An averaging result for switched DAEs with multiple modes

Luigi Iannelli ♮, Carmen Pedicini ♮, Stephan Trenn ♦, Francesco Vasca ♯

Abstract—The major motivation of the averaging technique for switched systems is the construction of a smooth average system whose state trajectory approximates in some sense the state trajectory of the switched system. Averaging of dynamic systems represented by switched ordinary differential equations (ODEs) has been widely analyzed in the literature. The averaging approach can be useful also for the analysis of switched differential algebraic equations (DAEs). Indeed by analyzing the evolution of the switched DAEs state it is possible to conjecture the existence of an average model. However a trivial generalization of the ODE case is not possible due to the presence of state jumps. In this paper we discuss the averaging approach for switched DAEs and an approximation result is derived for homogeneous switched linear DAE with periodic switching signals commuting among several modes. This approximation result extends a recent averaging result for switched DAEs with only two modes. Numerical simulations confirm the validity of the averaging approach for switched DAEs.

I. INTRODUCTION

Averaging theory is an useful approach to analyze nonlinear systems. The basic idea of the classical averaging theory exploits the time-scale separation between the time variations of the state of a dynamical system and the time variations of the derivative of the state [1]. For a wide class of switched systems, an explicit formulation of averaging is possible when the switching signal, which selects the state derivative, is faster than the continuous state space variables that can be considered as slow variables. Then the system can be approximated by a model consisting only of the slow continuous states, that is, an average model [2], [3]. Averaging for switched systems is a research topic where different approaches and points of view related to the switched system characteristics have been studied: non-periodic switching functions [4], [5], pulse modulations [6], dithering [7], effects of exogenous inputs [8], hybrid systems framework [9]. The paper [10] presents an overview on the averaging results for switched systems which commute among modes each representable by means of possibly nonlinear ordinary differential equations. Averaging of fast switching systems is also an effective technique used in many engineering applications [11]. However in modeling switched systems a representation by means of switching ordinary differential equations (ODEs) might limit the description of switched systems behavior. For instance a switched system characterized by modes with different algebraic constraints, which may imply state jumps at the switching time instants, cannot be represented by means of switched ODEs. In this case one can use a representation through switched differential algebraic equations (DAEs) [12]–[14]. Also in the case of switched DAEs an average model can be introduced and under certain assumption an approximation result can be proved. In [15], by considering the particular case of a switched DAE with two modes, making the assumption that the consistency projectors commute, we have shown that an average model exists and that solutions of the switched DAE converge to the average solution when the switching frequency increases. In this paper we extend the averaging result presented in [15] by considering switched DAEs with multiple modes. The extension is nothing but trivial because the presence of multiple consistency projectors complicates the analysis and makes the proof much more involved.

II. SWITCHED DIFFERENTIAL ALGEBRAIC EQUATIONS

A switched linear differential algebraic equation (switched DAE) is a system consisting of a family of linear DAEs and a policy that at each time instant selects the active subsystem among a set of possible modes. The selection policy is usually described by means of a switching function. In this paper we consider homogeneous linear switched DAEs of the form [12]

\[ E_\sigma \dot{x} = A_\sigma x \] (1)

with initial condition \( x(t_0) = x_0 \). The switching function \( \sigma(t) : [0, \infty) \to \Sigma \) is a piecewise constant right-continuous function, that selects at each time instant \( t \) the index of the active mode from the finite index set \( \Sigma = \{1, 2, \ldots, P\} \) and \( E_i, A_i \in \mathbb{R}^{n \times n} \) are constant matrices for each \( i \in \Sigma \).

A. The quasi Weierstrass form and the consistency projectors

For each mode, the system (1) can be represented by means of the following non-switched DAE

\[ E \dot{x} = Ax \] (2)

with \( E \) and \( A \in \mathbb{R}^{n \times n} \) and differentiable solutions \( x(t) : \mathbb{R}_+ \to \mathbb{R}^n \). When the matrix \( E \) is invertible, (2) reduces to a more familiar ordinary differential equation.

Assume that the matrix pair \((E, A)\) is regular, i.e. \( \det(sE - A) \) is not the zero polynomial. Then there exist invertible transformation matrices \( S, T \in \mathbb{R}^{n \times n} \) that put the matrices in the quasi Weierstrass form [16]

\[ (SET, SAT) = \left( \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} \right) \] (3)

where \( N \in \mathbb{R}^{n_2 \times n_2}, \) with \( 0 \leq n_2 \leq n \), is a nilpotent matrix, \( J \in \mathbb{R}^{n_1 \times n_1} \), with \( n_1 = n - n_2 \), is some matrix.
and $I \in \mathbb{R}^{n_1 \times n_1}$ is an identity matrix. The transformation matrices $T$ and $S$ can be obtained through the so called Wong sequences [17]. The form (3) consists of two independent parts: an “ODE part” and a “pure DAE part” [16]. This leads to the definition of the so called consistency projectors. The consistency projector $\Pi$ of the matrices pair $(E, A)$ is defined as

$$
\Pi = T \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T^{-1},
$$

(4)

where the block sizes correspond to the block size in the quasi Weierstrass form (3). The consistency projector characterizes the space within all solutions of (2) evolve, i.e. the consistency space is $\text{im} \Pi$. Moreover the matrix $\Pi$ plays a fundamental role when considering inconsistent initial values as they occur when switching between different DAEs. To describe the DAE solution it is possible to introduce the flow matrix

$$
A_{\text{diff}} = T \begin{bmatrix} J & 0 \\ 0 & 0 \end{bmatrix} T^{-1}.
$$

(5)

Note that, due to the special structure of the consistency projector $\Pi$ and $A_{\text{diff}}$, the following conditions hold

$$
A_{\text{diff}} \Pi = A_{\text{diff}} = \Pi A_{\text{diff}}.
$$

(6)

By using the flow matrix it is possible [18] to introduce an ODE system

$$
\dot{x} = A_{\text{diff}} x,
$$

(7)

and show that each solution of (2) also solves (7).

### B. Solutions of switched DAEs

Consider the switched DAE (1). To ensure the uniqueness of solutions we assume that each matrix pair $(E_i, A_i)$ is regular, and we assume knowledge of the quasi-Weierstrass form (3) with corresponding transformation matrices $T_i, S_i$, consistency projectors $\Pi_i$ and flow matrices $A_{\text{diff}}^i$. Moreover we assume impulse-free solutions for any switching signal, which can be characterized [12] by the condition

$$
E_j (I - \Pi_j) \Pi_i = 0, \forall i, j \in \{1, 2, \ldots, P\}.
$$

(8)

Any solution of each individual DAE $E_i \dot{x} = A_i x$ evolves within the consistency space starting from the time instant $t_i$ in which the $i$th mode has been activated. At the switching time $t_i$, a continuous extension of the solution of the previous mode does not exist in general, because the value $x(t_i^-)$ need not be within the consistency space corresponding to DAE $E_j \dot{x} = A_j x$ active after the switch. Therefore it is necessary to allow for solutions with jumps. Indeed, it can be shown [12] that a jump from an inconsistent to a consistent initial value is uniquely determined by using the consistency projector $\Pi_i$ corresponding to the system $(E_i, A_i)$ activated at the switching time $t_i$:

$$
x(t_i) := x(t_i^+) = \Pi_i x(t_i^-).
$$

(9)

Hence, invoking (7) and (8), the solution $x$ on the interval $[t_i, t_{i+1})$ is given by

$$
x(t) = e^{A_{\text{diff}}^i (t-t_i)} x(t_i) = e^{A_{\text{diff}}^i (t-t_i)} \Pi_i x(t_i^+), \quad t \in [t_i, t_{i+1}).
$$

(10)

Then the solution of the switched DAE (1) can be represented by cascading the solutions in the form (10) corresponding to the sequence of modes.

### III. AVERAGING FOR SWITCHED DAEs

Consider the switched DAE (1) on the time interval $[0, \infty)$ and assume that $\sigma : [0, \infty) \to \{1, 2, \ldots, P\}$ is periodic with period $p$ and, without loss of generality, also assume that it is increasing on each period:

$$
\sigma(t) = \begin{cases} 
1, & t \in [kp, (k+1)p), \\
2, & t \in [(k+1)p, (k+2)p), \\
\vdots & \\
P-1, & t \in [(k + (p-1)p), (k + p)p), \\
P, & t \in [(k + (p-2)p), (k + 1)p),
\end{cases}
$$

with $k \in \mathbb{N}$, $d_p$ is the time duration of the $i$th mode within the period $p$ and $d_i \in (0, 1)$ with $\sum_{i=1}^{P} d_i = 1$ is the duty cycle of the $i$th mode. For switched DAEs a crucial assumption to guarantee the convergence to trajectories of an average model, is that these trajectories must evolve within the intersection of the consistency spaces otherwise the jumps will not converge to zero as the frequency increases. Furthermore, at least one consistency projector must jump into the intersection, otherwise the limit cannot be within the intersection. It turns out that the crucial assumption is commutativity of the consistency projectors [19]:

$$
\Pi_i \Pi_j = \Pi_j \Pi_i \quad \forall i, j \in \{1, 2, \ldots, P\}.
$$

(12)

Let us define the matrix given by the product of all consistency projectors matrices:

$$
\Pi := \Pi_1 \Pi_2 \cdots \Pi_{p-1} \Pi_P.
$$

(13)

Then by applying Lemma 1, in Appendix, one obtains

$$
\text{im} \Pi \subseteq \text{im} \Pi_1 \cap \text{im} \Pi_2 \cdots \cap \text{im} \Pi_{p-1} \cap \text{im} \Pi_P.
$$

(14)

The candidate smooth average system for the approximation of the behavior of the switched system (1) is defined as

$$
\dot{x}_{av} = A_{av} x_{av},
$$

(15)

with initial condition $x_{av}(t_0) = \Pi \cap x_0$ and where

$$
A_{av} := \Pi \cap (A_1 + A_2 + \cdots + A_p) \Pi.
$$

(16)
Note that the following conditions hold
\[ \Pi_\gamma A_{av}\Pi_\gamma = A_{av} \quad (17a) \]
\[ \Pi_\gamma \tilde{A}\Pi_\gamma = A_{av}, \quad (17b) \]
with
\[ \tilde{A} := \Pi_\gamma A_{av}^\text{diff} d_1 + \Pi_\gamma p \ldots \Pi_\gamma A_{av}^\text{diff} \Pi_\gamma p_{d_2} + \ldots + \Pi_\gamma p \Pi_\gamma p_{d_1} - 1 A_{av}^\text{diff} \Pi_\gamma p_{d_1} + A_{av}^\text{diff} \Pi_\gamma p_{d_p}. \quad (18) \]

In the next section it will be shown that the properties (17) are important to prove the approximation result.

IV. MAIN RESULT

Before stating our main result let us recall the “big O notation”.

**Definition 1 (Big O notation):** Consider two functions \( f, g : (0, \infty) \rightarrow \mathcal{V} \), where \( \mathcal{V} \) is some normed vector space with norm \( \| \cdot \| \). We say that \( f(p) \) is an \( \mathcal{O}(g(p)) \) function if, and only if, there exist constants \( \alpha, \bar{p} > 0 \) such that
\[ \| f(p) \| \leq \alpha \| g(p) \| \quad \text{for all } 0 < p \leq \bar{p}. \quad (19) \]

With some abuse of notation in the following we indicate with \( \mathcal{O}(g(p)) \) a generic function which is an \( \mathcal{O}(g(p)) \) function. Then, if \( f_1(p) = O(g(p)) \) and \( f_2(p) = O(g(p)) \) are two different functions, it follows \( f_1(p) - f_2(p) = O(g(p)) \).

We are now ready to state our main result.

**Theorem 1:** Consider the switched DAE system (1) with initial value \( x_0 \in \mathbb{R}^n \) and \( P \) modes satisfying the following assumptions

(i) the switching signal \( \sigma \) is periodic of period \( p \) and given by (11) with \( d_i \in (0, 1) \) \( \forall i = 1, \ldots, P \);

(ii) the matrix pairs \( (E_i, A_i), i = 1, \ldots, P \), are regular with corresponding consistency projectors \( \Pi_i \) and flow matrices \( A_{av}^\text{diff} \);

(iii) the consistency projectors commute, i.e. (12) holds; in particular \( \Pi_i \) fulfills (14);

then for any given \( \tilde{t} > 0 \) and \( x_0 \in \mathbb{R}^n \) the following holds
\[ \| x(t) - x_{av}(\tilde{t}) \| = \mathcal{O}(p) \quad \forall t \in (0, \tilde{t}], \quad (20) \]

where \( x_{av}(t) \) is the solution of (15) with the initial value \( \Pi_\gamma x_0 \).

**Proof:** Consider the arbitrary but fixed time instant \( t^* \in (0, \tilde{t}] \). Choose \( k \in \mathbb{N} \) such that \( t^* = kp + \tau \) for \( \tau \in [0, p] \).

Note that \( t^* > 0 \) implies that \( k > 0 \) for sufficiently small \( p \).

The solution of the switched DAE is then
\[ x(t^*) = \tilde{M}(\tau) \left( e^{A_{av}^\text{diff} dp_1} Pi_1 \cdots e^{A_{av}^\text{diff} dp_2} Pi_2 e^{A_{av}^\text{diff} dp_1} Pi_1 \right)^k x_0, \quad (21) \]

where
\[ \tilde{M}(\tau) := \begin{cases} e^{A_{av}^\text{diff} \Pi_1} & \text{if } 0 \leq \tau \leq dp_1 \\
 e^{A_{av}^\text{diff} (\tau - dp_1)} Pi_2 e^{A_{av}^\text{diff} dp_1} Pi_1 & \text{if } dp_1 < \tau \leq (d_1 + d_2)p \\
 \vdots \end{cases} \]

For a matrix exponential one can write the following relation [20]
\[ e^{Ap} = I + Ap + \mathcal{O}(p^2) \quad (22) \]

By applying (22) to the exponentials in (21) we obtain
\[ x(t^*) = \tilde{M}(\tau)(I + A_{av}^\text{diff} dp_1 + \mathcal{O}(p^2))Pi_1 \cdots (I + A_{av}^\text{diff} dp_2 + \mathcal{O}(p^2))Pi_2 (I + A_{av}^\text{diff} dp_1 + \mathcal{O}(p^2))Pi_1 + \mathcal{O}(p^2) \]
\[ = M(\tau)(\Pi_\gamma + A_{av}^\text{diff} dp_1 + \mathcal{O}(p^2))Pi_1 \cdots (\Pi_\gamma + A_{av}^\text{diff} dp_2 + \mathcal{O}(p^2))Pi_2 + \mathcal{O}(p^2) \]
\[ = M(\tau)(\Pi_\gamma + Ap + \mathcal{O}(p^2))^k x_0 \]
\[ = M(\tau)(\Pi_\gamma + Ap + \mathcal{O}(p^2))^{k-2} N(p)x_0, \quad (23) \]

where \( \tilde{M} \) is defined by (18) and
\[ N(p) := \Pi_\gamma + Ap + \mathcal{O}(p^2) = \Pi_\gamma + O(p), \quad (24) \]

and
\[ M(\tau) := M(\tau)(\Pi_\gamma + Ap + \mathcal{O}(p^2)) = \Pi_\gamma + O(p), \quad (25) \]

where for the (25) the following relations have been used
\[ \Pi_1 + O(\tau) = \Pi_1 + O(p) \]
\[ \text{if } 0 \leq \tau \leq dp_1 \]
\[ \Pi_2 \Pi_1 + O(\tau - dp_2) = \Pi_2 \Pi_1 + O(p) \]
\[ \text{if } dp_1 < \tau \leq (d_1 + d_2)p \]
\[ \vdots \]
\[ \Pi_\gamma + O(\tau - dp_{d_1-1}p - \ldots - dp_1p) = \Pi_\gamma + O(p) \]
\[ \text{if } (d_1 + d_2 + \ldots + dp_{d_1-1})p < \tau < p. \quad (26) \]

Consider now the solution of the average model (15)
\[ x_{av}(t^*) = e^{A_{av}^\text{diff} kp} \Pi_\gamma x_0 = \tilde{M}_{av}(\tau) e^{A_{av}^\text{diff} kp} \Pi_\gamma x_0, \quad (27) \]

with
\[ \tilde{M}_{av}(\tau) := e^{A_{av}^\text{diff} \tau}. \quad (28) \]

By using (22), the state (27) can be written as
\[ x_{av}(t^*) = \tilde{M}_{av}(\tau)(\Pi_\gamma + A_{av}p + \mathcal{O}(p^2))^k x_0 \]
\[ = M_{av}(\tau)(\Pi_\gamma + A_{av}p + \mathcal{O}(p^2))^{k-2} N_{av}(p)x_0, \quad (29) \]

with
\[ M_{av}(\tau) := M_{av}(\tau)(\Pi_\gamma + A_{av}p + \mathcal{O}(p^2)) = \Pi_\gamma + O(p), \quad (30) \]

and
\[ N_{av}(p) := (\Pi_\gamma + A_{av}p + \mathcal{O}(p^2))\Pi_\gamma = \Pi_\gamma + O(p). \quad (31) \]
By choosing $k = O(1/p)$, (33) becomes

$$
\|x(t) - x_{av}(t)\| = O(p). \quad (34)
$$

V. SIMULATION RESULTS

To show the effectiveness of the proposed approach consider the following third order switched DAE system with

$$(E_1, A_1) = \left( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -2 & 0 \\ -1 & 0 & 3 \\ 0 & 0 & 1 \end{bmatrix} \right), \quad (35)
$$

$$(E_2, A_2) = \left( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} -3 & 0 & 0 \\ 0 & -4 & 0 \\ 1 & 0 & 0 \end{bmatrix} \right), \quad (36)
$$

$$(E_3, A_3) = \left( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \right), \quad (37)
$$

The second subsystem is an ODE, so $T_2 = S_2 = \Pi_2 = I$ and $A_{2,\text{diff}} = A_2$. For the first and third mode we have

$$(T_1, S_1) = \left( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right),
$$

$$(T_3, S_3) = \left( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right),
$$

$$\Pi_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Pi_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

$$A_{1,\text{diff}} = \begin{bmatrix} 0 & -2 & 0 \\ -3 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad A_{3,\text{diff}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 1 & 0 \end{bmatrix}.$$

and

$$\Pi_\gamma = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

By choosing $d_1 = 0.5$, $d_2 = 0.3$ and $d_3 = 0.2$ the average matrix is

$$A_{av} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\frac{3}{2} & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} 0 & 0 & 0 \\ -\frac{3}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \end{bmatrix}. \quad (38)$$

The solution of the corresponding switched DAEs described by (35)–(37), and that of the average system (15), (38) are shown in Fig. 1 with a switching period $p = 0.1$ s.

By decreasing the switching period the solution of the switched DAE and that of the average model become close to each other as shown in Fig. 2, where a switching period $p = 0.02$ s is chosen.

Fig. 1. Simulation results with $p=0.1$ s: (a) state–space solutions, (b) time evolution of $x_1$, (c) time evolution of $x_2$, (d) time evolution of $x_3$.

Fig. 2. Simulation results with $p=0.02$ s: (a) state–space solutions, (b) time evolution of $x_1$, (c) time evolution of $x_2$, (d) time evolution of $x_3$.

VI. CONCLUSIONS

The averaging method for linear switched differential algebraic equations (DAEs) with multiple mode has been proposed. By assuming that the consistency projectors commute we define an average model in the case of homogeneous linear switched DAEs with periodic switchings between multiple modes. The main result (Theorem 1) says that the error between the solution of the switched DAE and that of the corresponding average model is of the order of the switching period. The work reported here is a generalization of the averaging result presented in [15] for a switched DAEs with two modes, however it is not a trivial extension due to the presence of the consistency projectors that can not be moved outside the switched DAEs solution formula, as in the case of two modes. Dealing with the averaging theory
for switched DAEs opens many interesting lines of research. For instance it could be of interest to study the average system to infer stability property for the switched DAEs. Other directions of future research could be to investigate if in the case of solutions with impulses the average model is still valid. Also, the definition of an average model for switched DAEs with state-dependent switching functions is an interesting challenge. Finally, it is not clear yet how “necessary” commutativity of the consistency projectors is for the existence of an average system.

APPENDIX

Lemma 1: Consider \( P \) commuting matrices \( M_1, M_2, \ldots, M_p \), with \( M_i \) such that \( M_i^2 = M_i \) for \( i = 1, \ldots, (P - 1) \). Then the following holds

\[
im M_1M_2\ldots M_p = \im M_1 \cap \im M_2 \cap \ldots \cap \im M_p. \tag{39} \]

Proof: “\( \subseteq \)”. Clearly, it holds that

\[
im M_1M_2\ldots M_p \subseteq \im M_1M_2\ldots M_{p-1} \subseteq \im M_1M_2\ldots M_{p-2} \subseteq \ldots \subseteq \im M_1,
\]

by using that the matrix \( M_j \) commutes with \( M_i \) for \( j = 1, \ldots, i \) we have

\[
im M_1\ldots M_{i-1}M_i\ldots M_p = \im M_1\ldots M_{i-1}\im M_i\ldots M_p = \im M_1\ldots M_p \subseteq \im M_1M_2\ldots M_{p-1} \subseteq \ldots \subseteq \im M_p.
\]

This shows

\[
im M_1M_2\ldots M_p \subseteq \im M_1 \cap \im M_2 \cap \ldots \cap \im M_p.
\]

“\( \supseteq \)” Considering \( y \in \im M_1 \cap \im M_2 \cap \ldots \cap \im M_p \) the following holds:

\[
y \in \im M_1 \cap \im M_2 \cap \ldots \cap \im M_p \iff \exists x_1 \in \mathbb{R}^n, \exists x_2 \in \mathbb{R}^n \ldots \ x_p \in \mathbb{R}^n : y = M_1x_1 \wedge y = M_2x_2 \wedge \ldots \wedge y = M_p x_p \\
\iff \exists x_1 \in \mathbb{R}^n, \exists x_2 \in \mathbb{R}^n \ldots \ x_p \in \mathbb{R}^n : M_1x_1 = M_2x_2 = \ldots = M_p x_p = y \\
\iff y = \bullet M_1^2x_1 = \bullet M_1M_2x_2 = \ldots = M_1M_2 \ldots \ M_{i-1}M_i = \ldots \ = M_1M_2 \ldots M_{i-1}M_i \ldots M_{p-1}x_{p-1} \Longrightarrow \exists x \in \mathbb{R}^n : y = M_1M_2\ldots M_p x \\
y \in \im M_1M_2\ldots M_p.
\]

where for \( \bullet \) we used \( M_i^2 = M_i \), with \( i = 1, \ldots, (P - 1) \), and for \( \star \) \( M_{i-1}x_{i-1} = M_ix_i \), with \( i = 2, \ldots, P \). This concludes the proof.

Lemma 2: Let \( \ell(p) \in \mathbb{N} \) such that \( p\ell(p) = O(1) \). Then

\[
\tag{40}
\ell(p) = (\Pi_\cap + O(p))^{\ell(p)} = O(1).
\]

Proof: Since \( \Pi_\cap \) is a projector all finite products of \( \Pi_\cap \) are trivially bounded and a well known result, see e.g. [21, Thm. 3], ensures that we can find a norm \( \| \cdot \| \) on \( \mathbb{R}^n \) such that \( \| \Pi_\cap \|^2 = 1 \) for the induced matrix norm. Due to the equivalence of all norms on finite dimensional space we find a constant \( \alpha_1 \in \mathbb{R} \) such that

\[
\| (\Pi_\cap + O(p))^{\ell(p)} \| \leq \alpha_1 \| (\Pi_\cap + O(p))^{\ell(p)} \| \leq \alpha_1 (1 + O(p))^{\ell(p)} \leq \alpha_1 (1 + \alpha_2 p)^{\alpha_3/p} \leq \alpha_1 e^{2\alpha_3} = O(1)
\]

for sufficiently small \( p \) and where \( \alpha_2, \alpha_3 \in \mathbb{R} \) are chosen accordingly.

Lemma 3: Given the matrices \( \Pi_\cap, A_\cap \) and \( A_{av} \), and the period \( p \) define the function

\[
\tag{41}
\gamma(\ell, p) := \Pi_\cap(\Pi_\cap + A_\cap + O(p^2))^{\ell}A_\cap p
\]

with \( \ell \) a positive integer. The function (41) is \( \ell O(p^2) \) for any integer \( \ell \).

Proof: To prove our statement a recursive approach is used. For \( \ell = 1 \) one has

\[
\| \gamma(1, p) \| = \| \Pi_\cap(\Pi_\cap + A_\cap + O(p^2))^{\ell}A_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2)A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| = \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| = \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \|
\]

where we used (17). Then \( \gamma(1, p) \) is an \( O(p^2) \) function.

It is now possible to proceed with the inductive step. For the average theorem we are interest to the case that \( \ell \) goes to infinity when \( p \) goes to zero. Then it is important to explicitly indicate the possible presence of \( \ell \) terms of a certain order. Then by considering \( \gamma(\ell + 1, p) \) one obtains

\[
\| \gamma(\ell + 1, p) \| = \| \Pi_\cap(\Pi_\cap + A_\cap + O(p^2))^{\ell+1}A_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \| \leq \Pi_\cap(\Pi_\cap + A_{av}p) + O(p^2) \| A_{av}\Pi_\cap p \|
\]

(43)
Then
\[
\|\gamma(\ell+1,p)\| \leq \|\Pi_{\ell}(\Pi_{\ell} + \hat{A}p + O(p^2))^{\ell} \hat{A} \Pi_{\ell} p
+ \Pi_{\ell}(\Pi_{\ell} + A_{av}p + O(p^2))^{\ell} A_{av} \Pi_{\ell} p
- \Pi_{\ell} A_{av} (\Pi_{\ell} + A_{av}p + O(p^2))^{\ell} A_{av} \Pi_{\ell} p
\| + \alpha_2 p^3
\leq \|\gamma(\ell,p)\| + \|\Pi_{\ell}(\Pi_{\ell} + \hat{A}p + O(p^2))^{\ell} \hat{A} \Pi_{\ell} p
+ \|\Pi_{\ell} A_{av} (\Pi_{\ell} + A_{av}p + O(p^2))^{\ell} A_{av} \Pi_{\ell} p\| + \alpha_2 p^3
\leq \|\gamma(\ell,p)\| + \alpha_4 p^2 + \alpha_2 p^3,
\]
(44)
for some constants \(\alpha_2\) and \(\alpha_3\) and where for \(\diamond\) we used Lemma 2. Then from (43) the proof is complete.

**Lemma 4:** Given the matrices \(\Pi_{\ell}, A\) and \(A_{av}\), and the period \(p\) define the function
\[
g(\ell, p) := \Pi_{\ell}[(\Pi_{\ell} + \hat{A}p + O(p^2))^{\ell} - (\Pi_{\ell} + A_{av}p + O(p^2))]\Pi_{\ell}
\]
with \(\ell\) a positive integer. The function (45) is \(\ell O(p^2)\) for any integer \(\ell\).

**Proof:** To prove our statement a recursive approach is used. For \(\ell = 1\) one has
\[
\|g(1,p)\| = \|\Pi_{\ell}[(\Pi_{\ell} + \hat{A}p + O(p^2))^{\ell} - (\Pi_{\ell} + A_{av}p + O(p^2))]\Pi_{\ell}\|
\leq \|\Pi_{\ell} \hat{A} \Pi_{\ell} - \Pi_{\ell} A_{av} \Pi_{\ell}\| p + \alpha_4 p^2,
\]
(46)
for some constant \(\alpha_4\). Then by using (17) it is verified that \(g(1, p)\) is an \(O(p^2)\) function. It is now possible to proceed with the inductive step. For the average theorem we will be interested to the case that \(\ell\) goes to infinity when \(p\) goes to zero. Then it is important to explicitly indicate the possible presence of \(\ell\) terms of a certain order. Then by considering \(g(\ell+1, p)\) one obtains
\[
\|g(\ell+1, p)\| = \|\Pi_{\ell}[(\Pi_{\ell} + \hat{A}p + O(p^2))^{\ell+1} - (\Pi_{\ell} + A_{av}p + O(p^2))^{\ell+1}]\Pi_{\ell}\|
\leq \|\Pi_{\ell} \hat{A} \Pi_{\ell} - \Pi_{\ell} A_{av} \Pi_{\ell}\| p + \alpha_4 p^2
= \|g(\ell, p)\| + \|\gamma(\ell, p)\| + \alpha_6 p^2,
\]
(47)
for some constant \(\alpha_6\). Then from (47) and by using the fact that \(\gamma(\ell, p)\) is \(\ell O(p^2)\) the proof is complete.