psi_M \) is in the order of \( 1/M^2 \).

Theorem 2 also validates such an \( L_\infty \)-induced norm computation (i.e., analysis) method for \( \Sigma_{SD} \) that is based on piecewise linear approximation and computes \( \| F_{M1}(\hat{P}, \Psi) \| \) for a sufficiently large \( M \). In other words, when we are to compute the \( L_\infty \)-induced norm \( \| F(\hat{P}, \Psi) \| \) for a given \( \Psi \), its upper and lower bounds can be computed through \( \| F_{M1}(\hat{P}, \Psi) \| \) by (39). Furthermore, \( \| F_{M1}(\hat{P}, \Psi) \| \) converges to \( \| F(\hat{P}, \Psi) \| \) as \( M \) tends to infinity. To see this, for \( M \geq M_1 \), we restate (39) as follows:

\[ \frac{- K_0 \| F(\hat{P}, \Psi) \| + K_1}{M^2(1 + K_0/M_1^2)} \leq \frac{- K_0 \| F(\hat{P}, \Psi) \| + K_1}{M^2(1 + K_0/M^2)} \leq \frac{1}{M^2} \frac{K_0 \| F(\hat{P}, \Psi) \| + K_1}{1 - K_0/M_1^2} \]  

(46)

\[ \frac{1}{M^2} \frac{K_0 \| F(\hat{P}, \Psi) \| + K_1}{1 - K_0/M_1^2} \leq \frac{1}{M^2} \frac{K_0 \| F(\hat{P}, \Psi) \| + K_1}{1 - K_0/M^1} \]  

(47)

For a given \( \Psi \), Eq. (47) obviously implies that \( \| F_{M1}(\hat{P}, \Psi) \| \) converges to \( \| F(\hat{P}, \Psi) \| \) as \( M \) tends to infinity, where the convergence rate is \( 1/M^2 \). Hence, it is confirmed that the piecewise linear approximation given in this paper can be used also for the \( L_\infty \)-induced norm analysis of sampled-data systems.

We finally remark that the arguments of Theorem 2 are quite distinct from existing studies with the FSFH approximation technique. This is because the associated convergence rate for piecewise linear approximation is \( 1/M^2 \) for the fast-lifting parameter \( M \) as mentioned above, while the convergence rate shown for the conventional FSFH method is only \( 1/M \) for the FSFH parameter \( M \).

**VI. CONCLUSION**

In this paper, we developed a theoretical basis for piecewise linear approximation in the \( L_1 \) optimal controller synthesis problem of sampled-data systems. To take into account of intersample behavior, lifted representation of sampled-data systems was adopted. Piecewise linear approximation of the associated operators was then introduced via fast-lifting treatment, in which the sampling interval is divided into \( M \) subintervals with an equal width, and the input and output of sampled-data systems on each subinterval are approximated by piecewise linear functions through the introduction of two types of linearizing operators. We finally demonstrated the validity of the piecewise linear approximation approach by establishing Theorem 2 or the inequality (39) through the arguments of preadjoint operators. More precisely, this inequality was shown to play an important role in the piecewise linear approximation treatment for both analysis and synthesis in the \( L_1 \) optimal control problem of sampled-data systems. Furthermore, it was shown that the convergence rate associated with the piecewise linear approximation is \( 1/M^2 \) while the convergence rate in the conventional FSFH method is only \( 1/M \) with respect to the fast-lifting or FSFH parameter \( M \).

**REFERENCES**


APPENDIX

In this appendix, we give a brief proof of Lemma 1 for simplicity.

a) To show the existence of $(J_{M+1})^{-L}$ on $R(J_{M+1}B_{1+})$, we first show that $(J_{M+1}) : (L_{1}(0, h'))^{Mw} \rightarrow (L_{1}(0, h'))^{Mw}$ has no null space contained in $R(B_{1+})$. To this end, let $g \in R(B_{1+})$, which means that $g(\theta) = B_{1+}^{T} \exp((A\theta)^{T})x$ for some $x \in \mathbb{R}^{n}_{1}$. Then,

$$(J_{M+1}g)(\theta) = \begin{bmatrix} B_{1+}^{T}(A_d\theta)^{M-1}f(A, \theta)^{T} \\ \vdots \\ B_{1+}^{T}(A_0\theta)^{T} \end{bmatrix} x$$

where

$A_d := (A_d)^{T}$

$f(A, \theta) := f_{0}(\theta)A_{0d} + f_{1}(\theta)A_{1d}$

$A_{0d} := \int_{0}^{h'} \exp(A(h' - \tau'))d\tau'$

$A_{1d} := \int_{0}^{h'} \exp(A(h' - \tau'))\tau' d\tau'$(52)

Since $f_{0}(h'/2) = 1/h'$ and $f_{1}(h'/2) = 0$, we have

$$(J_{M+1}(g)(h'/2) = \frac{1}{h'2} \begin{bmatrix} B_{1+}^{T}(A_0\theta)^{T} \end{bmatrix} x = \frac{1}{h'2}B_{M,h'/2}x$$

Note that $(A_d, A_{0d}B_{1+})$ is controllable and thus $B_{M,h'/2}$ has full column rank from the assumptions. Since $J_{M+1}g$ is continuous, this implies that $J_{M+1}$ has no null space in $R(B_{1+})$. This implies that $J_{M+1}$ with its domain restricted to $R(B_{1+})$ (and its codomain restricted to $R(J_{M+1})$) is one to one, and thus so is $J_{M+1}B_{1+}$ (again by the controllability assumption). Hence, for any $J_{M+1}g$ with some $g = B_{1+}x$, $x \in \mathbb{R}^{n}_{1}$, we can uniquely determine the underlying $x$, which in turn implies that the underlying $g$ can be recovered. This ensures the existence of $(J_{M+1})^{-L}$.

As a preliminary step to obtain the bound (28), we first show the existence of a constant $c_{1}$ independent of $M$ such that

$$\|x\|_1 \leq c_{1} \|J_{M+1}B_{1+}x\|_1, \quad \forall x \in \mathbb{R}^{n}_{1}$$

(Note that $T_{BM}$ is formally the same as $B_{1+}J_{M+1}$ except that the domain is not $(L_{1}(0, h'))^{Mw}$ but $(L_{1}(0, h'))^{Mw}$ and the codomain is not $\mathbb{R}^{n}_{1}$ but $\mathbb{R}^{n}_{1}$.) Then, we have

$$T_{BM}J_{M+1}B_{1+} = \sum_{i=1}^{M} \int_{0}^{h'} F_{i}(\theta')F_{i}^{T}(\theta')d\theta' =: W_{BM}$$

where $F_{i}(\theta') := f(A, \theta')\exp(A(M - i)h')B_{1}$. Note that (48) implies that $(J_{M+1}g)(\theta') = F_{i}^{T}(\theta')x$ for $F(\theta') := [F_{1}(\theta'), \cdots, F_{M}(\theta')]$. Furthermore, $W_{BM} = \int_{0}^{h'} F(\theta')F^{T}(\theta')d\theta'$. Hence, $W_{BM}$ is nonsingular for $M \geq n$ by the above arguments. In addition, it can be shown that as $M$ tends to infinity, $W_{BM}$ converges to the controllability Grammian

$$W_{B} := \int_{0}^{h} \exp(A(h - \theta))B_{1}B_{1}^{T}\exp(A(h - \theta))d\theta$$

More precisely, by using the Taylor expansion arguments, we can show the existence of a constant $T_{1}$ independent of $M$ such that

$$\|W_{B} - W_{BM}\|_1 \leq \frac{T_{1}}{M^{2}}$$

because we have

$$\|W_{B} - W_{BM}\|_1 \leq \sum_{i=1}^{M} \int_{0}^{h'} \|G_{i}(\theta')G_{i}^{T}(\theta') - F_{i}(\theta')F_{i}^{T}(\theta')\|_1 d\theta'$$

$$\leq M \int_{0}^{h'} \|\exp(A(h' - \theta')) - f(A, \theta')\|_{1} d\theta'$$

$$\cdot \sup_{0 \leq \theta < h'} \|\exp(A(T(h' - \theta'))\|_1$$

$$+ M \int_{0}^{h'} \|\exp(A(T(h' - \theta')) - f(A, \theta')\|_{1} d\theta'$$

$$\cdot \sup_{0 \leq \theta < h'} \|f(A, \theta')\|_1$$

$$\leq \frac{h^{3}}{M^{2}} e^{2(\|A\| + \|A\|_{1})h}B_{1}B_{1}^{T}\left\|\frac{1}{2} e^{\|A\|h} + 5e^{\|A\|h}A^{2}\right\|$$

$$=: \frac{T_{1}}{M^{2}}$$

(59)
where $G_i(\theta_0) := \exp(A(h_0 - \theta_0)) \exp(A(M - i)h_0)B_1$. Hence, by taking $M$ such that $M^2 > 2T_1^2kW^{-1}$, it follows from [26, Theorem 10.11] that
\[\|W_{BM}^{-1}\|_1 \leq \|W_B^{-1}\|_1 + \|W_{BM}^{-1}\|_1^2 + 2\|W_B^{-1}\|_1^3T_1^2.\] (60)
This clearly implies the existence of $T_2$ independent of $M \geq n + n_u$ such that
\[\|W_{BM}^{-1}\|_1 \leq T_2.\] (61)

On the other hand, again by using the Taylor expansion arguments, we can see that
\[\|T_{BM}w\|_1 \leq \sum_{i=1}^{M} \int_0^{h'} \left\|F_i(\theta')w(i)(\theta')\right\|_1 d\theta'\]
\[\leq \sum_{i=1}^{M} \int_0^{h'} \left\|f(A, \theta')\right\|_1 \exp(A(M - i)h')B_1_1 \left\|w(i)(\theta')\right\|_1 d\theta'\]
\[\leq \sup_{0 \leq \theta' \leq h'} \left\|f(A, \theta')\right\|_1 e^{\|A\|h}B_1 \sum_{i=1}^{M} \int_0^{h'} \left\|w(i)(\theta')\right\|_1 d\theta'\]
\[\leq 10e^{2\|A\|h}B_1_1 \left\|w(\cdot)\right\|_1 =: T_3\|w(\cdot)\|_1\] (62)
Hence, because the inequalities
\[\|x\|_1 = \|W_{BM}^{-1}T_{BM}J_{M_1}B_1x\|_1\]
\[\leq \|W_{BM}^{-1}\|_1 \cdot \|T_{BM}\|_1 \cdot \|J_{M_1}B_1x\|_1\] (63)
are established, we are led to (54) with $c_1 = T_2T_3$.

Finally, let $g \in R(J_{M_1}B_1)$, i.e., $g = J_{M_1}B_1x$, $x \in \mathbb{R}^{n_1}$. Then, since $\|x\|_1 \leq c_1\|g\|_1$ by (54), a direct computation together with the Taylor expansion arguments leads to
\[\|((I - L_{M_1}J_{M_1})^{-L})g\|_1 = \|J_{M_1}B_1x - L_{M_1}B_1x\|_1\]
\[\leq \|(J_{M_1} - I)L_{M_1}B_1x\|_1\]
\[\leq \|B_1\|e^{\|A\|h}M \int_0^{h'} \|f(A, \theta')^T - \exp(A^T(h' - \theta'))\|_1 d\theta' \|x\|_1\]
\[\leq \|B_1\|e^{\|A\|h}M \frac{1}{2} \frac{h^3}{M^2} \|A'\| \cdot \|A\|^2 e^{\|A\|h} \|x\|_1\]
\[\leq \frac{h^3}{2M^2} e^{\|A\|h} \cdot \|A\|^2 \|B_1\| \|g\|_1 =: \frac{K_B}{M^2} \|g\|_1\] (64)
where we used (30). This implies (28) and the proof of part a) is completed.

b) To show the existence of $(H_{M_1})^{-L}$ on $R(H_{M_1}M_1)$, we first show that $(H_{M_1}) : (L_\infty(0, h))^{n_w} \to (L_\infty(0, h'))^{Mn_1}$ has no null space contained in $R(M_1)$. To this end, let $g \in R(M_1)$, i.e., $g(\theta) = C_0 \exp(A_2\theta)p$ for some $p \in \mathbb{R}^{n_1}$. Then,
\[\langle H_{M_1}g(\theta'), \phi \rangle = \left[\begin{array}{c} C_0 h(A, \theta') \\ \vdots \\ C_0(A_2')^{M-1} h(A, \theta') \end{array}\right] p\] (65)
where
\[h(A, \theta') := I + A_2\theta'\] (66)
Thus
\[\langle H_{M_1}g(0), \phi \rangle = \left[\begin{array}{c} C_0 \\ \vdots \\ C_0(A_2')^{M-1} \end{array}\right] \mu := C_{M,0}p\] (67)
Note that $(C_0, A_2')$ is observable from the assumptions. Since $H_{M_1}g$ is continuous, this implies that $H_{M_1}$ has no null space in $R(M_1)$.

As a preliminary step to obtain the bound (29), we show the existence of a constant $c_\infty$ independent of $M$ such that
\[\|p\| \leq c_\infty \|H_{M_1}M_1 p\|, \quad \forall p \in \mathbb{R}^{n_1+n_u}\] (68)
To show this, we introduce the operator $T_{CM} : (L_\infty(0, h'))^{Mn_1} \to \mathbb{R}^{n_1+n_u}$ given by
\[T_{CM} z = \sum_{i=1}^{M} h^T(A, \theta') \exp(A_2^{T}(i - 1)h') C_0 e^{A_2\theta'} d\theta'\] (69)
Then, we have
\[T_{CM}H_{M_1}M_1 = \sum_{i=1}^{M} \int_0^{h'} H_1^T(\theta') H_1(\theta') d\theta' =: W_{CM}\] (70)
where $H_1(\theta') := C_0(A_2')^{i-1}h(A, \theta')$. By using the similar arguments to part a), we see that $W_{CM}$ is nonsingular for $M \geq n + n_u$ from the above arguments, and we can also show the existence of a constant $L_1$ independent of $M$ such that
\[\|W_C - W_{CM}\| \leq \frac{L_1}{M^2}\] (71)
where $W_C$ is the observability Gramian
\[W_C := \int_0^{h} \exp(A_2^T \theta)C_0^T C_0 \exp(A_2 \theta) d\theta\] (72)
The remaining part of the proof proceed in a similar way to part a) with $\|\cdot\|_1$ replaced by $\|\cdot\|$.