Construction Methods of the Nearest Positive System

Kazuhiro Sato and Akiko Takeda

***Abstract*—Positive systems can be used as mathematical models for many practical systems, such as biological systems, communication networks, and interconnected systems. In this letter, we propose proximal alternating linearized minimization (PALM) and PALM-like algorithms to determine the nearest discrete-time linear positive system to a given system, with the same order as that of the considered system. Global convergence of the PALM algorithm to a critical point of the considered objective function is ensured by using the Kurdyka–Łojasiewicz and semi-algebraic properties. Numerical experiments are performed to compare the PALM and PALM-like algorithms.**

***Index Terms*—Identification, optimization, positive system, proximal alternating linearized minimization.**

I. Introduction

MANY practical systems such as biological systems [1], [2], communication networks [3], [4], and interconnected systems [5], [6] can be modeled as positive systems [7]. Although such systems may in fact be nonlinear positive systems [8], [9], linear positive systems can be considered as a first approximation. Therefore, theoretical results pertaining to linear positive systems have been established. In fact, we can find the existing results on positive realization problems [10], [11], controllability [12], [13], observability [11], [14], observer synthesis [14], robust control [15]–[17], and decentralized control [18], [19]. Nevertheless, as indicated in [20], data-driven modeling methods—which have become a popular research topic in recent years [21], [22]—for such systems have not been developed satisfactorily. In particular, it is difficult to identify a linear positive system by using the existing identification methods [23]–[27].

Manuscript received February 20, 2019; revised April 29, 2019; accepted May 27, 2019. Date of publication June 10, 2019; date of current version June 21, 2019. This work was supported by Japan Society for the Promotion of Science KAKENHI under Grant 18K13773, Grant 15K00031, and Grant 19H04069. Recommended by Senior Editor R. S. Smith. *(Corresponding author: Kazuhiro Sato.)*

K. Sato is with the Division of Information and Communication Engineering, Kitami Institute of Technology, Hokkaido 090-8507, Japan (e-mail: ksato@mail.kitami-it.ac.jp).

A. Takeda is with the Department of Creative Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan, and also with the RIKEN, Center for Advanced Intelligence Project, Tokyo 103-0027, Japan (e-mail: takeda@mist.i.u-tokyo.ac.jp).

Digital Object Identifier 10.1109/LCSYS.2019.2921838

Thus, in this letter, we propose methods for the establishment of a nearest discrete-time linear positive system to a given discrete-time linear system

$$\begin{cases} x_{t+1} = \tilde{A}x_t + \tilde{B}u_t, \\ y_t = \tilde{C}x_t, \end{cases} \tag{1}$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$, and $y_t \in \mathbb{R}^p$ respectively denote the state, input, and output vectors; and $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ are constant matrices with corresponding sizes over the set of real numbers $\mathbb{R}$. The primary objective is to best approximate system (1), which is not necessarily a positive system, by an internally positive system

$$\begin{cases} x_{t+1} = Ax_t + Bu_t, \\ y_t = Cx_t, \end{cases} \tag{2}$$

where A, B, and C are constant matrices with appropriate sizes over $\mathbb{R}$. Here, system (2) is termed internally positive if for any $x_0 \geq 0$ and $u_t \geq 0$, $t = 0, 1, \dots$, the solution x_t, $t = 0, 1, \dots$, to (2) and the output y_t, $t = 0, 1, \dots$, are nonnegative. In [28], it has been shown that system (2) is internally positive if and only if $A \geq 0$, $B \geq 0$, and $C \geq 0$.

A few relevant existing studies are as follows: In [29], the authors proposed scalable identification methods for stable discrete-time internally positive systems under the assumption that the input, output, and full state data can be obtained. As mentioned in [29], the assumption is rather restrictive, because the state data may not be consistent with a positive realization. Moreover, although identification methods for discrete-time externally positive systems were proposed in [20], several of the results provided in [5], [7], [15], [17]–[19] for internally positive systems cannot be used. Here, system (2) is called externally positive if $u_t \geq 0$ implies $y_t \geq 0$ for $t = 0, 1, \dots$ Clearly, an internally positive system is externally positive. Furthermore, in [30], the author developed a construction method for finding the nearest stable Metzler matrix by using the dissipative Hamiltonian theory; that is, a continuous-time case with no input and output was considered. However, this case does not correspond to the discrete-time case with input and output considered in this letter.

The contributions of this letter can be summarized as follows.

- We provide a formulation of the minimization problem for finding the nearest internally positive system of the same order as the given system, and then establish a simpler problem.

- We develop a proximal alternating linearized minimization (PALM) [31] algorithm for the established simpler problem, and subsequently propose a PALM-like algorithm for the general problem.
- For the PALM algorithm, global convergence to a critical point of the considered objective function is ensured by using the Kurdyka–Łojasiewicz (KL) and semi-algebraic properties.

The remaining paper is organized as follows. The formulation of the problem is presented in Section II. The development of the PALM and PALM-like algorithms for solving the considered problems is described in Section III. Section IV presents the analysis of the convergence of the sequence generated by the PALM algorithm. In Section V, the comparison of the PALM and PALM-like algorithms is described, and the conclusions are presented in Section VI.

Notation: The set of real numbers is denoted by $\mathbb{R}$. We define $\mathbb{R}_+ := \{x \in \mathbb{R} \,|\, x \geq 0\}$. If $A \in \mathbb{R}_+^{n\times m}$, $A \geq 0$. Given matrices $A, B \in \mathbb{R}^{n\times n}$, $\langle A, B\rangle$ is the Euclidean inner product, and $\|A\|_F$ is the Frobenius norm; i.e., $\langle A, B\rangle := \mathrm{tr}\,(A^\top B)$ and $\|A\|_F := \sqrt{\langle A, A\rangle}$, where the superscript $\top$ denotes the transpose and $\mathrm{tr}(A)$ denotes the sum of the diagonal elements of A. The symbols $GL(n)$ and $O(n)$ denote the general linear group and orthogonal group in $\mathbb{R}^{n\times n}$, respectively.

II. Problem Setting

This section describes the formulation of the two problems addressed in this letter.

Even if system (1) is not internally positive,

$$\begin{cases} \bar{x}_{t+1} = T^{-1}\tilde{A}T\bar{x}_t + T^{-1}\tilde{B}u_t, \\ y_t = \tilde{C}T\bar{x}_t \end{cases} \tag{3}$$

may be internally positive, where $x_t = T\bar{x}_t$ and $T \in GL(n)$. In other words, it may be possible to find a basis change matrix $T \in GL(n)$ such that system (3) is internally positive. In this case, there are $A \geq 0$, $B \geq 0$, $C \geq 0$, and $T \in GL(n)$ such that $F(A, B, C, T) = 0$, where

$$F(A, B, C, T) := \frac{1}{2}\|A - T^{-1}\tilde{A}T\|_F^2 + \frac{\gamma_1}{2}\|B - T^{-1}\tilde{B}\|_F^2 + \frac{\gamma_2}{2}\|C - \tilde{C}T\|_F^2. \tag{4}$$

Here, $\gamma_1 > 0$ and $\gamma_2 > 0$ are fixed parameters that represent the weights of each term. However, in general, $A \geq 0$, $B \geq 0$, $C \geq 0$, and $T \in GL(n)$ may not necessarily exist. Therefore, to construct the nearest positive system for all cases, we consider the following non-convex optimization problem.

Problem 1:

$$\begin{aligned} &\text{minimize} \quad F(A, B, C, T) \\ &\text{subject to} \quad A \geq 0,\ B \geq 0,\ C \geq 0,\ T \in GL(n). \end{aligned}$$

Note that the objective function $F(A, B, C, T)$ is not defined on $(\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}) \times \mathbb{R}^{n\times n}$, although it is defined on $(\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}) \times GL(n)$. This aspect is the reason for the difficulty in developing an efficient algorithm for solving Problem 1, as further explained in Section III-B.

To overcome this limitation, we replace the constraint $T \in GL(n)$ in Problem 1 with $U \in O(n)$, in which the orthogonal group $O(n)$ is the maximal compact subgroup of $GL(n)$. The modified problem is equivalent to the following problem.

Problem 2:

$$\begin{aligned} &\text{minimize} \quad \tilde{F}(A, B, C, U) \\ &\text{subject to} \quad A \geq 0,\ B \geq 0,\ C \geq 0,\ U \in O(n). \end{aligned}$$

Here, $\tilde{F}(A, B, C, U) := \frac{1}{2}\|A\|_F^2 - \langle A, U^\top\tilde{A}U\rangle + \gamma_1(\frac{1}{2}\|B\|_F^2 - \langle B, U^\top\tilde{B}\rangle) + \gamma_2(\frac{1}{2}\|C\|_F^2 - \langle C, \tilde{C}U\rangle)$. The equivalence follows from $F(A, B, C, U) = \tilde{F}(A, B, C, U) + \frac{1}{2}(\|\tilde{A}\|_F^2 + \gamma_1\|\tilde{B}\|_F^2 + \gamma_2\|\tilde{C}\|_F^2)$ subject to $U \in O(n)$, where $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ are constant matrices. Note that $\tilde{F}(A, B, C, U)$ is defined on $(\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}) \times \mathbb{R}^{n\times n}$ in contrast to the case of $F(A, B, C, T)$.

Remark 1: As mentioned already, the parameters γ_1 and γ_2 are the weights of each term in the objective function defined by (4). Thus, if the eigenvalues of A are desired to be as close as possible to those of $\tilde{A}$, it is preferable to choose sufficiently small values of γ_1 and γ_2. This is because the eigenvalues of $\tilde{A}$ and $T^{-1}\tilde{A}T$ coincide.

Remark 2: The search space for Problem 2 is narrower than that for Problem 1. However, Problems 1 and 2 are both non-convex. Consequently, in practice, one can determine only local optimal solutions to Problems 1 and 2. That is, a solution attained using an algorithm for Problem 2 may be better than that for Problem 1, as demonstrated later in Section V.

III. Algorithms for Solving Problems 1 and 2

This section describes two algorithms for solving Problems 1 and 2. Because it is easier to develop an algorithm for Problem 2 than Problem 1, first, a PALM [31] algorithm for Problem 2 is developed, and subsequently, we develop a PALM-like algorithm for Problem 1.

A. PALM Algorithm for Solving Problem 2

To develop the PALM algorithm for solving Problem 2, we reformulate Problem 2 into the following PALM-applicable form by introducing the indicator function

$$\mathcal{I}_S(X) := \begin{cases} 0, & X \in S, \\ \infty, & X \notin S, \end{cases}$$

where S is an arbitrary set.

Problem 3:

$$\begin{aligned} &\text{minimize} \quad \tilde{\alpha}(W, U) := f(W) + \mathcal{I}_{O(n)}(U) + \tilde{H}(W, U) \\ &\text{subject to} \quad W = (A, B, C) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}, \\ &\qquad\qquad\quad U \in \mathbb{R}^{n\times n}. \end{aligned}$$

Here,

$$f(W) := \mathcal{I}_{\mathbb{R}_+^{n\times n}}(A) + \mathcal{I}_{\mathbb{R}_+^{n\times m}}(B) + \mathcal{I}_{\mathbb{R}_+^{p\times n}}(C), \tag{5}$$

and $\tilde{H}(W, U) := \tilde{F}(A, B, C, U)$. Note that f is convex and nonsmooth; $\tilde{H}$ is differentiable in terms of W on $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}$ and U on $\mathbb{R}^{n\times n}$, and their gradients are Lipschitz-continuous, as shown later.

In the PALM algorithm [31], we consider the proximal regularization of the Gauss-Seidel scheme:

$$W_{k+1} \in \underset{W}{\mathrm{argmin}}\ \tilde{\Psi}_1(W, U_k), \tag{6}$$

Algorithm 1 PALM Algorithm for Problem 2

1: Set $(A_0, B_0, C_0, U_0) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n} \times \mathbb{R}^{n\times n}$, and $c_W, c_U > 1$.
2: **for** $k = 0, 1, 2, \ldots$ **do**
3: Set $\eta_k = c_W L_W$, and compute (8), (9), and (10).
4: Set $\mu_k = c_U L_U(W_{k+1})$, and compute (12).
5: **end for**

$$U_{k+1} \in \underset{U}{\operatorname{argmin}}\ \tilde{\Psi}_2(W_{k+1}, U), \tag{7}$$

where $\tilde{\Psi}_1(W, U_k) := \langle \nabla_W \tilde{H}(W_k, U_k), W - W_k\rangle + \frac{\eta_k}{2}\|W - W_k\|_F^2 + f(W)$ and $\tilde{\Psi}_2(W_{k+1}, U) := \langle \nabla_U \tilde{H}(W_{k+1}, U_k), U - U_k\rangle + \frac{\mu_k}{2}\|U - U_k\|_F^2 + \mathcal{I}_{O(n)}(U)$, and $\eta_k, \mu_k > 0$ are the parameters which are determined later. Here, $\nabla_W \tilde{H}(W, U) = (A - U^\top \tilde{A} U, \gamma_1(B - U^\top \tilde{B}), \gamma_2(C - \tilde{C}U))$ and $\nabla_U \tilde{H}(W, U) = -\tilde{A}^\top UA - \tilde{A}UA^\top - \gamma_1 \tilde{B}B^\top - \gamma_2 \tilde{C}^\top C$. The function $\tilde{\Psi}_1(W, U_k)$ is separable in terms of (A, B, C), because $\tilde{\Psi}_1(W, U_k) = \langle A_k - U_k^\top \tilde{A} U_k, A - A_k\rangle + \frac{\eta_k}{2}\|A - A_k\|_F^2 + \mathcal{I}_{\mathbb{R}_+^{n\times n}}(A) + \gamma_1\langle B_k - U_k^\top \tilde{B}, B - B_k\rangle + \frac{\eta_k}{2}\|B - B_k\|_F^2 + \mathcal{I}_{\mathbb{R}_+^{n\times m}}(B) + \gamma_2\langle C_k - \tilde{C}U_k, C - C_k\rangle + \frac{\eta_k}{2}\|C - C_k\|_F^2 + \mathcal{I}_{\mathbb{R}_+^{p\times n}}(C)$. Thus, the optimal solution to (6) is given by

$$A_{k+1} = \max\left(A_k - \frac{1}{\eta_k}(A_k - U_k^\top \tilde{A} U_k), 0\right), \tag{8}$$

$$B_{k+1} = \max\left(B_k - \frac{\gamma_1}{\eta_k}(B_k - U_k^\top \tilde{B}), 0\right), \tag{9}$$

$$C_{k+1} = \max\left(C_k - \frac{\gamma_2}{\eta_k}(C_k - \tilde{C}U_k), 0\right), \tag{10}$$

where the max operation is applied component-wise. Moreover, (7) is equivalent to

$$U_{k+1} \in \underset{U\in O(n)}{\operatorname{argmin}} \frac{\mu_k}{2}\|U - (U_k - \frac{1}{\mu_k}V_k)\|_F^2 \tag{11}$$

with $V_k := \nabla_U \tilde{H}(W_{k+1}, U_k)$. According to [32, Sec. 12.4], if the singular value decomposition of $U_k - \frac{1}{\mu_k}V_k$ is given by $\bar{U}_k \Sigma_k \bar{V}_k^\top$, the solution to (11) can be computed by

$$U_{k+1} = \bar{U}_k \bar{V}_k^\top. \tag{12}$$

Algorithm 1 describes the PALM algorithm for solving Problem 3. Here, L_W and $L_U(W)$ denote the Lipschitz constants for $\nabla_W \tilde{H}(W, U)$ and $\nabla_U \tilde{H}(W, U)$, respectively. That is, L_W satisfies $\|\nabla_W \tilde{H}(W_1, U) - \nabla_W \tilde{H}(W_2, U)\|_F \le L_W\|W_1 - W_2\|_F$ for any $W_1, W_2 \in \mathbb{R}^{n\times n}\times\mathbb{R}^{n\times m}\times\mathbb{R}^{p\times n}$, and $L_U(W)$ satisfies $\|\nabla_U \tilde{H}(W, U_1) - \nabla_U \tilde{H}(W, U_2)\|_F \le L_U(W)\|U_1 - U_2\|_F$ for any $U_1, U_2 \in \mathbb{R}^{n\times n}$. Via direct calculation, we can obtain

$$L_W = \max(1, \gamma_1, \gamma_2), \tag{13}$$

$$L_U(W) = 2\|A\|_F \cdot \|\tilde{A}\|_F. \tag{14}$$

Hence, $\eta_k = c_W L_W$ for each $k = 0, 1, 2, \ldots$ In other words, in practice, instead of calculating η_k at step 3, it is sufficient to calculate η_0 at step 1. Note that η_k and μ_k in Algorithm 1 ensure that the objective function decrease monotonically. The parameters $c_W, c_U > 1$ relate to the convergence speed, as illustrated in Section V.

Remark 3: To solve Problem 2, instead of Algorithm 1, we can also use the proximal alternating projection method (PAPM) proposed in [33]. In this case, similar update formulas on A, B, and C with (8), (9), and (10) can be obtained. However, update formula (11) is replaced with

$$U_{k+1} \in \underset{U\in O(n)}{\operatorname{argmin}}\ h(U), \tag{15}$$

where $h(U) := -\langle A_{k+1}, U^\top \tilde{A} U\rangle - \gamma_1\langle B_{k+1}, U^\top \tilde{B}\rangle - \gamma_2\langle C_{k+1}, \tilde{C}U\rangle + \frac{\mu_k}{2}\|U - U_k\|_F^2$. Unfortunately, it is difficult to obtain a closed-form solution for (15) unlike for (11). Although we can develop a Riemannian optimization algorithm for solving (15) by regarding the orthogonal group $O(n)$ as a Riemannian manifold [34], its use is less efficient than using a closed-form solution. Thus, it can be considered that Algorithm 1 is more efficient than the PAPM for several cases.

B. PALM-Like Algorithm for Solving Problem 1

To develop an algorithm for solving Problem 1, we reformulate the problem as follows.

Problem 4:

$$\begin{aligned} \text{minimize}\quad & \alpha(W, T) := f(W) + \mathcal{I}_{GL(n)}(T) + H(W, T) \\ \text{subject to}\quad & W = (A, B, C) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}, \\ & T \in \mathbb{R}^{n\times n}. \end{aligned}$$

Here, $f(W)$ is defined as in (5), and $H(W, T) := F(A, B, C, T)$. Although H is differentiable in terms of W on $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}$ and T on $GL(n)$, the function is not differentiable in terms of T on $\mathbb{R}^{n\times n}$. In particular, the function H is not defined on $(\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}) \times \mathbb{R}^{n\times n}$. Thus, to solve Problem 4, we cannot adopt the PALM algorithm unlike for Problem 3, because the algorithm requires that H is differentiable on $(\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}) \times \mathbb{R}^{n\times n}$. However, we can develop a PALM-like algorithm as follows.

To develop a PALM-like algorithm, we consider

$$W_{k+1} \in \underset{W}{\operatorname{argmin}}\ \Psi_1(W, T_k), \tag{16}$$

$$T_{k+1} \in \underset{T}{\operatorname{argmin}}\ \Psi_2(W_{k+1}, T), \tag{17}$$

where $\Psi_1(W, T_k) := \langle \nabla_W H(W_k, T_k), W - W_k\rangle + \frac{\eta_k}{2}\|W - W_k\|_F^2 + f(W)$ and $\Psi_2(W_{k+1}, T) := \langle \nabla_T H(W_{k+1}, T_k), T - T_k\rangle + \frac{\mu_k}{2}\|T - T_k\|_F^2 + \mathcal{I}_{GL(n)}(T)$. Here, $\nabla_W H(W, T)$ and $\nabla_T H(W, T)$ denote the gradients of $H(W, T)$ with respect to W on $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}$ and T on $GL(n)$, respectively, and they can be computed by $\nabla_W H(W, T) = (A - T^{-1}\tilde{A}T, \gamma_1(B - T^{-1}\tilde{B}), \gamma_2(C - \tilde{C}T))$ and $\nabla_T H(W, T) = (T^{-1})^\top (A - T^{-1}\tilde{A}T)T^\top \tilde{A}^\top (T^{-1})^\top - \tilde{A}^\top (T^{-1})^\top (A - T^{-1}\tilde{A}T) + \gamma_1 (T^{-1})^\top (B - T^{-1}\tilde{B})\tilde{B}^\top (T^{-1})^\top - \gamma_2 \tilde{C}^\top (C - \tilde{C}T)$. Similar to that for (6), the optimal solution for (16) is given by

$$A_{k+1} = \max\left(A_k - \frac{1}{\eta_k}(A_k - T_k^{-1}\tilde{A}T_k), 0\right), \tag{18}$$

$$B_{k+1} = \max\left(B_k - \frac{\gamma_1}{\eta_k}(B_k - T_k^{-1}\tilde{B}), 0\right), \tag{19}$$

$$C_{k+1} = \max\left(C_k - \frac{\gamma_2}{\eta_k}(C_k - \tilde{C}T_k), 0\right). \tag{20}$$

In addition, (17) is equivalent to

$$T_{k+1} \in \underset{T\in GL(n)}{\operatorname{argmin}}\, g(T), \tag{21}$$

Algorithm 2 Steepest Descent Method for Problem (21)

1: Set an initial point $T^0 \in GL(n)$.
2: **for** $j = 0, 1, 2, \ldots$ **do**
3: Determine the search direction $\xi_j \in T_{T^j}GL(n) \cong \mathbb{R}^{n\times n}$ using $\xi_j = -\mathrm{grad}\, g(T^j)$.
4: Compute the step size $t_j > 0$ and iterate

$$T^{j+1} = T^j \exp(t_j (T^j)^{-1}\xi_j).$$

5: **end for**

Algorithm 3 PALM-Like Algorithm for Problem 1

1: Set $(A_0, B_0, C_0, T_0) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n} \times GL(n)$.
2: **for** $k = 0, 1, 2, \ldots$ **do**
3: Set η_k, μ_k, and compute (18), (19), and (20).
4: **if** $T_k - \frac{1}{\mu_k}\nabla_T H(W_{k+1}, T_k) \in GL(n)$ **then**
5: Compute (22).
6: **else**
7: Compute (21) using Algorithm 2.
8: **end if**
9: **end for**

where $g(T) := \frac{\mu_k}{2}\|T - (T_k - \frac{1}{\mu_k}\nabla_T H(W_{k+1}, T_k))\|_F^2$. Specifically, the update formula for T can be obtained by solving the minimization problem of $g(T)$ on the general linear group $GL(n)$. Algorithm 2 represents the steepest descent method for solving problem (21) using the gradient

$$\mathrm{grad}\, g(T) = \mu_k(T - T_k) - \nabla_T H(W_{k+1}, T_k)$$

and the Lie group exponential map. In particular, because $GL(n)$ is a Lie group and $\mathbb{R}^{n\times n}$ can be regarded as Lie algebra of $GL(n)$, we can use the Lie group exponential map on $GL(n)$ in step 4, where exp denotes the usual matrix exponential map [35]. As the step size $t_j > 0$ at step 4, we can employ the Armijo step size [34]. Note that if $T_k - \frac{1}{\mu_k}\nabla_T H(W_{k+1}, T_k) \in GL(n)$, the optimal solution to (21) is given by

$$T_{k+1} = T_k - \frac{1}{\mu_k}\nabla_T H(W_{k+1}, T_k). \tag{22}$$

In this case, we do not need to use Algorithm 2.

Algorithm 3 describes the PALM-like algorithm for solving Problem 1. The differences between Algorithms 1 and 3 can be summarized as follows:

1) Although the initial point (A_0, B_0, C_0, U_0) in Algorithm 1 can be selected from $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n} \times \mathbb{R}^{n\times n}$, (A_0, B_0, C_0, T_0) in Algorithm 3 must be chosen from $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n} \times GL(n)$. That is, we cannot select T_0 in Algorithm 3 from an arbitrary element in $\mathbb{R}^{n\times n}$ unlike for the case of U_0 in Algorithm 1.
2) The parameters η_k and μ_k in Algorithm 1 were defined by $\eta_k = c_W L_W$ and $\mu_k = c_U L_U(W_{k+1})$ with $c_W, c_U > 1$, because in this case, we can guarantee the global convergence to a critical point of the objection function, as described later in Section IV. In contrast, the parameters in Algorithm 3 were not defined, because the Lipschitz constant corresponding to $L_U(W)$ cannot be defined in the case of Problem 4. As a result, it is difficult to guarantee the global convergence in the case of Algorithm 3.
3) Unlike in the case of Algorithm 1, we may need to use Algorithm 2 at step 8 in Algorithm 3. That is, in general, we must use the nested iterative schemes in the case of Algorithm 3.

Remark 4: Algorithms 1 and 3 can be used to identify discrete-time linear internally positive systems. The procedure can be summarized as follows.

1) Identify $(\tilde{A}, \tilde{B}, \tilde{C})$ for system (1) using the existing identification methods [23]–[27].
2) Construct (A, B, C) satisfying $A \geq 0$, $B \geq 0$, and $C \geq 0$ by using Algorithm 1 or 3.

Step 2 is required to obtain an internally positive system because, in general, $(\tilde{A}, \tilde{B}, \tilde{C})$ generated in step 1 is not internally positive.

IV. Convergence Analysis of Algorithm 1

This section proves the following theorem.

Theorem 1: Assume that $\|\tilde{A}\|_F > 0$, $\inf_k \|A_k\|_F > 0$, and $\{(W_k, U_k)\}$ is a bounded sequence generated by Algorithm 1, where $W_k := (A_k, B_k, C_k)$. Then, the sequence $\{(W_k, U_k)\}$ converges to a (limiting) critical point (W^*, U^*) of the objective function $\tilde{\alpha}$ in Problem 3; i.e., a point (W^*, U^*) satisfies $(0, 0) \in \partial\tilde{\alpha}(W^*, U^*)$, where $\partial\tilde{\alpha}(W, U)$ denotes the (limiting) subdifferential of $\tilde{\alpha}$ at $(W, U) \in (\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times m}\times\mathbb{R}^{p\times n})\times\mathbb{R}^{n\times n}$.

To this end, we need to show that the objective function $\tilde{\alpha}$ satisfies the following KL property [31], [33].

Definition 1: Let $\sigma : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be proper and lower semicontinuous. The function σ is said to have the KL property at $x^* \in \mathrm{dom}\,\partial\sigma := \{x \in \mathbb{R}^n \mid \partial\sigma(x) \neq \emptyset\}$ if there exist $\eta \in (0, \infty]$, a neighborhood U of x^*, and a continuous concave function $\phi : [0, \eta) \to \mathbb{R}_+$ with $\phi(0) = 0$ such that

1) ϕ is C^1 on $(0, \eta)$, and $\phi'(s) > 0$ for all $s \in (0, \eta)$, where ϕ' denotes the derivative of ϕ.
2) for all x in $U \cap \{x \in \mathbb{R}^n \mid \sigma(x^*) < \sigma(x) < \sigma(x^*) + \eta\}$, the KL inequality $\phi'(\sigma(x) - \sigma(x^*)) \inf\{\|y\| \mid y \in \partial\sigma(x)\} \geq 1$ holds.

Moreover, if σ satisfies the KL property at each point of $\mathrm{dom}\,\partial\sigma$, then σ is called a KL function.

According to [31], the following propositions hold.

Proposition 1: Suppose that $\{(W_k, U_k)\}$ is a bounded sequence generated using Algorithm 1. Moreover, suppose that the objective function $\tilde{\alpha}(W, U)$ in Problem 3 is a KL function that satisfies the following:

1) $\inf_{W,U} \tilde{\alpha}(W, U) > -\infty$, $\inf_W f(W) > -\infty$, and $\inf_U \mathcal{I}_{O(n)}(U) > -\infty$.
2) There exists $L_1(U)$ such that $\|\nabla_W \tilde{H}(W_1, U) - \nabla_W \tilde{H}(W_2, U)\|_F \leq L_1(U)\|W_1 - W_2\|_F$ for any $W_1, W_2 \in \mathbb{R}^{n\times n}\times\mathbb{R}^{n\times m}\times\mathbb{R}^{p\times n}$. Similarly, there exists $L_2(W)$ such that $\|\nabla_U \tilde{H}(W, U_1) - \nabla_U \tilde{H}(W, U_2)\|_F \leq L_2(W)\|U_1 - U_2\|_F$ for any $U_1, U_2 \in \mathbb{R}^{n\times n}$. Moreover, there exist positive real numbers λ_1^-, λ_1^+, λ_2^-, and λ_2^+ such that $\inf_k\{L_1(U_k)\} \geq \lambda_1^-$, $\inf_k\{L_2(W_k)\} \geq \lambda_2^-$, $\sup_k\{L_1(U_k)\} \leq \lambda_1^+$, and $\sup_k\{L_2(W_k)\} \leq \lambda_2^+$.
3) The gradient $\nabla\tilde{H}(W, U)$ is Lipschitz continuous on the bounded subsets of $(\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}) \times \mathbb{R}^{n\times n}$.

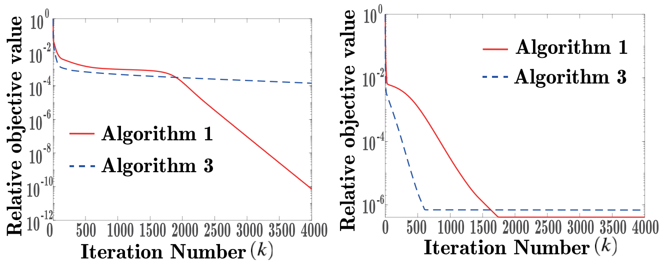

Fig. 1. $\frac{F(A_k,B_k,C_k,U_k)}{F(A_0,B_0,C_0,U_0)}$ obtained using Algorithm 1, and $\frac{F(A_k,B_k,C_k,T_k)}{F(A_0,B_0,C_0,T_0)}$ obtained using Algorithm 3 for Case 1. The left (right) figure shows the result for the case in which the system has an internally positive realization by acting the orthogonal (general linear) group.

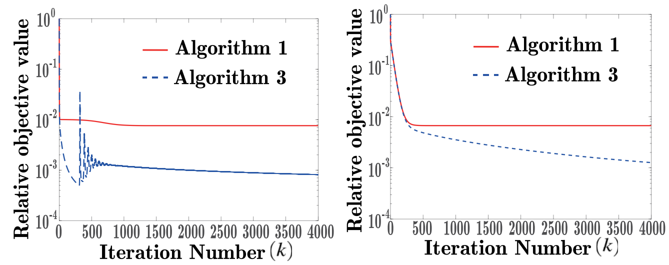

Fig. 2. $\frac{F(A_k,B_k,C_k,U_k)}{F(A_0,B_0,C_0,U_0)}$ obtained using Algorithm 1, and $\frac{F(A_k,B_k,C_k,T_k)}{F(A_0,B_0,C_0,T_0)}$ obtained using Algorithm 3 for Case 2. The left and right figures pertain to the cases in which $(c_W,c_U) = (2,2)$ and $(c_W,c_U) = (100,100)$, respectively.

Then, the sequence $\{(W_k, U_k)\}$ converges to a critical point (W^*, U^*) of $\tilde{\alpha}$.

Proposition 2: Let $\sigma : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a proper and lower semicontinuous function. If σ is semi-algebraic, then it satisfies the KL property at any point of $\operatorname{dom}\sigma$.

Because the objective function $\tilde{\alpha}(W, U)$ is semi-algebraic [31], $\tilde{\alpha}(W, U)$ satisfies the KL property. Moreover, $\tilde{\alpha}(W, U)$ satisfies 1), 2), and 3) in Proposition 1 under the assumption of Theorem 1. In fact, 1) is clear, and 3) follows from the mean value theorem. To see that $\tilde{\alpha}(W, U)$ satisfies 2) in Proposition 1 under the assumption of Theorem 1, note that $L_1(U)$ and $L_2(W)$ are equal to L_W of (13) and $L_U(W)$ of (14), respectively. Thus, $\inf_k\{L_1(U_k)\} \geq 1$, $\sup_k\{L_1(U_k)\} \leq \max(1, \gamma_1, \gamma_2)$, and $\inf_k\{L_2(W_k)\} > 0$. Moreover, the boundedness of $\{(W_k, U_k)\}$ implies that $\{A_k\}$ is a bounded sequence. That is, there exists $a > 0$ such that $\|A_k\|_F < a$ for any k, and thus $\sup_k\{L_2(W_k)\} \leq 2a \cdot \|\tilde{A}\|_F$. Therefore, $\tilde{\alpha}(W, U)$ satisfies 2) in Proposition 1, because we have assumed $\|\tilde{A}\|_F > 0$.

Considering this discussion, Propositions 1 and 2 imply Theorem 1.

V. Numerical Experiments

This section provides the comparison of Algorithms 1 and 3. It should be noted that the parameters η_k and μ_k in Algorithm 3 were defined as $\eta_k = c_W \max(1, \gamma_1, \gamma_2)$ and $\mu_k = 2c_U\|A\|_F \cdot \|\tilde{A}\|_F$, respectively. That is, the parameters in Algorithm 3 were determined in a similar manner as in Algorithm 1. The parameters γ_1 and γ_2 in the objective function were both equal to 1. The initial points (A_0, B_0, C_0) in Algorithms 1 and 3 were equal and chosen randomly. Moreover, we selected $U_0 = T_0 = I_n$.

A. Case 1 (Internally Positive System)

We generated two systems with $(n, m, p) = (4, 2, 1)$, which can be expressed using (1). The first system admits an internally positive realization by acting $O(4)$; i.e., there exists $U \in O(4)$ such that $U^\top \tilde{A} U \geq 0$, $U^\top \tilde{B} \geq 0$, and $\tilde{C} U \geq 0$. The second system admits an internally positive realization by acting $GL(4)$; i.e., there exists $T \in GL(4)$ such that $T^{-1}\tilde{A}T \geq 0$, $T^{-1}\tilde{B} \geq 0$, and $\tilde{C}T \geq 0$.

The left (right) part of Fig. 1 shows the relative objective values attained by Algorithms 1 and 3 in the case of the first (second) system, where (c_W, c_U) in Algorithm 1 were assigned values of $(2, 2)$. According to the left part of Fig. 1, Algorithm 1 produced A, B, C, and U satisfying $F(A, B, C, U) \approx 0$; however, A, B, C, and T generated by Algorithm 3 did not satisfy $F(A, B, C, U) \approx 0$. Moreover, the right figure showed that Algorithm 1 generated a better solution than Algorithm 3 in the iterations. These cases indicate that for Problem 1, Algorithm 3 does not always generate better solutions than Algorithm 1, although the search space of Problem 1 is broader than that of Problem 2. This finding is a result of non-convexity of the considered problem, as mentioned in Remark 2.

B. Case 2 (Externally Positive System)

We set

$$\tilde{A} = 0.9\begin{pmatrix} \cos(\frac{\sqrt{2}}{\pi}) & \sin(\frac{\sqrt{2}}{\pi}) & 0 \\ -\sin(\frac{\sqrt{2}}{\pi}) & \cos(\frac{\sqrt{2}}{\pi}) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} 0.5 \\ 0.6 \\ 1 \end{pmatrix},$$

$$\tilde{C} = \begin{pmatrix} 0.5 & 0.5 & 1 \end{pmatrix}.$$

According to [20], this system is an externally positive system that does not admit an internally positive system.

Fig. 2 shows the relative objective values obtained using Algorithms 1 and 3; the left and right parts pertain to the cases of $(c_W, c_U) = (2, 2)$ and $(c_W, c_U) = (100, 100)$, respectively. According to Fig. 2, Algorithm 3 generated a better solution than Algorithm 1 for this case. In particular, the left and right figures indicate that the oscillation of the sequence generated by Algorithm 3 decreased as c_W and c_U increased; however, the convergence speeds of Algorithms 1 and 3 were low. Moreover, it was verified that the relative objective values in this case were larger than those for Case 1, because the system did not admit an internally positive system in contrast to the system in Case 1. However, the dominant eigenvalue of $\tilde{A}$ nearly coincided with those of A generated using Algorithms 1 and 3. That is, from the perspective of the eigenvalues, we could construct an internally positive system near the original system.

Remark 5: We also generated a number of random $(\tilde{A}, \tilde{B}, \tilde{C})$ which were not externally positive. These cases occur frequently when we perform step 1, as mentioned in Remark 4. In these cases, Algorithm 3 generated better solutions than Algorithm 1. However, similar to in Case 2, the sequence generated by Algorithm 3 oscillated when c_W and c_U

were small. In contrast, the sequence generated by Algorithm 1 always decreased monotonically.

VI. Conclusion

We developed PALM and PALM-like algorithms for determining the nearest discrete-time linear positive system for a given system discrete-time system, with the same order as that of the considered system. For the PALM algorithm, the global convergence property to a critical point of the considered objective function was guaranteed by using the KL and semi-algebraic properties. Through numerical experiments, we demonstrated the effectiveness of the proposed algorithms.

The following issues must be addressed in future studies.

1) The analysis of the convergence rate of the PALM algorithm can be reduced to studying the KL exponent in several cases [31], [33]. Because $\tilde{\alpha}(W, U)$ in Problem 3 is semi-algebraic, the function has a KL exponent in $[0, 1)$. However, the determination of the KL exponent is considerably difficult because of the term $\langle A, U^\top \tilde{A} U\rangle$ in $\tilde{\alpha}$. In fact, although a few results pertaining to these determination have been reported, as in [36], these results cannot be used due to the existence of the term. Consequently, the analysis of the convergence rate of Algorithm 1 would be investigated in future work.
2) Although we defined η_k and μ_k in Algorithm 3 in a similar manner as in Algorithm 1 in Section V, the sequence generated using Algorithm 3 oscillated in contrast to that of Algorithm 1. The development of a method for setting η_k and μ_k in Algorithm 3 such that a monotonically decreasing sequence is obtained would be considered in future work.
3) The findings for Case 2, as reported in Section V, demonstrated that there may be an internally positive system near a given externally positive system. This leads to the conjecture that the set of all internally positive systems is a dense subset of the set of all externally positive systems in a sense. The analysis of the conjecture would be considered in future work.

References

[1] D. S. Bernstein and D. C. Hyland, "Compartmental modeling and second-moment analysis of state space systems," *SIAM J. Matrix Anal. Appl.*, vol. 14, no. 3, pp. 880–901, 1993.
[2] E. Hernandez-Vargas, P. Colaneri, R. Middleton, and F. Blanchini, "Discrete-time control for switched positive systems with application to mitigating viral escape," *Int. J. Robust Nonlin. Control*, vol. 21, no. 10, pp. 1093–1111, 2011.
[3] R. Shorten, F. Wirth, and D. Leith, "A positive systems model of TCP-like congestion control: Asymptotic results," *IEEE/ACM Trans. Netw.*, vol. 14, no. 3, pp. 616–629, Jun. 2006.
[4] R. Shorten, C. King, F. Wirth, and D. Leith, "Modelling TCP congestion control dynamics in drop-tail environments," *Automatica*, vol. 43, no. 3, pp. 441–449, 2007.
[5] Y. Ebihara, D. Peaucelle, and D. Arzelier, "Analysis and synthesis of interconnected positive systems," *IEEE Trans. Control Netw. Syst.*, vol. 62, no. 2, pp. 652–667, Feb. 2017.
[6] H. Ichihara, S. Kajihara, Y. Ebihara, and D. Peaucelle, "Formation control of mobile robots based on interconnected positive systems," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8441–8446, 2017.
[7] W. M. Haddad, V. Chellaboina, and Q. Hui, *Nonnegative and Compartmental Dynamical Systems*. Princeton, NJ, USA: Princeton Univ. Press, 2010.
[8] D. Angeli and E. D. Sontag, "Monotone control systems," *IEEE Trans. Control Netw. Syst.*, vol. 48, no. 10, pp. 1684–1698, Oct. 2003.
[9] F. Forni and R. Sepulchre, "Differentially positive systems," *IEEE Trans. Control Netw. Syst.*, vol. 61, no. 2, pp. 346–359, Feb. 2016.
[10] H. Maeda and S. Kodama, "Positive realization of difference equations," *IEEE Trans. Circuits Syst.*, vol. 28, no. 1, pp. 39–47, Jan. 1981.
[11] Y. Ohta, H. Maeda, and S. Kodama, "Reachability, observability, and realizability of continuous-time positive systems," *SIAM J. Control Optim.*, vol. 22, no. 2, pp. 171–180, 1984.
[12] L. Caccetta and V. G. Rumchev, "A survey of reachability and controllability for positive linear systems," *Ann. Oper. Res.*, vol. 98, nos. 1–4, pp. 101–122, 2000.
[13] M. E. Valcher, "Controllability and reachability criteria for discrete time positive systems," *Int. J. Control*, vol. 65, no. 3, pp. 511–536, 1996.
[14] H. M. H'ardin and J. H. van Schuppen, "Observers for linear positive systems," *Linear Algebra Appl.*, vol. 425, nos. 2–3, pp. 571–607, 2007.
[15] C. Briat, "Robust stability and stabilization of uncertain linear positive systems via integral linear constraints: L_1-gain and L_∞-gain characterization," *Int. J. Robust Nonlin. Control*, vol. 23, no. 17, pp. 1932–1954, 2013.
[16] M. Colombino and R. S. Smith, "A convex characterization of robust stability for positive and positively dominated linear systems," *IEEE Trans. Control Netw. Syst.*, vol. 61, no. 7, pp. 1965–1971, Jul. 2016.
[17] T. Tanaka and C. Langbort, "The bounded real lemma for internally positive systems and H-infinity structured static state feedback," *IEEE Trans. Control Netw. Syst.*, vol. 56, no. 9, pp. 2218–2223, Sep. 2011.
[18] N. K. Dhingra, M. Colombino, and M. R. Jovanović, "Structured decentralized control of positive systems with applications to combination drug therapy and leader selection in directed networks," *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 1, pp. 352–362, Mar. 2019.
[19] A. Rantzer, "Scalable control of positive systems," *Eur. J. Control*, vol. 24, pp. 72–80, Jul. 2015.
[20] C. Grussler, J. Umenberger, and I. R. Manchester, "Identification of externally positive systems," in *Proc. IEEE 56th Annu. Conf. Decis. Control (CDC)*, Melbourne, VIC, Australia, 2017, pp. 6549–6554.
[21] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*. Philadelphia, PA, USA: SIAM, 2016.
[22] K. Sato, "Optimal graph Laplacian," *Automatica*, vol. 103, pp. 374–378, May 2019.
[23] T. Katayama, *Subspace Methods for System Identification*. London, U.K.: Springer, 2006.
[24] W. E. Larimore, "Canonical variate analysis in identification, filtering, and adaptive control," in *Proc. 29th IEEE Conf. Decis. Control*, Honolulu, HI, USA, 1990, pp. 596–604.
[25] P. Van Overschee and B. De Moor, "N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica*, vol. 30, no. 1, pp. 75–93, 1994.
[26] M. Verhaegen and P. Dewilde, "Subspace model identification part 1. The output-error state-space model identification class of algorithms," *Int. J. Control*, vol. 56, no. 5, pp. 1187–1210, 1992.
[27] M. Verhaegen and A. Hansson, "N2SID: Nuclear norm subspace identification of innovation models," *Automatica*, vol. 72, pp. 57–63, Oct. 2016.
[28] L. Farina and S. Rinaldi, *Positive Linear Systems: Theory and Applications*. New York, NY, USA: Wiley, 2011.
[29] J. Umenberger and I. R. Manchester, "Scalable identification of stable positive systems," in *Proc. IEEE 55th Conf. Decis. Control (CDC)*, Las Vegas, NV, USA, 2016, pp. 4630–4635.
[30] J. Anderson, "Distance to the nearest stable Metzler matrix," in *Proc. IEEE 56th Conf. Decis. Control*, Melbourne, VIC, Australia, 2017, pp. 6567–6572.
[31] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *Math. Program.*, vol. 146, nos. 1–2, pp. 459–494, 2014.
[32] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: Jones Hopkins Univ. Press, 1996.
[33] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, "Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka–Łojasiewicz inequality," *Math. Oper. Res.*, vol. 35, no. 2, pp. 438–457, 2010.
[34] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ, USA: Princeton Univ. Press, 2008.
[35] J. M. Lee, *Introduction to Smooth Manifolds*. New York, NY, USA: Springer, 2013.
[36] G. Li and T. K. Pong, "Calculus of the exponent of Kurdyka–Łojasiewicz inequality and its applications to linear convergence of first-order methods," *Found. Comput. Math.*, vol. 18, no. 5, pp. 1199–1232, 2018.